Pankaj Shrivastava
Hirak Ranjan Dash
Jose A. Lorente
Jahangir Imam *Editors*

Forensic DNA Typing: Principles, Applications and Advancements

Forensic DNA Typing: Principles, Applications and Advancements

Pankaj Shrivastava • Hirak Ranjan Dash •
Jose A. Lorente • Jahangir Imam
Editors

Forensic DNA Typing: Principles, Applications and Advancements

Editors
Pankaj Shrivastava
DNA Fingerprinting Unit, State Forensic Science Laboratory, Department of Home (Police)
Government of Madhya Pradesh
Sagar, Madhya Pradesh, India

Hirak Ranjan Dash
DNA Fingerprinting Unit
Regional Forensic Science Laboratory
Bhopal, Madhya Pradesh, India

Jose A. Lorente
Department of Legal Medicine
University of Granada
Granada, Granada, Spain

Jahangir Imam
DNA Fingerprinting Unit
State Forensic Science Laboratory
Ranchi, Jharkhand, India

ISBN 978-981-15-6654-7 ISBN 978-981-15-6655-4 (eBook)
https://doi.org/10.1007/978-981-15-6655-4

© Springer Nature Singapore Pte Ltd. 2020
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Contents

Part I Fundamentals of Forensic DNA Typing

1 Forensic DNA Typing: Inception, Methodology, and Technical Advancements 3
Kamayani Vajpayee, D. C. Sagar, and Hirak Ranjan Dash

2 STR Typing and Available Multiplex Kits Including Validation Methods 27
Anna Barbaro

3 Sequential Advancements of DNA Profiling: An Overview of Complete Arena 45
Kriti Nigam, Ankit Srivastava, Subhasish Sahoo, I. P. Dubey, I. P. Tripathi, and Pankaj Shrivastava

4 Forensic DNA Evidence: From Crime Scene to Conviction 69
Ashish Badiye, Neeti Kapoor, and Pankaj Shrivastava

5 RNA- and DNA-Based Identification of Body Fluids 87
Arjun Rao, Manisha Rana, Ashish Pradhan, and Moumita Sinha

6 Statistical Softwares Used in Evaluation of Forensic DNA Typing 105
R. K. Kumawat, Aditi Mishra, and Pankaj Shrivastava

Part II Challenges in Forensic DNA Typing

7 Ancient DNA Analysis and Its Relevance in Forensic DNA Fingerprinting 137
Mian Sahib Zar and M. Aslamkhan

8 Analyses of Second World War Skeletal Remains Using a Forensic Approach 153
Irena Zupanič Pajnič

9 Molecular Tools for Analysis of Archaeological and Prehistoric Human Bones: A Perspective of Anthropological and Forensic Relevance 181
Muktikanta Panda, Awdhesh Narayan Sharma, Pankaj Shrivastava, and Hirak Ranjan Dash

10 Usefulness of Mini-STRs in Analyzing Degraded DNA Samples and Their Forensic Relevance 205
Mukesh Thakar, Bhawana Joshi, and Pankaj Shrivastava

11 Capillary Electrophoresis Issues in Forensic DNA Typing 223
Shuchi Kaushik and Vivek Sahajpal

12 Human Trafficking and DNA Analysis 239
María Saiz, Maria Jesus Alvarez-Cubero, Begona Lorente-Remon, Luis Javier Martínez-Gonzalez, and Jose A. Lorente

Part III Case Studies/Applications in Forensic DNA Typing

13 Autosomal STR Typing and Case Studies 257
Pankaj Shrivastava, Shivani Dixit, Varsha Srivastava, Hirak Ranjan Dash, and R. K. Kumawat

14 Y Chromosome Short Tandem Repeats Typing 277
Rashed Alghafri

15 X-STRs: Potentials and Applications 301
Sonia Kakkar, R. K. Kumawat, and Pankaj Shrivastava

16 Applications of Mitochondrial DNA in Forensic Science 329
Moumita Sinha, Manisha Rana, and Pushpesh Kushwaha

17 SNP in Forensic DNA Testing 345
Rohit Saluja, Pankaj Shrivastava, and Swapnil Sinha

18 SNP Testing in Forensic Science 365
Moumita Sinha and I. Arjun Rao

Part IV Non-human Forensic DNA Typing

19 DNA Analysis of Domestic Animals 379
Kamlesh Kaitholia, Pushpesh Kushwaha, Manisha Rana, Ila Gautam, and Ankit Srivastava

20 DNA Forensics in Combating Illegal Wildlife Trade: Present, Past, and Future Perspectives 399
Gul Jabin, Vivek Sahajpal, Kailash Chandra, and Mukesh Thakur

21 The Utility of DNA Barcoding Technology in the Authentication of Medicinal Plants in Illegal Trade: A Critical Review 405
Mukesh Kumar Thakar and Tina Sharma

22 DNA Barcoding in Forensic Mycology: Concepts, Limitations, and Future Prospects 431
Spriha Sharma, Jaskirandeep Kaur Jossan, and Rajinder Singh

Part V Advances in Forensic DNA Typing

23 Applications of Next-Generation Sequencing in Forensic Field 451
María Saiz, Luis Javier Martínez-Gonzalez, Jose Antonio Lorente, and Maria Jesus Alvarez-Cubero

24 Utility and Possibility of Next-Generation Sequencing in Forensic DNA Typing 473
Vijay Nema

25 Oral Microbes: A Hidden Yet Powerful Evidence for Futuristic Forensic Investigation 497
Bhargav C. Patel and Mayuri Ganguly Soni

26 MALDITOF the Fourth Generation Techniques Still at Its Infancy to Identify Forensically Important Insects 519
Vandana Vinayak and Anshuman Rai

27 Forensic DNA Phenotyping 547
Dolly Mahna, Meenu Dutt, Sanjeev Puri, Tanzeer Kaur, and Shweta Sharma

28 Rapid DNA Typing 561
Pankaj Shrivastava, Aditi Mishra, Anand Kumar, Sumit K. Chaudhary, Sonia Kakkar, and R. K. Kumawat

Part VI Guideline, Quality Control, Legal Aspects and DNA Database in Forensic DNA Typing

29 Guidelines for Collection and Preservation of Samples for Forensic DNA Testing 573
Pankaj Shrivastava, Shivani Dixit, and Hirak Ranjan Dash

30 Quality Control in Forensic DNA Typing 585
Ashish Badiye, Neeti Kapoor, Prachi Kathane, and Ritesh K. Shukla

31 Legal Aspects of Forensic DNA Typing 607
Ankit Srivastava, Abhimanyu Harshey, and Pankaj Shrivastava

32 DNA Databases 629
Anna Barbaro

33 Building of the World's Largest DNA Database: The China Case . . . 639
Ausma Bernotaite

34 DNA Databases: Risks, Benefits, Privacy, and Human Rights 659
Sachil Kumar

Appendix A: SWGDAM Guidelines (Web Link) 679

Appendix B: International Society for Forensic Genetics (ISFG) 681

Appendix C: Forensic Science Journals and Publications 685

About the Editors

Pankaj Shrivastava received his Ph.D. in Microbiology with a specialization in Biotechnology from the Department of Biological Science, Rani Durgawati University, Jabalpur, Madhya Pradesh, India. He is working as In charge of the Forensic DNA Fingerprinting facility of Govt. of Madhya Pradesh, India at Sagar. He has more than 13 years of experience as the bench worker for examining and reporting a wide range of criminal cases using DNA technology and deposing the court evidence. His research interests include population DNA database, improvement of methods in forensic DNA typing, finding rapid protocols for the technology, and microbial forensics. He is the recipient of the prestigious Pt. Govind Vallabh Pant Award from the Bureau of Police Research and Development, Ministry of Home Affairs, Govt. of India twice, and Anusrajan Award from AIISECT University, Bhopal for his authored books. He is also a recipient of the FICCI smart Policing award for developing a fast DNA typing protocol. He is visiting faculty of many universities, Police training Institutes, and Judicial officers training Institutes. He has written 13 books and published 65 research papers in the journals of repute in forensic science, 21 book chapters, 84 conference proceedings, 15 popular science articles.

Hirak Ranjan Dash completed his Ph.D. from the Department of Life Science, National Institute of Technology, Rourkela, India, and is currently working as a DNA fingerprinting expert at the Forensic Science Laboratory, Bhopal, Madhya Pradesh, India. He received his M.Sc. in Microbiology from Orissa University of Agriculture and Technology, Odisha, India. His research interests include forensic microbiology, thanatomicrobiome analysis, molecular microbiology, DNA fingerprinting and genetic manipulation. He has written four books and published 29 research papers, 14 book chapters, 12 conference proceedings, and four popular science articles.

Jose A. Lorente is a Professor of Forensic Medicine at the University Of Granada, Spain. After graduating at the Faculty of Medicine of Granada in 1985, he received Ph.D. Honors in Medicine and Surgery at the University of Granada, in 1989. Dr. Lorente has published over 140 peer-reviewed papers and several books and book chapters. He has a special focus on the use of DNA and its application to

human rights (he launched the first-ever database to identify missing people, the Spanish Phoenix Program, back in 1999); he also created and launched the DNA-PROKIDS Program in 2004. His areas of interest in forensics also deal with population variability and analysis of old and ancient DNA samples and databases expansion and control. Dr. Lorente is also actively working in medical genomics and he’s the scientific director of the Center for Genomics and Oncological Research (GENYO), where his team focuses on liquid biopsy and cancer interception. Dr. Lorente is the founder and first president of the AICEF (Ibero-Latin American Network of Forensic Sciences), and an honorary member of the AFSN (Asian Forensic Sciences Network).

Jahangir Imam completed his Ph.D. from the Department of Microbiology, Maharshi Dayanand University, Rohtak, India, in Collaboration with the National Rice Research Institute (ICAR), Cuttack, India, and is currently working as Forensic DNA expert at the Forensic Science Laboratory, Jharkhand, India. He received his M.Sc. in Biotechnology from Jiwaji University, Gwalior, India. His research interests include forensic biology, molecular marker, DNA fingerprinting and genetic manipulation. He has written 16 research papers, 11 book chapters, eight conference proceedings, and two popular science articles.

Part I

Fundamentals of Forensic DNA Typing

Forensic DNA Typing: Inception, Methodology, and Technical Advancements

1

Kamayani Vajpayee, D. C. Sagar, and Hirak Ranjan Dash

Abstract

Forensic DNA typing has overpowered all other known methods in forensics. Starting with the method which required a reasonably high amount of good quality of DNA, time taking and a complex process, this little more than three decade old technology has gone through much advancement to overcome these issues. The chapter details different steps involved in forensic DNA typing from its inception to the present including the advancements.

Keywords

Forensic DNA · STR · PCR · RFLP · Capillary electrophoresis

1.1 Introduction

In the 1891 novel, Sherlock Holmes, a fictional private detective said "it has long been an axiom of mine that the little things are infinitely the most important" (Doyle and Roden 1993). Certainly he would have never imagined that a thing as little as a DNA molecule could become a powerful tool aiding the criminal justice system (Jobling and Gill 2004). Gradually, the field of forensic genetics developed to solve the legal issues like identification of the victims of mass disasters and

K. Vajpayee
Department of Forensic Science and Anthropology, University of Lucknow, Lucknow, Uttar Pradesh, India

D. C. Sagar
Disaster Management, Police Head Quarters Bhopal, Bhopal, Madhya Pradesh, India

H. R. Dash (✉)
DNA Fingerprinting Unit, Integrated High-Tech Laboratory, Regional Forensic Science Laboratory, Bhopal, Madhya Pradesh, India

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_1

individualization of the culprits of the crime through analysis of forensic DNA (Giardina et al. 2011).

Earlier forensic scientists around the world were using the serological markers like RBC antigens, enzymes and polymorphic serum proteins for solving the forensic genetic case works using immunological assays and electrophoresis (Martin et al. 2001). The serological markers however did not prove to be a helping hand while working with the forensic evidences like hair, semen, saliva, etc. The evidences collected from a crime scene are limited in quantity and may be degraded, so only handful of the information could be deciphered from them. In addition to this it went difficult while accessing the biological evidences other than blood like hair and semen as in the cases of sexual assault. Thus, the reports of forensic scientists lacked sufficient amount of information that could help them reach a conclusion (Carracedo and Sanchez-Diz 2005).

In 1985, while studying the genetic diseases in the lab of Leicester University, Sir Allec Jeffery and his coworkers experienced “eureka” moment which has opened a new gateway in the field of genetics. The technology was later named as genetic profiling, DNA typing, or DNA fingerprinting. He observed polymorphism in repetitive segment of genome using multi locus probes (DNA markers) (Roewer 2013). The DNA profiling technology was then tested for the very first time in a forensic casework of the two young girls who were raped and murdered. The suspect refused to accept his involvement in the murder of the second girl. The police then chose to genetically profile the entire male population of the area and to match them with the profile generated from the materials collected from crime scene. Unfortunately, no match was found. Later Colin Pitchfork was overheard of saying how he escaped this exercise by police. There the case got solved. With the advent of these markers (DNA markers), a great sense of excitement and enthusiasm among the forensic experts was felt. These technologies lead to a deep insight into the forensic materials since DNA can be extracted from any biological evidence collected like hair, semen, saliva, etc. Moreover, the technology could even access the degraded material which serological markers failed to provide (Carracedo and Sanchez-Diz 2005). Thus, DNA profiling established itself as a gold standard in forensic science.

Mid of twentieth century seen several technological advancements relating to the genetic markers which is still continuing till date. During inception of this technology, restriction fragment length polymorphism (RFLP) was used as a pilot method which uses hybridized DNA probes—multi locus and single locus DNA probes.

1.1.1 Restriction Fragment Length Polymorphism (RFLP)

In this method, the DNA molecule is treated with restriction enzymes. Restriction enzymes have a characteristic property of cutting the DNA at specific base sequences, the restriction sites. This results in the formation of DNA fragments of variable length. Further, agarose gel electrophoresis is carried out to separate such variable length DNA fragments followed by southern blotting. The fragments get hybridized to short radio labeled probes and can be easily detected by radiography as

bands. The loss or absence of such few restriction sites forms the basis of individuality in humans. Thus, following this technique, a unique RFLP signature for an individual can be generated. Forensic experts match such signatures originating from different sources with the one lifted from the crime scene. The DNA probes most importantly the single locus probes (SLPs) were routinely in this technique for analyzing the paternity dispute cases.

By the mid of 1980s, variable number of tandem repeats (VNTR) mini satellite loci became the common method for analysis of DNA. VNTR are the sequences of few nucleotides commonly called as motifs which are repeated in variable number of times in every DNA molecule. This variation in number of repeats creates a scope of individualization among organisms. The mini satellites are the type of VNTR sequences consisting of 9–100 bp sequences. Minisatellites which were widely in forensic routine was D1S80 (pMCT118) and other probes like 33.15, 33.6, pentameric CAC/GTG were also in use in countries like Germany, the USA, and the UK.

For so long, the RFLP technology proved to be a reliable tool in DNA analysis but it had few limitations which lead the research community for further discovery of other markers. The limitations posed were the need of large quantity of non-degraded DNA, cost-effectiveness, involvement of huge time and labor, and requirement of much statistical work for analysis. These limitations were then overcome by PCR-based markers (Primorac et al. 2000; Panneerchelvam and Norazmi 2003; Roewer 2013; Grover and Sharma 2014).

1.1.2 PCR and Related Markers

The polymerase chain reaction (PCR) technique was invented by Kary Mullis in 1985. It allowed the in vitro amplification of DNA sequences using extension of the complementary primers. The entire process or a single cycle involved—denaturation of DNA strands caused by high temperature, annealing of complementary primer and their extension. Since the temperature during the entire process remains very high which is crucial for denaturation process to occur; a special type of DNA polymerase, the Taq polymerase aids in extending DNA primers. It takes 25–30 cycles for a DNA segment to be amplified for precise analysis.

Earlier PCR-based marker systems used single nucleotide polymorphisms in HLA-DQA1 GENE7. This system was based on detection of variations in DNA using probes. But the discrimination power was low. New set of PCR markers were the need of an hour. In the mid-1990s, the forensic community started using microsatellites repeats. The microsatellite repeats are a class of VNTRs consisting of 2–9 bp. The microsatellite VNTRs also called as short tandem repeats (STRs) proved to have enormous discriminating power and thus, ideally suited for forensic caseworks (Budowle et al. 1990; Panneerchelvam and Norazmi 2003; Carracedo and Sanchez-Diz 2005).

1.2 Forensic Application

Since the advancements in DNA typing technology, the forensic materials can now be easily assessed for individuality. The first use of DNA fingerprinting in forensic case work was witnessed in 1986. Sir Allec Jeffreys used the technique to verify the suspect on a double rape and murder case. Robert Melias was the first suspect to be identified and convicted by DNA fingerprint technology in 1987. The famous cases of Glen Woodal v/s State of Virginia in 1992 and the multiple murder trial of Timothy Wilson Spencer versus the state of Virginia in 1994 gave thrust to the use of DNA fingerprinting in solving the forensic cases (Panneerchelvam and Norazmi 2003).

In the present scenario the DNA technology is been used in both civil and criminal cases. During a disaster or a crime the DNA evidence serves as corroborative evidence and aids in identification of the victim and criminal as the case maybe. Meanwhile the DNA profiling also helps in civil disputes related to paternity and kinship determination. With the advancement in the field of DNA typing methodologies and its analysis DNA is now serving the criminal justice system while solving the caseworks. In India several cases of kinship, paternity, murder, and rape have now been solved with the help of DNA profiling technique (Verma and Goswami 2014). Few of them are:

- ***Rajiv Gandhi Assassination case***: Mr. Rajiv Gandhi, the ex-prime minister of India was assassinated in 1991 using human bomb. The flesh collected from the crime scene (The belt) was handed over to the forensic team who using DNA typing methods proved the identity of the individuals involved. Similarly the case associated with the assassination of Mr. Beant Singh, the Chief Minister of Punjab, was solved using DNA profiling technology.
- ***Naina Sahani Tandoor case***: The famous “Tandoor case” of India was also solved using DNA typing methodology. Naina Sahani was a young lady. She was murdered and the body was thrown in the “tandoor” oven. The charred body of the lady was recovered by the police and the flesh and bones were analyzed for DNA testing. The analysis proved that the body belonged to Naina Sahani.
- ***Santosh Kumar Singh vs. State***: A young student named Ms. Priyadarshini Mattoo was murdered by her friend in 2006. Since there was not any evidence that could prove the culprit guilty, he was acquitted from the trail court. The blood samples collected from the item found at crime scene and semen collected from the lady’s undergarments was sent for DNA analysis. The High court on the basis of the results of DNA profiling held the accused guilty of Ms. Mattoo’s murder. The decision was also upheld by the apex court of India.
- ***Gautam Kundu case***: The Gautam Kundu case of paternity issue is amongst the important cases since it lead to the formulation of important guidelines explaining the purpose and circumstances by Hon’ Supreme of India. The guidelines included: “(1) that courts in India cannot order blood test as matter of course; (2) wherever applications are made for such prayers in order to have roving inquiry, the prayer for blood test cannot be entertained; (3) There must be a strong

prima facie case in that the husband must establish non-access in order to dispel the presumption arising under section 112 of the Evidence Act, 1872; (4) The court must carefully examine as to what would be the consequence of ordering the blood test; whether it will have the effect of branding a child as a bastard and the mother as an unchaste woman; and (5) No one can be compelled to give sample of blood for analysis." These guidelines were asked to be undertaken when DNA Evidence Acts were not framed (Panneerchelvam and Norazmi 2003).
- ***N. D. Tiwari case***: Rohit Shekhar born in 1979 claimed himself as the putative son of Mr. N. D. Tiwari. Although his mother married another man in 1963, she opted for divorce in 2006. As per the section 112 of the Indian Evidence Act, 1872, paternity can be decided but the court ordered for DNA test, a deviation from the guidelines posed in Gautam Kundu case. The DNA analysis indicated that N. D. Tiwari is the biological father of Rohit shekhar. Both parties agreed upon settlement outside the court but the questions related to distribution of rights and duties remained unanswered.
- ***The Black Buck Poaching Case of Jodhpur***: It was in 1998 while shooting for a movie in Jodhpur (India), a very famous cine star got accused of hunting the black buck (endangered species). The samples of meat were analyzed and the DNA study revealed the species of the animal. Further the number of animals hunted was also established.

1.3 Methodology

1.3.1 STR-Based Technique

While working with the SLP and PCR methods scientists found novel method involving variation within these minisatellites. The method was then termed as MVR-PCR (minisatellite variant repeat). The technology, however, was complicated. The analysis of mixed samples proved to be a tough task. These altogether lead to research for new improved methods which could have a better discriminating power and on the other hand are easy to analyze.

In 1992, STRs came into existence and were used to analyze forensic samples. These STRs are microsatellite VNTRS having 2–9 nucleotides repeats. Soon they became the basis of DNA analysis. The STRs, as regards to their structure, are either complex STRs or simple STRs. The complex STRs show the property of hypervariability and these are the loci which contains one or more than one run of repeats having more than one repeat type. Whereas simple STR are the loci consists of uninterrupted run of one repeat type. In simple STRs the rate of mutation remains low and can be standardized easily. The STRs are thus selected on the basis of certain criteria like robustness, size, and presence of artifacts for forensic use. The STRs being small in size tend to remain stable which had simplified the analysis of degraded age old DNA samples (Jobling and Gill 2004; Carracedo and Sanchez-Diz 2005).

The STRs are polymorphic in nature. For each STR locus, the STR alleles are found in variable number of repeat sequences, giving STRs a high discriminating power. The experts hence agreed upon using more STR loci which will increase the probability of presence of exact same number of repeat units amongst individuals to almost rare.

The STRs alleles for a particular locus show codominant behavior. An individual thus inherits one allele from mother and another from father. This forms the basis of inclusion in paternity case works. In such cases, the STR alleles of the disputed person are matched against the STR alleles of alleged father and mother and a perfect match indicates the inclusion whereas a mismatch excludes them. Thus the discovery of STRs proved to be a boon for forensic experts (Martin et al. 2001; Panneerchelvam and Norazmi 2003).

The forensic groups from different countries started using various techniques for STR typing. Earlier denaturing electrophoresis with native gel was used for the analysis of STRs. The electrophoretic method of analysis was somehow prone to errors thus hampers the analysis. Recently, multiplex technique for STR typing is developed which uses primers probed with fluorescent dyes and DNA sequencers. The technique allowed amplification of STR loci while separating them. Today many commercial multiplexes are available in the market like SGM Plus (Applied Biosystems), Promega (Madison, WI), Poweplex16 (Promega) and the Identifiler (Applied Biosystems) (Martin et al. 2001).

1.3.2 Types of STR Markers

1.3.2.1 Sex Determining Markers

Determination of the sex from the biological evidence found from the crime scene is a routine process in forensic laboratories. There are several methods of sex determination from the skeletal remains like skull, pelvic girdle, long bones, etc. But finding a complete skull or a complete long bone with all the features leading to the conclusion is impossible. To rescue, DNA typing technology emerges out as a simple, rapid, and sophisticated method. The STR markers for sex determination are found in Y and X chromosomes of the individual. These are called as Amelogenin marker. Today, every commercial DNA typing kit includes Amelogenin marker so as to determine sex in cases of sexual assault or mixed samples. The gene Amelogenin produces several mRNAs which are processed in various manners to generate a group of proteins. The proteins are responsible for tooth enamel formation. In humans, this gene is found in X as well as in Y chromosomes at p22.1-22.3 region and at Yp11.2 region, respectively. These genes are subsequently called as AMELX and AMELY gene. The mutation in the gene effects the enamel production (Amelogenesis imperfect). The Amelogenin protein is synthesized maximally by AMELX. The two genes are homologous (89%) and have highly conserved sequences between 2 and 6 exons. Thus, homogeneity and presence of distinguishing features they serve as reliable sex determining markers. The AMELX gene has a 6 bp deletion. Since females possess only X

chromosome only a single fragment is observed whereas one fragments each of X and Y is observed. This forms the basis of determination and differentiation between male- and female-originated samples.

The Amelogenin being a reliable marker for sex determination also have few drawbacks. Issues related with false positive results due to allele dropout, effects of PCR inhibitors and failure to process degraded samples are being faced by experts on daily basis (Dash et al. 2020).

1.3.2.2 Autosomal STR Markers

Autosomal STR markers show wide variation due to mutations, independent assortment and recombination. Initially two STR loci, THO1 and vWA were in use. The US federal bureau of Investigation (FBI) soon formed an Index system called Combined DNA Index System or CODIS which included use of 13 STR loci for investigation purpose. These were: CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, and D21S11. Lately to increase the usefulness of DNA fingerprinting technique seven ethnic group specific markers were also added to the list. The newly added STR alleles were: D1S1656, D2S441, D2S1338, D10S1248, D12S391, D19S433, and D22S1045 (Dash et al. 2019). The STR alleles included in the CODIS system are tetranucleotide repeats except D22S1045 which is a trinucleotide repeat (Dash et al. 2019).

1.3.2.3 Y chromosome STR Markers

Human males exclusively have a haploid Y chromosome. This Y chromosome is extremely helpful in solving forensics case works. The specific regions of the Y chromosome can be looked for the features which can possibly differentiate them with female counterparts. Such set of microsatellites found on Y chromosome are collectively called as YSTRs. Since these are male specific they help in analysis of the mixed samples-originating from both male and females in the case of rape, sexual assault, or otherwise in the case of mass disaster. The Y STR typing in above said cases can reveal about the male component. The Y STRs aids in better analysis of degraded samples. Further, if the sample comes from an azospermic male, the Y STR typing can establish his identity. Thus, the discriminatory power of Y STRs is no less than Autosomal STRs.

Autosomal STRs are variable in nature as they are inherited following the law of independent assortment, crossing over and recombination, whereas, Y STRs are only prone to mutation. Hence, are less diverse in nature. The set of Y STRs that are routinely practiced in labs are: DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS385ab, DYS438, DYS439, DYS437, DYS448, DYS456, DYS458, DYS635, and YGATAH4. YSTRs also have a drawback of having patrilineal occurrence. In case of a match, the whole male lineage of the family comes under question. But such patrilineal characteristic on other hand could aid in individualizing the missing person or the victim of a mass disaster (Jobling and Gill 2004; Giardina et al. 2011).

1.3.2.4 X-STR Analysis

In certain complex cases such as kinship analysis as well as cases with limitations of reference samples, autosomal STR analysis alone cannot give a definitive conclusion. In such a scenario, analysis of STRs located on sex chromosomes becomes imperative. Thus, analysis of X-STRs plays a significant role in forensic DNA typing. As men are hemizygous and women are dizygous, men receive single X from their mother. This unique inheritance pattern of X chromosome makes its suitability in deciphering the deficient paternity cases. For a disputed girl child, X-STR analysis between the father and daughter can be performed. However, it cannot be performed for a male child in paternity dispute case due to non-inheritance of X chromosome from father. The analysis of X-STR in supplement with autosomal STR analysis increases the discrimination power besides affecting the paternity exclusion probability. It is useful in the analysis of complex kinship cases such as paternal half-sisters, paternal aunt/uncle-niece, and maternal uncle-nephew. X-STR analysis is highly useful in incestuous relationships involving grandparent, half-sibling, and uncle/aunt by distinguishing between such relationships. However, the major limitation of X-STR analysis is that as all X-STR markers represent on same X chromosome, it leads to linkage disequilibrium among them due to linkage. Additionally, the currently available X-STR kits harbor a limited number of markers and generate a large amplicon size.

1.3.3 mtDNA Analysis

The mitochondrion is a semi-autonomous organelle found in every cell type (except the RBCs) in the organism. They have their own circular DNA molecule called mitochondrial DNA (mtDNA). The mtDNA consists of 16,569 nucleotide pairs and acts as an important tool for forensic scientists. The mtDNA, similar to Y chromosome, is a lineage molecule. It is maternally inherited and is found in high copy number within the cell. The copy number ranges from 100 to 10,000 per cell. Mitochondrial DNA resists the phenomenon of independent assortment, recombination, and crossing over. The only process which lies behind the variation in mtDNA is high rate of mutation, making the mtDNA molecule less diverse in nature. In mtDNA molecule, the variation among the individual is seen in the hypervariable regions present in the control region or the D loop. The nucleotide repeats in these hypervariable regions are highly polymorphic. The hypervariable regions are found in two segments denoted as HV-I and HV-II. The HV-I region also called as low resolution segment ranges from 16,024 to 16,383 bp whereas the HV-II segment is called as high resolution segment and is numbered from 57 to 372. Thus, the molecule has high discriminating power.

The mtDNA typing is considered beneficial in the cases where long aged DNA samples are to be analyzed or the sample is degraded to such an extent that nuclear genome (nuDNA) cannot be analyzed. The mtDNA is present in high copy number and hence the chance of recovery surpasses to that of nuclear DNA (nuDNA). MtDNA analysis also aids in the interpretation of samples like hair shafts or

Table 1.1 List of databases useful for mtDNA analysis

S. No.	Mitochondrial DNA websites	Links
1	Armed Forces DNA Identification Laboratory (AFDIL)	http://www.afip.org/Departments/oafme/dna/
2	EMPOP Mitochondrial DNA Control Region Database	http://www.empop.org
3	FBI Laboratory DNA Unit II	http://www.fbi.gov/hq/lab/org/dnau.htm
4	Human Mitochondrial Protein Database	http://bioinfo.nist.gov/
5	Mitochondria Research Society	http://www.mitoresearch.org/
6	Mitochondrion (Journal of the Mitochondria Research Society)	http://www.sciencedirect.com/science/journal/15677249
7	Mitomap—A Human Mitochondrial Genome Database	http://www.mitomap.org/
8	Mitotyping Technologies	http://www.mitotyping.com
9	mtDB—Human Mitochondrial Genome Database	http://www.genpat.uu.se/mtDB/
10	mtDNA Manager	http://mtmanager.yonsei.ac.kr
11	SWGDAM mtDNA Database	http://www.fbi.gov/about-us/lab/forensic-science-communications/fsc/april2002/index.htm/miller1.htm
12	SWGDAM Guidelines for mtDNA Interpretation	http://www.swgdam.org

bones, where the nuDNA is fragmented upto such an extent so that its analysis becomes cumbersome. Cases of missing and mass disaster can be solved using mtDNA typing technique as they are maternally inherited. The websites useful for mitochondrial DNA analysis is listed in Table 1.1.

Though the mtDNA typing seems to be a beneficial tool in identification but like Y chromosome typing it too has got its limitations. The mtDNA typing procedure in a laboratory requires more skill and care with regards to the contamination. The molecule gets easily contaminated and would result in mis-analysis. Due to its matrilineal inheritance it is found present in all the members of matriline. The property of heteroplasmy also hinders the application of mtDNA typing. Moreover, the typing methodology is time taking and costly. Thus, making it a complex process and is only practiced when traditional nuDNA typing fails (Jobling and Gill 2004; Roewer 2013).

1.4 Advancements in DNA Typing Technology

DNA typing or the DNA profiling refers to the process of determining the sequence of nucleotides in the genome of an individual. The nucleotide sequence is analyzed for genetic variations among the individuals. The DNA typing involves isolation of DNA from the cell, quantification, amplification, and sequencing. With the advancement in the field of molecular biology and genetics, the tools to extract DNA

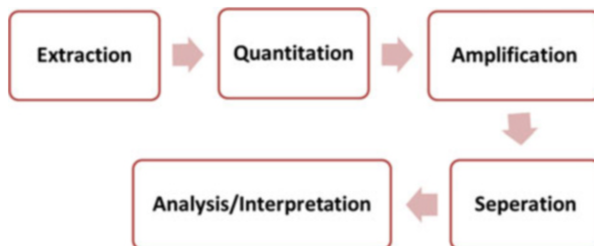


Fig. 1.1 Steps of routine DNA fingerprinting experiments

simultaneously became rapid, cost effective. Starting from the manual DNA extraction procedure used by Friedrich Miescher, the process modified several times by research groups around the globe. Now there are techniques available which extract sufficient amount of DNA even from tough biological samples like bone and teeth. These advancements were accompanied by modifications in amplification process by PCR. The PCRs now days, use high quality efficient polymerases which can also withstand high temperatures. Thus, provides a better typing methodology (Fig. 1.1).

1.4.1 Advancements in DNA Extraction Methodologies

The first step in DNA profiling is the DNA isolation/extraction process. The process involves sequential steps to isolate DNA from the cell/nucleated sample. Since DNA typing involves the analysis of DNA, the DNA extraction process should be such that enough and sufficient amount of DNA can be isolated and the chemicals involved in the process do not interact with the DNA so as to hamper its integrity. Techniques ideal to various types and forms of forensic samples have been developed. Forensic laboratories practice several manual methods to extract DNA automated techniques as per the requirement of the sample to be analyzed. The various techniques can be briefly described as under:

1.4.1.1 Manual Techniques

The manual technique of DNA extraction requires the physical extraction of DNA from the sample using combination of chemicals and enzymes. DNA can be extracted manually by using any of the two methodologies—organic extraction of DNA in phenol-chloroform solution or the salting out method given by miller in 1988.

Organic Extraction of DNA

The phenol-chloroform method of DNA isolation relies on the principle of liquid–liquid extraction of biomolecules. The DNA is isolated based on its solubility in aqueous phase, leaving the debris and other protein parts of the cell in organic phase.

The procedure involves cell lysis by digesting the cellular proteins and other biomolecules present with the help of sodium dodecylsulfate commonly known as SDS and proteinase K. To partition the solution into organic and aqueous phase a

mixture of phenol:chloroform:isoamyl alcohol in the ratio of 25:24:1 is added. This leads to the separation of DNA molecule which can be easily purified by centrifugation. Another method of purification could be by using chilled ethanol which precipitates the DNA.

This is a reliable method of DNA extraction as the double stranded DNA molecule can be recovered which is preferably best suited for RFLP method of DNA typing. Being an advantageous process it also have some of the limitations like, the method is time consuming, uses a lot of hazardous chemicals (McKiernan and Danielson 2017; Köchl et al. 2005).

Salting Out Method

Salting out method was proposed by Miller in 1988 as a method which outcomes the limitations of the phenol-chloroform method of DNA extraction. As per Miller, it is a safe and inexpensive method which has modified the de-proteinization part of DNA isolation. The method relies on the fact that the cellular proteins can be precipitated out using NaCl Solution. In this method the cell is lysed using lysis buffer which contains 10 mM Tris-HCl, 400 mM NaCl and 2 mM Na_2EDTA, pH 8.2. The cell lysates are then incubated at 37 °C with 0.2 mL of 10% SDS and 0.5 mL of a protease K solution (1 mg protease K in 1% SDS and 2 mM Na_2EDTA) for digestion. 1 mL of concentrated NaCl (6 M) is added to it followed by vigorous shaking for about 10–15 s. The above solution is taken for centrifugation (2500 rpm for 15 min). The supernatant is then transferred to another test tube and chilled ethanol is added to it for purification. The precipitated DNA strand can then be separated out using a pipette and is transferred to another test tube having 100–200 mL TE buffer (10 mM Tris-HCl, 0.2 mM Na_2EDTA, pH 7.5). This method thus allows the isolation of double stranded DNA without using the toxic reagents (Miller et al. 1988).

Few other researchers have modified the salting out procedure of DNA extraction either by using sodium hydroxide (Klintschar and Neuhuber 2000) or by using temperature resistant proteinase to denature proteins (Moss et al. 2003).

1.4.1.2 Automatic and Semiautomatic Techniques

Column-Based Approach

The column-based approach is one of the solid-phase extraction in which the DNA is adsorbed over the silica bed layered on to the glass beads using highly concentrated guanidine salts like guanidine hydrochloride, guanidine isothiocyanate, or sodium iodide. These salts tend to dehydrate the proteins resulting in stable binding of the DNA. After this, the low ionic solution with pH more than 7.5 is allowed to flow through the spin column which will efficiently elute the DNA out from the silica bed (Tereba et al. 2004; Duncan et al. 2003; Boom et al. 1990; Vogelstein and Gillespie 1979). Spin columns from Qiagen (QIAamp® spin columns) are available in the market which uses this methodology to separate the DNA from cell debris (McKiernan and Danielson 2017; Butler 2012a, b).

Magnetic Beads Based Approach

Magnetic beads based approach to DNA isolation is another type of solid-phase extraction. Unlike column based methods this approach uses magnets to isolate DNA. It uses the paramagnetic resin bed coated with silica and DNA is separated onto it. Products like DNA IQ system marketed by Promega Corporation, works on this format. The technique follows the binding of DNA molecule to magnetic beads in the presence of highly ionic solutions with low pH. The DNA bound magnetic beads are then drawn to the bottom of the test tube using a magnet. This in turn leaves all the impurities, cell debris, proteins, etc. in the solution. The magnetic beads are then washed several times so as to recover the isolated DNA efficiently. Another product from Applied Biosystems named PrepFiler has been designed to work on the similar fashion. It uses magnetic beads smaller to the ones used by the DNA IQ system. The size of beads enables the efficient DNA binding since they provide the large surface area.

The magnetic bead approach to DNA isolation is a rapid and simple methodology. The magnetic bead method has been automated and used by several forensic laboratories (Greenspoon et al. 2004). Some of the workstations automated for DNA IQ system are Beckman 2000 robot workstation, Maxwell 16 robot (Promega Corporation), and Tecan liquid-handling robots which provides high throughoutput. Prepfiler has also been validated (Barbaro et al. 2009; Brevnov et al. 2009) and Technologies like Tecan Freedom EVO automated liquid-handling workstation, AutoMate Express Forensic DNA Extraction System by Applied Biosystems and BTA (bone, teeth, adhesive) lysis buffer protocol have also been developed to extract DNA efficiently from tough samples (Butler 2012a, b).

The method though enables the rapid isolation of DNA, it also prevents and filters out the hazardous chemicals and makes the extraction process safe. One the other hand if any chemical or the substance hinders the binding of DNA to magnetic beads the whole process of DNA isolation fails. Further it does not involves transfer of solutions from a vial to another, reducing the chance of contamination (McKiernan and Danielson 2017; Butler 2012a, b; Tan and Yiap 2009).

FTA-Based Technique

Burgoyne in 1994 described the use of FTA paper to store DNA. The paper is made up of cellulose fibers and contains chemical which preserves DNA for longer period of time (Burgoyne 1996). The bactericidal enzyme in the paper prevents the growth of bacteria. To isolate the DNA, a small section of stain containing FTA paper is punched out and is washed briefly and the DNA can be purified using FTA purification reagent. The bound heme and other proteins elute out in the supernatant. Since the application of stain on FTA paper causes the cells to lyse, the DNA gets bound to the fibers of the paper, preserved. This DNA can now be isolated using techniques like chelex or can be directly used in PCR (Lorente et al. 1998; Kline et al. 2002). Several robotic automated workstations have been developed which isolate DNA from FTA paper. The methodology is not as easy as it seems—the static energy between the paper and the plastic well makes it difficult to stay (Butler 2012a, b).

Chelex Technique

In 1991, forensic community started using ion exchange resins for DNA isolation. Chelex 100 resin from Bio Rad was amongst the few popular resins. Chelex 100 was a copolymer of styrene divinylbenzene copolymer along with paired iminodiaetate ions. As per the protocol provided by Walsh et al. the sample is directly added to the chelex resin. The mixture is now heated which in turns denatures the DNA and helps in cell lysis. With the help of centrifugation the resin along with cell debris is removed and the DNA is contained in the supernatant. This supernatant can directly be used in PCR. In this approach the resin chelates the divalent ions and thereby deactivates the nucleases. Although the technique fails to filter out the contaminants present in sample like heme. It proves to be a simple, single step and rapid procedure to extract single stranded DNA. This method of DNA extraction is simple, rapid, and avoids contamination of the DNA. The main disadvantage of the technique lies in its failure to filter out the contaminants present in sample like heme (McKiernan and Danielson 2017; Elkins 2013; Butler 2012a, b; Walsh et al. 1991).

1.4.2 Advancements in PCR Technique

Most of the currently available amplification kits available today contain the recommended global core STR loci. These kits have a high level of tolerance to potential PCR inhibitors. Besides they can amplify the STR loci from the degraded samples. The extensive list of PCR kits commercially available now-a-days along with their marker features is given in Table 1.2.

1.4.3 Capillary Electrophoresis and Analysis

Capillary electrophoresis is a technique used by the forensic experts to separate and analyze the STRs present in the sample DNA. Applied Biosystems, introduced the technology for the very first time in 1995. The technology was called as the single-capillary ABI Prism 310 Genetic Analyzer. Ever since many technologies on the similar platforms were developed namely: 3100, 3130, 3700, 3730, and 3500 series. The instrument works on the principle of separation of DNA molecule while moving inside the gel filled capillary, under the effect of electric field. The liquid media forms a sieve-like structure through which the DNA fragments pass. The smaller fragment will pass with an ease and will take smaller amount of time to pass through the capillary as compared to the larger fragments. Hence, the fragments get separated on the basis of their migration speed. To aid in detection, the fragments are fluorescent labeled. The instrument consists of—a narrow capillary made primarily of glass having internal diameter of 50 μm, vials for the buffer (2 in number), electrodes, power supply, laser source (excitation), sample injector, and detector. The whole instrument is connected to a computer for storage of data and analysis. The capillaries range from 30 to 80 cm in length. To start with the process, the capillaries are filled with the viscous fluid media. The fluid serves as a medium

Table 1.2 List of commercially available STR genotyping kits and their marker features

S. No.	Name of the kit	List of STR markers	Manufacturer's Name and Address
1	AmpF*l*STR® Identifiler® kit	D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, D2S1338, D19S433, vWA, TPOX, D18S51, Amelogenin, D5S818, FGA	ThermoFisher Scientific, Foster City, CA, USA
2	AmpF*l*STR® Identifiler® Plus kit	CSF1PO, D2S1338, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D19S433, D21S11, FGA, TH01, TPOX, vWA, Amelogenin	ThermoFisher Scientific, Foster City, CA, USA
3	AmpF*l*STR® Identifiler® Direct kit	D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, D2S1338, D19S433, vWA, TPOX, D18S51, D5S818, FGA, Amelogenin	ThermoFisher Scientific, Foster City, CA, USA
4	AmpF*l*STR® MiniFiler™ kit	CSF1PO, FGA, D16S539, D18S51, Amelogenin, D2S1338, D21S11, D7S820	ThermoFisher Scientific, Foster City, CA, USA
5	AmpF*l*STR® NGM™ kit	D3S1358, vWA, D16S539, D2S1338, D8S1179, D21S11, D18S51, D19S433, TH01, FGA, D1S1656, D12S391, D10S1248, D22S1045, D2S441, Amelogenin	ThermoFisher Scientific, Foster City, CA, USA
6	AmpF*l*STR® NGM SElect™ kit	D3S1358, vWA, D16S539, D2S1338, D8S1179, D21S11, D18S51, D19S433, TH01, FGA, Amelogenin, D10S1248, D22S1045, D2S441, D1S1656, D12S391, SE33	ThermoFisher Scientific, Foster City, CA, USA
7	AmpF*l*STR® NGM SElect™ Express	D3S1358, vWA, D16S539, D2S1338, D8S1179, D21S11, D18S51, D19S433, TH01, FGA, Amelogenin, D10S1248, D22S1045, D2S441, D1S1656, D12S391, SE33	ThermoFisher Scientific, Foster City, CA, USA
8	AmpF*l*STR® NGM Detect™ kit	FGA, TH01, vWA, D3S1358, D8S1179, D18S51, D21S11, D12S391, D1S1656, D2S441, D10S1248, D22S1045, D16S539, D2S1338, D19S433, SE33, IQCS, IQCL, Y indel, Amelogenin	ThermoFisher Scientific, Foster City, CA, USA
9	GlobalFiler™ kit	D3S1358, vWA, D16S539, CSF1PO, TPOX, D8S1179, D21S11, D18S51, D2S441, D19S433, TH01, FGA, D22S1045, D5S818, D13S317, D7S820, SE33, D10S1248, D1S1656, D12S391, D2S1338, Amelogenin, DYS391, Y indel	ThermoFisher Scientific, Foster City, CA, USA

(continued)

Table 1.2 (continued)

S. No.	Name of the kit	List of STR markers	Manufacturer's Name and Address
10	GlobalFiler™ Express kit	D3S1358, vWA, D16S539, CSF1PO, TPOX, D8S1179, D21S11, D18S51, D2S441, D19S433, TH01, FGA, D22S1045, D5S818, D13S317, D7S820, SE33, D10S1248, D1S1656, D12S391, D2S1338, Amelogenin, DYS391, Y indel	ThermoFisher Scientific, Foster City, CA, USA
11	VeriFiler™ Express kit	D3S1358, vWA, D16S539, CSF1PO, TPOX, D8S1179, D21S11, D18S51, D2S441, D19S433, TH01, FGA, D22S1045, D5S818, D13S317, D7S820, D10S1248, D1S1656, D12S391, D2S1338, D6S1043, Penta D, Penta E, Y indel, Amelogenin	ThermoFisher Scientific, Foster City, CA, USA
12	AmpF*l*STR® Yfiler® kit	DYS19, DYS385a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS438, DYS439, DYS437, DYS448, DYS456, DYS458, DYS635 (YGATAC4), YGATAH4	ThermoFisher Scientific, Foster City, CA, USA
13	AmpF*l*STR® Yfiler® Direct kit	DYS19, DYS385a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS438, DYS439, DYS437, DYS448, DYS456, DYS458, DYS635 (YGATAC4), YGATAH4	ThermoFisher Scientific, Foster City, CA, USA
14	Yfiler™ Plus kit	DYS576, DYS389I, DYS635, DYS389II, DYS627, DYS460, DYS458, DYS19, YGATAH4, DYS448, DYS391, DYS456, DYS390, DYS438, DYS392, DYS518, DYS570, DYS437, DYS385 a/b, DYS449, DYS393, DYS439, DYS481, DYS387S1, DYS533	ThermoFisher Scientific, Foster City, CA, USA
15	PowerPlex® 18D System	D18S51, D21S11, TH01, D3S1358, FGA, TPOX, D8S1179, vWA, CSF1PO, D16S539, D7S820, D13S317, D5S818, Amelogenin, Penta E, Penta D, D2S1338, and D19S433	Promega Corporation, Madison, WI
16	PowerPlex® 21 System	Amelogenin, D1S1656, D2S1338, D3S1358, D5S818, D6S1043, D7S820, D8S1179, D12S391, D13S317, D16S539, D18S51, D19S433, D21S11, Amelogenin, CSF1PO, FGA, Penta D, Penta E, TH01, TPOX, and vWA	Promega Corporation, Madison, WI

(continued)

Table 1.2 (continued)

S. No.	Name of the kit	List of STR markers	Manufacturer's Name and Address
17	PowerPlex® CS7 System	LPL, F13B, FESFPS, F13A01, Penta D, Penta C, and Penta E	Promega Corporation, Madison, WI
18	PowerPlex® ESX 16 and ESI 16 Fast Systems	D18S51, D21S11, TH01, D3S1358, Amelogenin, D16S539, D2S1338, D1S1656, D10S1248, FGA, D8S1179, vWA, D22S1045, D19S433, D12S391, and D2S441	Promega Corporation, Madison, WI
19	PowerPlex® ESX 17 and ESI 17 Fast Systems	D3S1358, D8S1179, D18S51, D21S11, FGA, TH01, vWA, D2S441, D10S1248, D22S1045, D1S1656, D12S391, D2S1338, D16S539, D19S433, SE33, and Amelogenin	Promega Corporation, Madison, WI
20	PowerPlex® Fusion 6C System	CSF1PO, FGA, TH01, vWA, D1S1656, D2S1338, D2S441, D3S1358, D5S818, D7S820, D8S1179, D10S1248, D12S391, D13S317, D16S539, D18S51, D19S433, D21S11, Amelogenin, DYS391, Penta D, Penta E, D22S1045, TPOX, SE33, DYS570, DYS576	Promega Corporation, Madison, WI
21	PowerPlex® Fusion System	CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D10S1248, D22S1045, D2S441, D1S1656, D12S391, Amelogenin, DYS391, Penta D, Penta E, D2S1338, D19S433	Promega Corporation, Madison, WI
22	PowerPlex® Y23 System	DYS576, DYS389I/II, DYS448, DYS19, DYS391, DYS481, DYS549, DYS533, DYS438 (penta), DYS437, DYS570, DYS635, DYS390, DYS439, DYS392, DYS643 (penta), DYS393, DYS458, DYS385a/b, DYS456, and Y-GATA-H4	Promega Corporation, Madison, WI
23	Investigator 24plex QS Kit	Amelogenin, TH01, D3S1358, vWA, D21S11, TPOX, DYS391, D1S1656, D12S391, SE33, D10S1248, D22S1045, D19S433, D8S1179, D2S1338, D2S441, D18S51, FGA, QS1, D16S539, CSF1PO, D13S317, D5S818, D7S820, QS2	QIAGEN, Hilden, Germany
24	Investigator 24plex GO! Kit	Amelogenin, TH01, D3S1358, vWA, D21S11, TPOX, DYS391, D1S1656, D12S391, SE33, D10S1248, D22S1045, D19S433, D8S1179, D2S1338, D2S441, D18S51, FGA, QS1, D16S539, CSF1PO, D13S317, D5S818, D7S820, QS2	QIAGEN, Hilden, Germany

(continued)

Table 1.2 (continued)

S. No.	Name of the kit	List of STR markers	Manufacturer's Name and Address
25	Investigator ESSplex SE QS Kit	QS1, Amelogenin, TH01, D3S1358, vWA, D21S11, QS2, D163539, D1S1656, D19S433, SE33, D10S1248, D22S1045, D12S391, D8S1179, D2S1338, D2S441, D18S51, FGA	QIAGEN, Hilden, Germany
26	Investigator ESSplex SE GO! Kit	Amelogenin, TH01, D3S1358, vWA, D21S11, D16S539, D1S1656, D19S433, SE33, D10S1248, D22S1045, D12S391, D8S1179, FGA, D2S1338, D2S441, D18S51	QIAGEN, Hilden, Germany
27	Investigator ESSplex SE Plus Kit	Amelogenin, TH01, D3S1358, vWA, D21S11, D16S539, D1S1656, D19S433, D8S1179, D2S1338, D10S1248, D22S1045, D12S391, FGA, D2S441, D18S51	QIAGEN, Hilden, Germany
28	Investigator ESSplex Plus Kit	Amelogenin, TH01, D3S1358, vWA, D21S11, D16S539, D1S1656, D19S433, D8S1179, D2S1338, D10S1248, D22S1045, D12S391, FGA, D2S441, D18S51	QIAGEN, Hilden, Germany
29	Investigator IDplex Plus Kit	Amelogenin, TH01, D3S1358, vWA, D21S11, TPOX, D7S820, D19S433, D5S818, D2S1338, D16S539, CSF1PO, D13S317, FGA, D18S51, D8S1179	QIAGEN, Hilden, Germany
30	Investigator IDplex GO! Kit	Amelogenin, TH01, D3S1358, vWA, D21S11, TPOX, D7S820, D19S433, D5S818, D2S1338, D16S539, CSF1PO, D13S317, FGA, D18S51, D8S1179	QIAGEN, Hilden, Germany
31	Investigator Argus X-12 QS Kit	QS1, Amelogenin, DXS10103, DXS8378, DXS10101, DXS10134, DXS10074, DXS7132, DXS10135, DXS7423, DXS10146, DXS10079, DXSHPRTB, DXS10148, D21S11	QIAGEN, Hilden, Germany
32	Investigator HDplex Kit	Amelogenin, D7S1517, D3S1744, D12S391, D2S1360, D6S474, D4S2366, D8S1132, D5S2500, D18S51, D21S2055, D10S2325, SE33	QIAGEN, Hilden, Germany
33	Investigator Argus Y-12 QS Kit	DYS439, DYS437, DYS390, DYS385, DYS391, DYS389-I, DYS19, DYS389-II, DYS393, DYS438, DYS392	QIAGEN, Hilden, Germany

(continued)

Table 1.2 (continued)

S. No.	Name of the kit	List of STR markers	Manufacturer's Name and Address
34	COrDIS Plus kit	D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, CSF1PO, FGA, TH01, TPOX, VWA, D1S1656, D2S441, D10S1248, D12S391, D22S1045 and SE33, Amelogenin	Gordiz, Moskva, Russia
35	COrDYS-Y kit	DYS19, DYS389I/II, DYS390, DYS391, DYS392, DYS393, and DYS385 a/b, DYS438 and DYS439, DYS437, DYS456, DYS635, DYS448, DYS576, DYS481, DYS449	Gordiz, Moskva, Russia
36	COrDX kit	DXS10148, DXS10135, DXS8378, DXS7132, DXS10079, DXS10075, DXS10074, DXS10103, HPRTB, DXS10101, DXS10146, DXS8377, DXS10134, DXS7423	Gordiz, Moskva, Russia
37	iPLEX-STR™ Kit	Amelogenin, D3S1358, TH01, D12S391, D1S1656, D10S1248, D2S441, D7S820, D13S317, FGA, TPOX, D18S51, D16S539, D8S1179, CSF1PO, D5S818, vWA, D21S11, SE33	Independent Forensics, 500 Waters Edge Lane, Lombard
38	iPLEX-STR™ Y Kit	DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS385 a/b, DYS438, DYS439, DYS437, DYS456, DYS635, DYS448, DYS576, DYS481, DYS449	Independent Forensics, 500 Waters Edge Lane, Lombard
39	iPLEX X Kit	DXS10148, DXS10135, DXS8378, DXS7132, DXS10079, DXS10075, DXS10074, DXS10103, HPRTB, DXS10101, DXS10146, DXS8377, DXS10134, DXS7423	Independent Forensics, 500 Waters Edge Lane, Lombard
40	SureID® 23comp Human DNA Identification Kit	Amelogenin, D18S1364, D1S1656, D13S325, D9S1122, D4S2366, D3S1744, D12S391, D11S2368, D21S2055, D20S482, D8S1132, D7S3048, D2S441, D19S253, D10S1248, D17S1301, D22-GATA198B05, D16S539, D6S474, D14S1434, D15S659	HEALTH Gene Technologies Co. Ltd., China
41	SureID® 21G Human STR Identification Kit	CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, and D21S11, D12S391, D19S433, D1S1656, D2S1338, D6S1043, Penta E, Penta D, Amelogenin	HEALTH Gene Technologies Co. Ltd., China

(continued)

Table 1.2 (continued)

S. No.	Name of the kit	List of STR markers	Manufacturer's Name and Address
42	SureID® 27Y Human STR Identification Kit	DYS456, DYS576, DYS570, DYS481, DYF387S1, DYS627, DYS458, DYS460, DYS437, DYS439, DYS392, DYS385, DYS393, DYS391, DYS390, DYS456, DYS635, DYS449, DYS533, DYS438, DYS389I, DYS448, DYS389II, DYS19, GATA_H4, DYS518	HEALTH Gene Technologies Co. Ltd., China
43	DNATyper™15 PCR Genotyping System	D6S1043, D21S11, D7S820, CSF1PO, D2S1338, D3S1358, D13S317, D8S1179, D16S539, Penta E, D5S818, vWA, D18S51, FGA, Amelogenin	Institute of Forensic Science, Ministry of Public Security, Beijing, China

through which fragments pass. The sample is injected via electrokinetic injection ports. These injection ports allow high resolution in short distances.

The sample from PCR is mixed with the STR typing kits available in the market and is diluted so that the sample does not interfere with the electrokinetic injection process. To ensure DNA denaturation sufficient amount of formamide is added to sample mix. The sample is now allowed to separate using the high voltage across the electrodes. The factors effecting the movement and thus separation of the DNA fragments are the polymer matrix, capillary dimensions, buffer used (100 mmol/L *N*-tris-(hydroxymethyl)-methyl-3-aminopropane-sulfonic acid and 1 mmol/L EDTA, pH 8.0). The polymer solutions generally in use are POP-4, POP-6 and POP-7 with urea (performance optimized polymer). These are un-cross linked liner polymers of dimethyl polyacrylamide with 4% and 6% concentrations, respectively. The fragments are detected by the fluorescent detectors namely a photomultiplier tube (PMT) or a charge-coupled device (CCD) and is measured in relative fluorescence units (RFU). The PCR primers are fluorescent labeled and hence the dye gets incorporated into the DNA fragments. With each fragment passing the window—carved in the capillary, the laser is focused upon them—illuminating the DNA. The signals are captured by the detectors and are plotted as peaks (Butler 2012a, b).

1.5 SNP-Based Typing

Single nucleotide polymorphisms are the single base pair (biallelic) variations such as insertions, deletions, and substitutions found in the genome of an individual. SNPs are diverse in nature and estimated to occur once in every 1000 bp. They are the commonly occurring polymorphism in the human genome. The SNPs occur at coding as well as noncoding gene sequences. The forensically important SNPs occur mainly in noncoding regions of the genome. Apart from being useful to forensic

community they aid in understanding ancestry and population genetics. It is thought that SNPs are the future of DNA typing methodologies and will soon replace the use of STRs in laboratories. The forensic community is more inclined toward use of SNPs despite of STRs in analysis since they are biallelic hence degraded forensic samples requiring small amplicons can be analyzed smoothly. Further, these are less prone to mutations compared to STRs. SNPs can be readily used with high performance techniques—required for large DNA databases. This feature also aids in interpretation of the cases which require analysis of Y chromosome and mtDNA typing. Since SNPs can occur in coding region of genome, a single variation may cause change in phenotype of an individual. Thus, in future SNPs can aid in phenotyping of individuals. For forensic purposes SNPs can be grouped as—(1) Identity-testing SNPs: SNPs used for individualization. These require being heterozygous and having low inbreeding coefficient. (2) Lineage SNPs— these are the group of linked SNPs also called as haplotype markers and are useful in missing cases. (3) Ancestry informative SNPs—These SNPs aid in establishing phenotypic identity on the basis of ancestry. These have low heterozygosity and high inbreeding coefficient. (4) Phenotype informative SNPs— the SNPs from which phenotypic characters can be inferred like eye color, skin color, etc.

These SNPs are regularly typed in forensic laboratories where forensic sample is either degraded or is age old. The results are then compared with reference which establishes the identity of the individual. There are several methods for SNP typing. Each laboratory has its own framework and as per the requirement of the case the method of choice is adapted. The chosen method needs to be highly accurate, sensitive. SNP typing methods constitute any one of the molecular mechanism among allele-specific hybridization, primer extension, oligonucleotide ligation, and invasive cleavage. The detection method may however depend on the type of method chosen. A set of 50–75 SNPs is used in forensic caseworks for analysis. However, the number of SNPs required depends on the type of sample. Some samples require large number of SNP markers whereas few samples require a very small number of markers. A core group of SNPs are selected by European Consortium SNPforID (Growth Program GRD1-2002-71802) based on linkage, validation, quality, and polymorphism. Several strategies while choosing the SNPs for Y chromosome, autosomal chromosome are being introduced to upgrade the technologies (Table 1.3). The PCR amplification acts as a limiting factor in the SNP typing technology (Sobrino et al. 2005; Budowle and van Daal 2008; Salwa Teama 2018).

1.6 Next Generation Sequencing

The National Cancer Institute has defined the genetic term next generation sequencing as an advanced technology which is used to sequence the DNA of individual. NGS can sequence large amount of DNA simultaneously hence is rightly called as Massively Parallel Sequencing (MPS) technology. Sequencing techniques have been developed since 1977 when sanger had demonstrated the newly developed

Table 1.3 List of available SNP panels (https://strbase.nist.gov/)

S. No.	SNP assay	References
1	35plexY-SNPminisequencing assay	Sanchez et al. (2003)
2	50 Y-SNPs in 8 multiplexes	Vallone and Butler (2004)
3	11plex mtSNP minisequencing assay	Vallone et al. (2004)
4	21plex Autosomal IISNP assay	Dixon et al. (2005)
5	29plex Y-SNP minisequencing assay	Brión et al. (2005)
6	52plex Autosomal IISNP assay	Sanchez et al. (2006)
7	34plex AISNP assay	Phillips et al. (2007)
8	90 Autosomal IISNPs in two multiplexes	Fang et al. (2009)
9	6plex PISNP assay for eye color	Walsh et al. (2011)

DNA sequencing technique called Sanger sequencing. This methodology involved PCR like reactions using the deoxynucleotidetriphosphates (dNTPs) primarily along with polymerases (chain termination method). The Sanger sequencing technology laid the foundation of the DNA sequencing was used in labs as first generation technology. Since the methodology proved to be a boon for genetic research it wasn't cost effective and handy. Thus further improvements led to the adoption of next generation sequencing tools. The second generation applications followed amplification and sequencing of the DNA fragment having adapter sequences at both ends.

In the same format instruments like GE Healthcare MegaBACE™ system, the system of Intelligent Bio-Systems Inc., sequencing with nanoballs by Complete Genomics™, the Polonator (developed by the research group of Church at Harvard), 454 system (Roche) and the systems from Illumina® and Applied Biosystems™/Life Technologies became available in the market. Technologies brought by the third generation sequencing were PacBio (Pacific Biosciences) and HeliScope (Helicos Biosciences) which were less prone to error and use enzymatic template replication system to sequence the single molecule templates. The fluorescent dyed pyrophosphate helps in analyzing the sequence. The advancements in science and technology have now led to the development of fourth generation technology or as said the advance model of NGS. The nanopore technology has been introduced with which the single molecules can be read directly. In this format Oxford Nanopore Technologies (ONT) has launched two platforms—"GridION™" and "MinION™." Since there are numerous instruments available in the market, the technology to be used depends directly upon the sample type and requirements of the case (Bruijns et al. 2018; Yang et al. 2014; Kim et al. 2016; Børsting and Morling 2015).

1.7 DNA Phenotyping

Forensic scientists are constantly working toward development of new techniques to make human identification from the forensic sample easy, reliable, and effective. Among such techniques DNA profiling using STR markers is the most accepted

method. While these STR markers have several advantages like having high discriminatory power, the technique holds few drawbacks as well. The STR markers consist of long nucleotide repeated sequences which are not suitable for the analysis of degraded forensic samples. Further the profile generated needs to be compared either to a reference profile or to the profiles selected through databases.

To overcome such limitations the novel concept of forensic DNA phenotyping has been introduced. In this methodology, DNA sequences are used to infer about the morphology of the individual. The observable traits or the phenotype of an individual is controlled by both genetic factors and external environmental factors (multi-factorial). The genetic data is thus used to predict the characteristic feature of an individual upto an extent. Phenotypic characters like hair, eye color, skin color, etc. are now phenotyped using genetic information. Pigmentation genes responsible for eye color-HERC2, OCA2, SLC24A4, SLC45A2, TYR, and IRF4 are used by Irisplex System, first tool developed for phenotyping. It has the high accuracy of likelihood. Similarly, MC1R, SLC45A2, SLC24A5, and HERC2 and an SNP-based model has been used for hair color determination having 88% accuracy. Later, HIrisplex System was developed and validated for both hair and eye color markers. For the skin color SLC24A5, HERC2, and SLC45A2 markers have been incorporated into the phenotyping kits available in the market. Thus, this methodology is still novel and new phenotyping markers are continuously being researched and incorporated into the kits (Marano and Fridman 2019; Samuel and Prainsack 2019; Kayser 2015).

References

Barbaro A et al (2009) Validation of PrepFiler forensic DNA extraction kit (applied biosystems). For Sci Int Genet Suppl Ser 2:176–177

Boom R et al (1990) Rapid and simple method for purification of nucleic acids. J Clin Microbiol 28:495–503

Børsting C, Morling N (2015) Next generation sequencing and its applications in forensic genetics. For Sci Int Genet 18:78–89. https://doi.org/10.1016/j.fsigen.2015.02.002

Brevnov MG et al (2009) Validation of the PrepFiler forensic DNA extraction kit for extraction of genomic DNA from biological samples. J Forensic Sci 54:599–607

Brión M, Sanchez J, Balogh K, Thacker C, Blanco-Verea A, Børsting C et al (2005) Introduction of an single nucleodite polymorphism-based “Major Y-chromosome haplogroup typing kit” suitable for predicting the geographical origin of male lineages. Electrophoresis 26(23):4411–4420. https://doi.org/10.1002/elps.200500293

Bruijns B, Tiggelaar R, Gardeniers H (2018) Massively parallel sequencing techniques for forensics: a review. Electrophoresis 39(21):2642–2654. https://doi.org/10.1002/elps.201800082

Budowle B, Giusti A, Allen R (1990) Analysis of PCR products (pMCT118) by polyacrylamide gel electrophoresis. Advances in forensic haemogenetics, pp 148–150. https://doi.org/10.1007/978-3-642-75496-8_44

Budowle B, van Daal A (2008) Forensically relevant SNP classes. Biotechniques 44(5):603–610. https://doi.org/10.2144/000112806

Burgoyne LA (1996) Solid medium and method for DNA storage. U.S. Patent 5,496,562

Butler J (2012a) Advanced topics in forensic DNA typing. Academic Press, London, pp 141–165

Butler J (2012b) Advanced topics in forensic DNA typing. Academic Press, London, pp 29–47

Carracedo A, Sánchez-Diz P (2005) Forensic DNA-typing technologies: a review. In: Forensic DNA typing protocols, pp 001–012. https://doi.org/10.1385/1-59259-867-6:001
Dash H, Rawat N, Das S (2020) Alternatives to amelogenin markers for sex determination in humans and their forensic relevance. Mol Biol Rep 47(3):2347–2360. https://doi.org/10.1007/s11033-020-05268-y
Dash H, Shrivastava P, Das S (2019) Expediency of tetra- and pentanucleotide repeat autosomal STR markers for DNA typing in Central Indian Population. Proc Natl Acad Sci India B Biol Sci. https://doi.org/10.1007/s40011-019-01156-z
Dixon L, Murray C, Archer E, Dobbins A, Koumi P, Gill P (2005) Validation of a 21-locus autosomal SNP multiplex for forensic identification purposes. Forensic Sci Int 154(1):62–77. https://doi.org/10.1016/j.forsciint.2004.12.011
Doyle A, Roden C (1993) The memoirs of Sherlock Holmes. Oxford University Press, Oxford
Duncan E et al (2003) Isolation of genomic DNA. In: Bowien B, Dürre P (eds) Nucleic acids isolation methods. American Scientific Publishers, Stevenson Ranch, CA, pp 7–19
Elkins K (2013) DNA extraction. In: Forensic DNA biology. Academic Press, Oxford, pp 39–52. https://doi.org/10.1016/b978-0-12-394585-3.00004-3
Fang R, Pakstis A, Hyland F, Wang D, Shewale J, Kidd J et al (2009) Multiplexed SNP detection panels for human identification. For Sci Int Genet Suppl Ser 2(1):538–539. https://doi.org/10.1016/j.fsigss.2009.08.161
Giardina E, Spinella A, Novelli G (2011) Past, present and future of forensic DNA typing. Nanomedicine 6(2):257–270. https://doi.org/10.2217/nnm.10.160
Greenspoon SA et al (2004) Application of the BioMek 2000 laboratory automation workstation and the DNA IQ system to the extraction of forensic casework samples. J Forensic Sci 49:29–39
Grover A, Sharma P (2014) Development and use of molecular markers: past and present. Crit Rev Biotechnol 36(2):290–302. https://doi.org/10.3109/07388551.2014.959891
Jobling M, Gill P (2004) Encoded evidence: DNA in forensic analysis. Nat Rev Genet 5 (10):739–751. https://doi.org/10.1038/nrg1455
Kayser M (2015) Forensic DNA Phenotyping: predicting human appearance from crime scene material for investigative purposes. For Sci Int Genet 18:33–48. https://doi.org/10.1016/j.fsigen.2015.02.003
Kim E, Lee H, Yang I, Jung S, Yang W, Shin K (2016) Massively parallel sequencing of 17 commonly used forensic autosomal STRs and amelogenin with small amplicons. For Sci Int Genet 22:1–7. https://doi.org/10.1016/j.fsigen.2016.01.001
Kline MC et al (2002) Polymerase chain reaction amplification of DNA from aged blood stains: quantitative evaluation of the "suitability for purpose" of four filter papers as archival media. Anal Chem 74:1863–1869
Klintschar M, Neuhuber F (2000) Evaluation of an alkaline lysis method for the extraction of DNA from whole blood and forensic stains for STR analysis. J Forensic Sci 45:669–673
Köchl S, Niederstätter H, Parson W (2005) DNA extraction and quantitation of forensic samples using the phenol–chloroform method and real-time PCR. For DNA Typ Protoc 297:13–30. https://doi.org/10.1385/1-59259-867-6:013
Lorente JA et al (1998) Newborn genetic identification: expanding the fields of forensic haemogenetics. Prog Forensic Genet 7:114–116
Marano L, Fridman C (2019) DNA phenotyping: current application in forensic science. Res Rep For Med Sci 9:1–8. https://doi.org/10.2147/rrfms.s164090
Martin P, Schmitter H, Schneider P (2001) A brief history of the formation of DNA databases in forensic science within Europe. Forensic Sci Int 119(2):225–231. https://doi.org/10.1016/s0379-0738(00)00436-9
McKiernan H, Danielson P (2017) Molecular diagnostic applications in forensic science. In: Molecular diagnostics. Academic, New York, NY, pp 371–394. https://doi.org/10.1016/b978-0-12-802971-8.00021-3

Miller S, Dykes D, Polesky H (1988) A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16(3):1215–1215. https://doi.org/10.1093/nair/16.3.1215

Moss D et al (2003) An easily automated, closed-tube forensic DNA extraction procedure using a thermostable proteinase. Int J Legal Med 117:340–349

Panneerchelvam S, Norazmi MN (2003) Forensic DNA profiling and database. Malaysian J Med Sci 10(2):20–26

Phillips C, Salas A, Sánchez J, Fondevila M, Gómez-Tato A, Álvarez-Dios J et al (2007) Inferring ancestral origin using a single multiplex assay of ancestry-informative marker SNPs. For Sci Int Genet 1(3-4):273–280. https://doi.org/10.1016/j.fsigen.2007.06.008

Primorac D, Schanfield MS, Primorac D (2000) Application of forensic DNA testing in the legal system. Croat Med J 41:32–46

Roewer L (2013) DNA fingerprinting in forensics: past, present, future. Investig Genet 4(1):22. https://doi.org/10.1186/2041-2223-4-22

Samuel G, Prainsack B (2019) Forensic DNA phenotyping in Europe: views “on the ground” from those who have a professional stake in the technology. New Genet Soc 38:119–141. https://doi.org/10.1080/14636778.2018.1549984

Sanchez J, Børsting C, Hallenberg C, Buchard A, Hernandez A, Morling N (2003) Multiplex PCR and minisequencing of SNPs—a model with 35 Y chromosome SNPs. Forensic Sci Int 137(1):74–84. https://doi.org/10.1016/s0379-0738(03)00299-8

Sanchez J, Phillips C, Børsting C, Balogh K, Bogus M, Fondevila M et al (2006) A multiplex assay with 52 single nucleotide polymorphisms for human identification. Electrophoresis 27(9):1713–1724. https://doi.org/10.1002/elps.200500671

Sobrino B, Brión M, Carracedo A (2005) SNPs in forensic genetics: a review on SNP typing methodologies. Forensic Sci Int 154(2-3):181–194. https://doi.org/10.1016/j.forsciint.2004.10.020

Tan S, Yiap B (2009) DNA, RNA, and protein extraction: the past and the present. J Biomed Biotechnol 2009:1–10. https://doi.org/10.1155/2009/574398

Teama S (2018) DNA polymorphisms: DNA-based molecular markers and their application in medicine. InTech, Rijeka. https://doi.org/10.5772/intechopen.79517

Tereba AM et al. (2004) Simultaneous isolation and quantitation of DNA. U.S. Patent 6,673,631

Vallone P, Butler J (2004) Y-SNP typing of U.S. African American and Caucasian samples using allele-specific hybridization and primer extension. J Forensic Sci 49(4):1–10. https://doi.org/10.1520/jfs2003303

Vallone P, Just R, Coble M, Butler J, Parsons T (2004) A multiplex allele-specific primer extension assay for forensically informative SNPs distributed throughout the mitochondrial genome. Int J Legal Med 118(3):147–157. https://doi.org/10.1007/s00414-004-0428-5

Verma S, Goswami G (2014) DNA evidence: current perspective and future challenges in India. Forensic Sci Int 241:183–189. https://doi.org/10.1016/j.forsciint.2014.05.016

Vogelstein B, Gillespie D (1979) Preparative and analytical purification of DNA from agarose. Proc Natl Acad Sci U S A 76:615–619

Walsh PS et al (1991) Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques 10:506–513

Walsh S, Liu F, Ballantyne K, van Oven M, Lao O, Kayser M (2011) IrisPlex: a sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. For Sci Int Genet 5(3):170–180. https://doi.org/10.1016/j.fsigen.2010.02.004

Yang Y, Xie B, Yan J (2014) Application of next-generation sequencing technology in forensic science. Genom Proteom Bioinformatics 12(5):190–197. https://doi.org/10.1016/j.gpb.2014.09.001

STR Typing and Available Multiplex Kits Including Validation Methods

2

Anna Barbaro

Abstract

Since the first introduction of DNA typing in 1987, more than 30 years have been passed and an enormous growth in forensics occurred. DNA test routinely represents the most important tool for solving forensic caseworks and biological relationships. Forensic laboratories carry out annually hundreds of DNA tests because modern technologies allow obtaining conclusive results from a wide range of biological evidences even if degraded or in small quantities. Repeated DNA regions selected by the international scientific community for forensic applications include *minisatellites (VNTR)* and *microsatellites (STR)*. In particular STRs became rapidly the first-choice markers for human identification because of their features and the good polymorphism degree. Once loci were definitively established, commercial multiplex were developed including initially only few loci till to current large systems that allow now simultaneous amplification of European and USA database core loci. Thus nowadays, large commercial STR kits are widely used either in forensic than in paternity tests because of their easy use and high discriminatory power. Obviously, to assess new forensic DNA technologies, two types of validation are required: *developmental* performed directly by the manufacturer and *internal* done by the user—laboratory. According to the Scientific Working Group on DNA Analysis Methods (SWGDAM) recommendations, several different steps are required in both cases.

Keywords

STR · Polymorphism · PCR · Multiplex · Validation

A. Barbaro (✉)
Department of Forensic Genetics, Studio Indagini Mediche E Forensi (SIMEF), Reggio Calabria, Italy

Worldwide Association of Women Forensic Experts (WAWFE), Rome, Italy
e-mail: simef_dna@tiscali.it

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_2

2.1 Introduction

DNA analysis was firstly introduced in the mid-1980s and rapidly developed so that it now represents the most important tool for solving forensic caseworks, resolving biological relationships, identifying missing persons or mass disaster victims.

In humans only a small part (0.3%) of genome differs among individuals: this difference is mainly due to the presence of highly repeated DNA traits available in almost every chromosome, which constitute more than 40% of the human genome. Repeated sequences differ on the basis of the total length of the repeated area, the length of the repeated units and the number of units. Since these regions are generally placed between genes, they may change from individual to individual, without affecting the person's health. Repeated regions selected by the international scientific community for forensic applications include *minisatellites* and *microsatellites*.

2.2 Variable Number of Tandem Repeats (VNTRs)

Minisatellites (called also variable number of tandem repeats "VNTRs") generally have repeats of 9–80 bp length and they are typically repeated 5–50 times (Fig. 2.1). Some minisatellites show also a "core unit" including some nucleotides "GGGCAGGANG" (where N can be any base). Minisatellites are found at more than 1000 locations in the human genome; they are generally placed in the centromeric or telomeric area of a chromosome, mainly in noncoding regions even if sometimes they are part of genes. They show high mutation rate and high diversity in the populations (Jeffreys et al. 1984).

In 1985, Prof. Alec Jeffreys, from the University of Leicester (UK), developed a technique called by him "DNA fingerprinting" able to detect length variation of some different VNTRs and to produce a pattern similar to a "bar code" that is unique for each individual. The procedure used by Jeffreys to examine VNTRs involved the use of a restriction enzyme to cut the regions of DNA surrounding the VNTRs so to obtain "restriction fragment length polymorphism (RFLPs)". In this procedure

Fig. 2.1 Example of variable number of tandem repeats (VNTRs)

fragments were separated in size by electrophoresis on agarose gel and then transferred to a nylon membrane for subsequent detection by radioactive or chemiluminescent labelled probes (Southern Blotting) (Jeffreys et al. 1985a; Southern 1975).

Since resulting pattern is unique for each person and it's the same for every cell, tissue, and organ, without the possibility to be altered by any known treatment, the use of this technique made possible the first human DNA identity test. The test was used in UK in 1985 to solve a controversial immigration case and then in 1987 the double-murder case of two teenagers, Lynda Mann and Dawn Ashworth, raped and killed in Narborough (Leicestershire), respectively in 1983 and 1986. Even if VNTRs analysis was successfully used in forensic caseworks for several years, anyway its use was limited because a large amount of not degraded DNA was required for a successful typing and VNTR patterns interpretation was often problematic, specially in case of mixed samples (Jeffreys et al. 1985b).

The introduction of the polymerase chain reaction (PCR), invented in 1983 by K. Mullis, was a great revolution that allowed to forensic scientists the analysis of small biological samples not analyzable with the original procedure (Mullis 1990).

2.3 Short Tandem Repeats (STRs)

Microsatellites, also named short tandem repeats (STRs), are simpler than VNTRs and generally contain 2–6 bp repeats (most common is a 4 bases repetition): the unit is generally repeated many times in "head-tail" manner (Fig. 2.2). While minisatellites are mainly interspersed in genoma, STRs are generally found in the

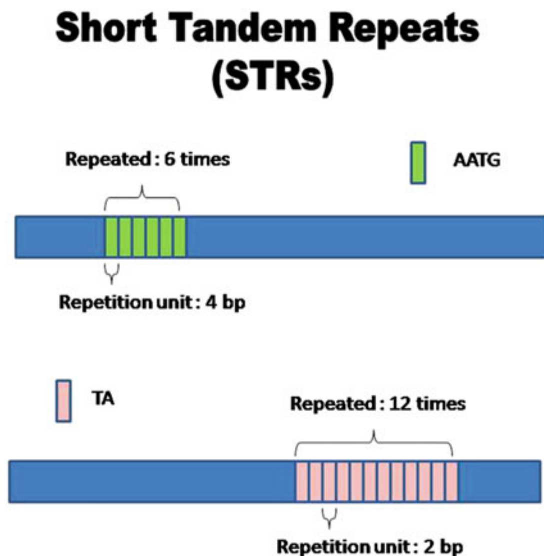


Fig. 2.2 Example of short tandem repeats (STR)

Table 2.1 Types of STR according to length (Butler 2001)

Classification	Length of repeat
Dinucleotides	2 bases repeat (i.e. AC, AT)
Trinucleotides	3 bases repeat (i.e. AAC, AAT)
Tetranucleotides	4 bases repeat (i.e. AAAC, AAAT)
Pentanucleotides	5 bases repeat (i.e. AAAAC, AAAAT)
Hexanucleotides	6 bases repeat (i.e. AAAAAC, AAAAAT)

Table 2.2 Types of STR according to pattern (Butler 2001)

Classification	Pattern of repeat
Simple	Repeat units with identical length and sequence
Compound	Two or more adjacent simple repeats
Complex	Repeats of variable length or sequences

centromeric area of a chromosome. It has been estimated the presence of one STR per every 6–10 kb in the human genome (Jin et al. 1994).

In unrelated individuals STR markers have different numbers of repeat units and repetition patterns, but the sequence generally does not differ (Table 2.1). It has been observed that GATA and AGAT are the most common repeats in forensic STR loci. According to Pattern (Butler 2001) STR may be classified as simple, compound or complex (Table 2.2). There are hundreds of STR markers in the human genome and several dozen have been studied to be used in human identity testing (Butler 2012).

In particular, for human identification purposes, the scientific community selected only sufficiently polymorphic STRs with the following characteristics (Table 2.3) (Gill et al. 1995; Carracedo and Lareu 1998; Bär et al. 1992, 1997):

- High discriminating power >0.9
- Heterozygosity (Ho) >70%
- Different locations on chromosomes (no closely linked loci)
- Robustness and reproducibility of results (low artefacts)
- Alleles length in the range of 90–500 bp (to favourite the analysis of degraded DNA)
- Low stutters rates

Loci placed in "noncoding DNA region" are chosen in order to avoid giving any additional information about person phenotype or health status. STRs are used in several fields such as genetic mapping, disease diagnosis, evolutionary biology, but they rapidly become the first-choice markers for human identification because of some benefits, which may be summarized in the following ones (Butler et al. 2003):

- regular repetition units (1–6 nucleotides)
- wide range of distinct alleles (polymorphism)
- ability to be analyzed by PCR

Table 2.3 Information about most commonly used Autosomal STR Loci (Butler 2012)

Locus	Chromosome	Chromosomal location	Repeat motif	Repeat category	Allele range	GenBank accession
D1S1656	Chr 1	1q42	TAGA	Compound	8–20.3	G0782
D2S441	Chr 2	2p14	TCTA/TCAA	Compound	8–17	AC079112
TPOX Thyroid peroxidase intron	Chr 2	2p25.3	GAAT	Simple	4–16	M68651
D2S1338	Chr 2	2q35	TGCC/TTCC	Compound	15–28	AC010136
D3S1358	Chr 3	3p21.31	TCTG/TCTA	Compound	8–21	AC099539
FGA Alpha Fibrinogen intron	Chr 4	4q31.3	CTTT/TTCC	Compound	12.2–51.2	M64982
D5S818	Chr 5	5q23.2	AGAT	Simple	7–18	AC008512
CSF1PO Proto-oncogene intron	Chr 5	5q33.1	TAGA	Simple	5–16	X14720
D6S1043	Chr 6	6q15	AGAT	Simple	9–25	G08539
SE33 Beta actin-related pseudogene	Chr 6	6q14	AAAG	Complex	4.2–37	V00481
D7S820	Chr 7	7q21.11	GATA	Simple	5–16	AC004848
D8S1179	Chr 8	8q24.13	TCTA/TCTG	Compound	7–20	AF216671
D10S1248	Chr 10	10q26.3	GGAA	Simple	7–19	AL391869
TH01 Tyrosine Hydroxylase intron	Chr 11	11p15.5	TCAT	Simple	3–14	D00269
D12S391	Chr 12	12p13.2	AGAT/AGAC	Compound	13–27.2	G08921
VWA Von Willebrand factor intron	Chr 12	12p13.31	TCTG/TCTA	Compound	10–25	M25858
D13S317	Chr 13	13q31.1	TATC	Simple	5–16	AL353628
Penta E	Chr 15	15q26.2	AAAGA	simple	5–32	AC027004
D16S539	Chr 16	16q24.1	GATA	simple	5–16	AC02459
D18S51	Chr 18	18q21.33	AGAA	Simple	7–40	AP001534

(continued)

Table 2.3 (continued)

Locus	Chromosome	Chromosomal location	Repeat motif	Repeat category	Allele range	GenBank accession
D19S433	Chr 19	19q12	AAGG/TAGG	Compound	9–17.2	AC008507
D21S11	Chr 21	21q21.1	TCTA/TCTG	Complex	12–41.2	AP000433
Penta D	Chr 21	21q22.3	AAAGA	Simple	1.1–19	AP001752
D22S1045	Chr 22	22q12.3	ATT	Simple	7–20	AL022314

- heterozygotes alleles similar in size
- low mutation rate (average value per locus is 10^{-3} to 10^{-4} mutations per generation)
- high heterozygosity and good discrimination degree

STRs small size (2–6 bp repeat) allows easy PCR amplification in multiplex assays, without the problem of differential amplification. Tetranucleotide STR repeats are generally preferred to dinucleotide or trinucleotide repeats, while pentanucleotide and hexanucleotide repeats are less frequently used because they are less diffused in the human genome. A common problem with STR typing is the presence of *stutters* peaks that are produced by Taq-Polymerase slippage during PCR amplification. Stutters have generally one repeat unit smaller than the true allele ($n - 1$), but sometimes they show one repeat larger ($n + 1$). Stutter percentage varies from locus to locus: in tetranucleotides it's generally less than 15% of the main allele, while with dinucleotides or trinucleotides it may increase till to 30% or more. This makes mixture interpretation very difficult especially when working with low or degraded DNA samples (Sparkes et al. 1996a, b).

2.4 Mini-STRs

DNA from forensic evidences, found for example at crime scene or in missing person identifications caseworks, is often in limited quantity or show high degradation due to the exposition to harsh environmental factors (humidity, heat, UV, etc.) or to microorganism's effect. When degradation occurs, DNA is cleaved in small fragments (average range 80–200 bp) and PCR primers may fail in annealing with the target regions. In these cases, standard STRs typing may be not produce results or yields only partial profiles: in fact as DNA molecule fragmentation increases, dropout of the larger STR loci occurs and this produces a loss of information mainly from larger PCR products. The problem obviously is enhanced when using large multiplex PCR, due to the wide size range of amplicons produced (Hill et al. 2008; Coble and Butler 2005).

Partial DNA profiles unfortunately cannot provide the same discrimination power than full profiles and often do not give enough information able to include or exclude an individual as a donor of a sample. The recovery of data from these difficult samples may be increased analyzing smaller STRs named *Mini-STRs*. Mini-STRs were used for the first time during the identification of the World Trade Center victims: in fact, even if most bone were analyzed by traditional STR analysis, unfortunately many samples were in so hard conditions, because of heat, fire or bacterial degradation, that only very poor or negative results were obtained.

In 2001 John Butler and Bruce McCord, from the National Institute of Science and Technology (NIST), were developing new small primers with the aim to produce PCR products smaller than the traditional ones. They obtained reduced-size amplicons by moving the forward and reverse PCR primers very close to the STR repeat region (Fig. 2.3). Due to this, the use of Mini-STRs increases the chance

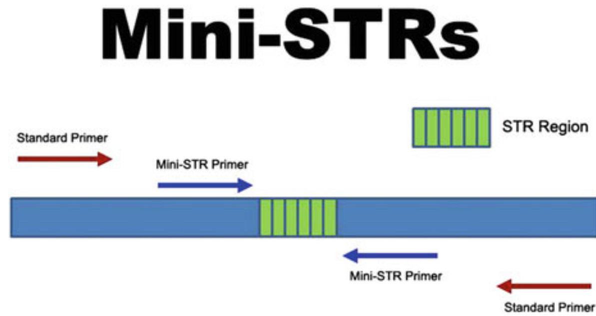


Fig. 2.3 Example of Mini-STRs

of amplification for larger loci, providing a useful tool for obtaining data from compromised samples (i.e. bones, teeth, burnt items).

In addition this technology enhances the sensitivity of DNA detection and produces results even with inhibited or scarce DNA samples such as contact traces left after a casual handling of objects ("touch DNA") or cold cases samples, previously not useful for a DNA test. Another important advantage is that DNA profiles obtained by typing Mini-STRs are compatible with DNA profiles of samples processed using traditional STR multiplexes: this allows comparison with old profiles already stored in national databases.

2.5 STRs Nomenclatures

In October 1993, the DNA Commission of the International Society of Forensic Genetics (ISFG) established the STRs nomenclature, now universally used, with the aim to encourage international reproducibility of DNA analysis and to allow comparisons of data obtained by different laboratories worldwide. It was established that DNA sequences must be read in the 5′ to 3′ direction and to consider as first repeat, the first 5′ nucleotide that defines the repeat unit.

Loci falling within a gene are named by the gene (i.e. FGA, is named after the Fibrinogen A alpha Gene) while loci not falling within genes are named with the letter "D" + the chromosome number where the locus is placed + letter "S" + the order in which the chromosomal location is described (i.e. D21S11).

Alleles name depends on the number of repeat units: if the allele contains a partial repeat, then it is named by the number of repeat units and the number of base pairs of the partial repeat (e.g. Allele 9.3). As STR typing expanded, many microvariant alleles have been observed and they have been studied in order to confirm the effective sequence (Bär et al. 1992, 1997).

The database "STRBase", freely accessible online, was created in 1997 by the Applied Genetics Group of the National Institute of Standards and Technology (NIST). It includes information about STR typing and an updated list of new alleles as they are found.

The "GenBank", created in 1982 by the *National Center for Biotechnology Information (NCBI)* is another free online database that collects all available DNA sequences with the aim to provide them to the scientific community. Generally updates are released every 2 months.

2.6 STRs Multiplexes

STRs analysis originally involved silver-stain detection of PCR amplicons after their separation by gel electrophoresis, but the introduction of a laser multicolour fluorescence detection greatly revolutionized DNA typing. To allow laser detection, in each marker, a fluorescent label is added to the 5′ end of one of the PCR primers: this permits analyzing different markers in the same PCR tube. Nowadays till to five different fluorescent dyes may be now used for labelling primers, while another different fluorescent colour is reserved for the internal size standard (Gill et al. 1995).

The ability to amplify multiple STR markers using fluorescent-labelled PCR primers together with automated laser detection increases the amount of information collected from a sample. Obviously an accurate selection of different fluorescent dye colours is required in order to avoid overlapping between PCR products with the same allele size range, to reduce spectral overlap between dye colours and to ensure full compatibility with instrumental detection. In comparison with standard silver-stain detection, less DNA sample is necessary to obtain results from multiplex amplification; in addition labour-time and cost are reduced since multiple loci are analyzed in parallel in the same assay, allowing high-throughput analysis (Kimpton et al. 1994).

In 1994, the Forensic Science Service (FSS) introduced the first-generation quadruplex that included TH01, vWA, FES/FPS, and F13A1 markers. The discrimination power of this first multiplex was approximately 1 in 10,000. Due to the low PD value several efforts were done in order to build a larger multiplex. Then, when in April 1995 the UK National DNA Database started, the *SGM* system including seven STR database core loci (VWA, D8S1179, D21S11, D18S51, TH01, FGA, D3S1358, D16S539, D2S1338 and D19S433) was developed. In October 1998 FBI launched USA National DNA Index System (NDIS) including 13 core STR loci (TH01, vWA, FGA, D8S1179, D18S51, D21S11, CSF1PO, TPOX, D3S1358, D5S818, D7S820, D13S317 and D16S539) (Werrett 1997; Budowle et al. 1998).

Starting from 1998 a second generation multiplex named *SGM Plus* was used for the UK database: SGM Plus kit had eight loci (FGA, TH01, VWA, D3S1358, D8S1179, D16S539, D18S51 and D21S11) in common with CODIS and included as additional markers D2S1338, D19S433 and Amelogenin for sex determination. The discrimination power for SGM Plus was around 1 in 13 trillion for African-Americans and 1 in 3.3 trillion Caucasian Americans, while the PI for CODIS loci was around 1.14×10^{-15} for African-Americans and 2.97×10^{-15} for Caucasians. Once core loci were established, commercial multiplex began to be developed allowing initially simultaneous amplification of few loci and then of more markers till to large multiplexes which are now commonly used in forensic laboratories.

In particular STR multiplex able to co-amplify 16 different loci (including Amelogenin) rapidly becomes the most used for solving forensic and paternity cases. In 1999, the ENFSI (European Network of Forensic Science Institutes) DNA Working proposed a European Standard Set (ESS) of seven loci (TH01, vWA, FGA, D21S11, D3S1358, D8S1179 and D18S51) as a minimum number of markers required for DNA profiles comparison between different countries. This was accepted and confirmed by the EU-Council resolution no. 2001/C 187/01.

In 2001, Interpol decided to include as Interpol Standard Set of Loci (ISSOL) the same ESS loci plus Amelogenin as sex marker. A comparison based on seven core loci was initially enough for a sporadic exchange of data, but in 2005 some countries in Europe firmed the Treaty of Prum with the aim of cross-border cooperation. In June 2008, the Treaty of Prüm was converted by the Council of the European Union into EU legislation (The EU-Prüm-Decision): this allowed all 27 EU countries to be able to exchange DNA profiles. Unfortunately, it was observed that performing a massive exchange of DNA profiles, the chance of occasional matches was greatly increased, so seven loci were not enough for large scale comparisons. In 2009, in order to reduce the chance of false positive matches, ENFSI recommended the extension of the European Standard Set of Loci including five additional loci (Gill et al. 2006).

This proposal was accepted and confirmed by a resolution of the European Council no. 2009/C 296/01. It was established to include, between the five new core loci, also three new mini-STR loci (D10S1248, D2S441, D22S1045) in order to favour the amplification of degraded/low DNA sample. In 2012, in USA, FBI proposed to expand also CODIS core loci number from 13 to 20 loci, with the aim not only to reduce false matches, but also to increase international compatibility and discrimination power in cases of missing person's identifications (Hares 2012; Butler and Hill 2012).

In March 2015, FBI published the expansion to 20 loci of the original CODIS core: as new core CODIS loci were included the five new ESS loci (D1S1656, D2S441, D10S1248, D12S391 and D22S1045) plus D2S1338 and D19S443, which are markers already commonly used worldwide. This became effective since 1st January 2017. In this perspective, in order to cover European and USA core loci, commercial companies started to produce new large multiplex able to amplify till to 27 markers (Table 2.4). Commercially STR kits are nowadays widely used either in forensic than in paternity tests because of their easy use and high discriminatory power (Guo et al. 2014; Schumm et al. 2013).

Multiplex are designed using specific softwares in order to select primers that anneal at similar temperatures and that show low affinity to each other and towards the other DNA regions different from the specific target: this in order to avoid mismatches that may to influence the stability of the primer-template combination and the PCR efficiency. Primer concentrations must be optimized for obtaining in all loci balanced signals after PCR reaction. Commercial multiplexes are easy to use because they include all reagents required for PCR amplifications such as reaction buffer with Taq Polymerase, Primer Mix, male or female Positive Controls (i.e. Control DNA 007, Control 9947, Control 9948), and Allelic Ladder. The allelic

Table 2.4 Forensic STR multiplex kits commercially available

Manufacturer	Commercial name	Production year	Number of markers
Applied Biosystems	AmpFlSTR Blue	1996	3
Applied Biosystems	AmpFlSTR Green I	1997	3
Applied Biosystems	Profiler	1997	9
Promega	Gamma STR	1997	4
Promega	PowerPlex 1.1	1997	8
Applied Biosystems	Profiler Plus	1997	10
Applied Biosystems	COfiler	1998	7
Promega	PowerPlex 1.2	1998	10
Applied Biosystems	SGM Plus	1999	11
Promega	PowerPlex 2.1	1999	9
Promega	PowerPlex 16	2000	16
Applied Biosystems	Identifiler	2001	16
Applied Biosystems	Profiler Plus ID	2001	10
Applied Biosystems	SEfiler	2002	11
Promega	PowerPlex ES	2002	9
Promega	PowerPlex Y	2003	12
Applied Biosystems	Yfiler	2004	17
Applied Biosystems	MiniFiler	2007	8
Applied Biosystems	SEfiler Plus	2007	11
Promega	PowerPlex S5	2007	4
Applied Biosystems	Sinofiler	2008	15
Promega	PowerPlex CS7	2009	7
Applied Biosystems	Identifiler Direct	2009	16
Applied Biosystems	NGM	2009	16
Promega	PowerPlex 16 HS	2009	16
Promega	PowerPlex ESX 16	2009	16
Promega	PowerPlex ESX 17	2009	17
Promega	PowerPlex ESI 16	2009	16
Promega	PowerPlex ESI 17	2009	17
Applied Biosystems	Identifiler Plus	2010	16
Applied Biosystems	NGM Select	2010	17
Qiagen	Nonaplex ESS	2010	9
Qiagen	ESSplex	2010	12
Qiagen	ESSplex SE QS	2010	13
Qiagen	Decaplex SE	2010	10
Qiagen	Idplex Plus GO	2010	16
Qiagen	Idplex Plus	2010	16
Qiagen	Hexaplex ESS	2010	10
Qiagen	HDplex kit	2010	11
Qiagen	Triplex AFS QS	2010	2
Qiagen	Triplex DSF	2010	3
Qiagen	Argus X-12 QS	2010	12

(continued)

Table 2.4 (continued)

Manufacturer	Commercial name	Production year	Number of markers
Promega	PowerPlex 18D	2011	18
Applied Biosystems	GlobalFiler	2012	24
Promega	PowerPlex Y23	2012	23
Promega	PowerPlex 21	2012	21
Promega	PowerPlex Fusion	2012	24
Applied Biosystems	GlobalFiler Express	2014	24
Applied Biosystems	YFiler plus	2014	27
Promega	PowerPlex ESX/ESI Fast	2014	16/17
Promega	PowerPlex Fusion 6C	2015	27
Qiagen	Investigator 24plex GO	2015	24
Qiagen	Investigator 24plex QS	2015	24
HealthGene Technologies	SureID® 21G	2015	21
HealthGene Technologies	SureID® 27Y	2015	27
Applied Biosystems	NGM Detect	2016	18
Applied Biosystems	VeriFiler™ Express	2016	25
HealthGene Technologies	SureID® PanGlobal	2018	27
Applied Biosystems	GlobalFiler IQC	2019	24

ladder is used for genotype determination: it's a mixture of all the alleles available for a particular STR marker and hence it's artificially obtained by combining DNA from individuals which have different alleles, representatives of the locus variations in the population. According to the recommendations of the ISFG DNA Commission, allelic ladder sequencing is required. Alleles from investigated loci are assigned by a comparison with ladder alleles even if rare size variants may be found, showing intermediate mobility respect to ladder alleles. A representative STR DNA profile is given in Fig. 2.4.

Before to introduce a new kit, obviously manufacturers spend time and resource in research, thus the advantage for forensic labs in the use of commercial multiplexes, is the possibility to work with robust STR multiplex assays, in already optimized and validated reaction protocols using carefully designed and well balanced primers, labelled in adequate different fluorescent-dye combinations.

2.7 Validation Studies

Validation is "a process by which a procedure is evaluated to determine its efficacy and reliability for forensic casework and/or database analysis" (SWGDAM 2010). The Scientific Working Group on DNA Analysis Methods (SWGDAM) includes around 50 forensic experts from several Federal, State and Local DNA laboratories

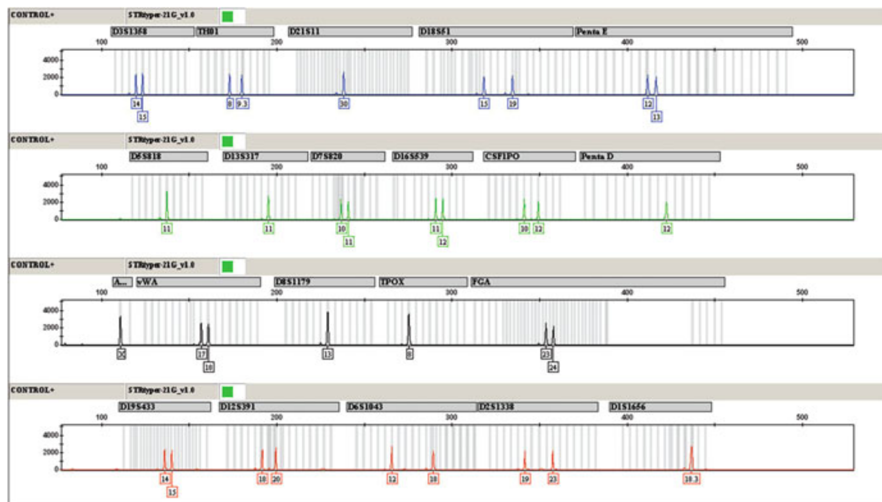


Fig. 2.4 Control DNA9948 amplified by Sure ID 21G

in the USA and Canada. They developed the guidelines for the validation of DNA analysis procedures, in accordance with the FBI Director's Quality Assurance Standards (QAS) (FBI 2004).

Initially, in 1998, the Federal DNA Advisory Board recommended quality assurance standards to the FBI Director and then the SWGDAM (that replaced TWGDAM) was charged with this responsibility. Two types of validation (each one involving several studies) are required to assess new forensic DNA technologies: developmental and internal. Obviously the laboratory should evaluate which validation studies are more relevant for the procedure tested and it should establish the right number of samples required for each study (SWGDAM 2010, 2001).

2.7.1 Developmental Validation

Guidelines describe the main studies that should be performed by the manufacturer during the developmental validation process in order to evaluate if methods are effectively reliable and to establish which are the conditions and the limitations of the new DNA methodology prior to be effectively introduced in forensics. According to SWGDAM recommendations, main studies include: Species Specificity, Sensitivity Studies, Precision and Accuracy (Repeatability and Reproducibility), Mixture Studies. In addition the method performance on some caseworks-like samples should be evaluated and population studies should be done to create a population database for the main populations groups (Table 2.5).

Table 2.5 Main studies required for developmental validation (SWGDAM 2010)

Study	Description
Species specificity	Ability to detect nonhuman DNA
Sensitivity	Evaluation of the assay range useful to obtain reliable results and determination of the upper and lower detection limits
Stability	Ability to obtain results from biological samples placed on different substrates
Precision	Degree of agreement among a series of measurements, values or results
Accuracy	Degree of conformity of a measured quantity to its real value
Repeatability	Precision and accuracy of results of the same operator or instrument
Reproducibility	Precision and accuracy of results among different operators or instruments
Mixture studies	Ability to obtain reliable results from mixed samples
Case-type samples	Ability to obtain reliable results using samples that are representative of real laboratory samples.
Population studies	Distribution of genetic markers in relevant population groups

2.7.2 Internal Validation

Before applying in routine caseworks a new forensic procedure, a laboratory is required to perform internal validation studies with the aim to accumulate data able to demonstrate that a method works according to the laboratory expectations. All data must be accurately documented. According to SWGDAM recommendations, the internal validation process includes the following studies: Sensitivity and Stochastic Studies, Precision and Accuracy (Repeatability and Reproducibility), Mixture Studies, Contamination Assessment (Table 2.6). In addition, the analysis of known samples, non-probative, evidence samples or mock evidence samples is required.

2.8 Applications of DNA Typing

Since the first introduction of DNA typing in 1987, in the last 30 years an enormous growth has occurred in the use of DNA test either in criminal investigations than in paternity testing, missing person's identification, mass disasters (DVI). In particular, the use of the polymerase chain reaction (PCR) and of large multiplexes allows to analyze samples in a short time, offering the possibility to obtain useful DNA profiles and conclusive results from a wide range of biological evidences even if degraded or in scarce quantity.

Because of this, annually forensic laboratories perform hundreds of test and DNA typing represents the most powerful tool for identification.

In fact main applications of DNA typing include:

Table 2.6 Main studies required for internal validation (SWGDAM 2010)

Study	Description
Sensitivity	Determination of the target range, limit of detection, heterozygote balance (e.g. peak height ratio) and the signal to noise ratio.
Stochastic	Evaluation of stochastic effects obtained from the analysis of low quantity/quality DNA samples.
Precision	Degree of agreement among a series of measurements, values or results
Accuracy	Degree of conformity of a measured quantity to its real value
Repeatability	Precision and accuracy of results of the same operator or instrument
Reproducibility	Precision and accuracy of results among different operators or instruments
Mixture studies	Ability to obtain reliable results from mixed samples
Contamination	Detection of exogenous DNA in reagents, consumables, samples and laboratory environment
Known, non-probative or mock evidence samples	Results from these samples should be compared with data previously obtained from the same samples to ensure concordance

1. **Forensics Caseworks**
   Biological samples are recovered at crime scene and genotyped: evidences DNA profiles are compared between them or with the ones of known people (i.e. victim, suspect) or with data in national DNA databases in order to find a match useful to identify the evidence donor and to link different caseworks.
2. **Personal Identification**
   DNA test is the best method to establish the identity of a person, especially in cases of mass disasters (DVI) or missing persons identification when only parts of bodies (i.e. bones, teeth) are available.
3. **Biological Relationship**
   Due to the inheritance modality, DNA analysis is routinely used for paternity/maternity test, kinship analysis as well as to reconstruct complex relationships or entire family pedigree.

References

Bär W, Brinkmann B, Lincoln P, Mayr W, Rossi U, Budowle B, Eisenberg A, Fourney R, Gill P, Rand S (1992) Editorial: recommendations of the DNA Commission of the International Society for Forensic Haemogenetics relating to the use of PCR-based polymorphisms. Forensic Sci Int 55:1–3

Bär W, Brinkmann B, Budowle B, Carracedo A, Gill P, Lincoln P, Mayr W, Olaisen B (1997) DNA recommendations. Further report of the DNA Commission of the ISFG regarding the use of short tandem repeat systems. Forensic Sci Int 87(3):179–184

Budowle B, Moretti TR, Niezgoda SJ, Brown BL (1998) CODIS and PCR-based short tandem repeat loci: law enforcement tools. In: Proceedings from the Second European Symposium on Human Identification

Butler JM (2001) Forensic DNA typing: biology, technology, and genetics of STR. Academic Press, Elsevier

Butler JM (2012) Advanced topics in forensic DNA, typing methodology. Academic Press, Elsevier

Butler JM, Shen Y, Mc Cord BR (2003) The development of reduced size STR amplicons as tools for analysis of degraded DNA. J Forensic Sci 48:1054–1064

Butler JM, Hill CR (2012) Biology and genetics of new autosomal STR loci useful for forensic DNA analysis. Forensic Sci Rev 24:15

Carracedo A, Lareu MV (1998) Development of new STRs for forensic casework: criteria for selection, sequencing and population data and forensic validation. In: Proceedings of the Ninth International Symposium on Human Identification, pp 89–107

Coble MD, Butler JM (2005) Characterization of new MiniSTR loci to aid analysis of degraded DNA. J Forensic Sci 50:43–53

FBI (2004) Quality assurance standards for forensic DNA testing laboratories. Revised validation guideline—scientific working group on DNA analysis methods (SWGDAM). Forensic Sci Commun 6(3):1

Gill P, Kimpton CP, Urquhart A, Oldroyd N, Millican ES, Watson SK, Downes TJ (1995) Automated short tandem repeats (STR) analysis in forensic casework - a strategy for the future. Electrophoresis 16:1543–1552

Gill P, Fereday L, Morling N, Schneider PM (2006) The evolution of DNA databases. Forensic Sci Int 156(2–3):242–244

Guo F et al (2014) Development of a 24-locus multiplex system to incorporate the core loci in the Combined DNA Index System (CODIS) and the European Standard Set (ESS). Forensic Sci Int Genet 8(1):44–54

Hares DR (2012) Expanding the CODIS core loci in the United States. Forensic Sci Int Genet 6: e52–e54

Hill CR, Kline MC, Coble MD, Butler JM (2008) Characterization of 26 MiniSTR loci for improved analysis of degraded DNA samples. J Forensic Sci 53:1

Jeffreys AJ, Wilson V, Thein SW (1984) Hypervariable ‘minisatellite’ regions in human DNA. Nature 314:67–73

Jeffreys AJ, Wilson V, Thein SL (1985a) Individual-specific ‘fingerprints’ of human DNA. Nature 316:76

Jeffreys AJ, Brookfield JF, Semeonoff R (1985b) Positive identification of an immigration test-case using human DNA fingerprints. Nature 6:317

Jin L, Zhong Y, Chakraborty R (1994) The exact numbers of possible microsatellite motifs. Am J Hum Genet 55:582–583

Kimpton C, Fisher D, Watson S, Adams M, Urquhart A, Lygo J, Gill P (1994) Evaluation of an automated DNA profiling system employing multiplex amplification of four tetrameric STR loci. Int J Legal Med 106:302–311

Mullis K (1990) The unusual origin of the polymerase chain reaction. Sci Am 262(4):56–61. 64

Schumm JW, Gutierrez-Mateo C, Tan E, Selden R (2013) A 27-locus STR assay to meet all United States and European law enforcement agency standards. J Forensic Sci 58(6):1584–1592

Scientific Working Group on DNA Analysis Methods (2010) Interpretation guidelines for autosomal STR typing by forensic DNA testing laboratories. Approved at the Scientific Working Group on DNA Analysis Methods meeting, Fredericksburg, Virginia, January 2010

Scientific Working Group on DNA Analysis Methods (SWGDAM) (2001) Training guidelines. Forensic Sci Commun 3(4):1

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503–517

Sparkes R, Kimpton C, Watson S, Oldroyd N, Clayton T, Barnett L, Arnold J, Thompson C, Hale R, Chapman J, Urquhart A, Gill P (1996a) The validation of a 7-locus multiplex STR test

for use in forensic casework (I) Mixtures, ageing, degradation and species studies. Int J Legal Med 109:186–194
Sparkes R, Kimpton C, Gilbard S, Carne P, Andersen J, Oldroyd N, Thomas D, Urquhart A, Gill P (1996b) The validation of a 7-locus multiplex STR test for use in forensic casework (I) Artefacts, casework studies and success rates. Int J Legal Med 109:195–204
Werrett DJ (1997) The national DNA database. Forensic Sci Int 88:33–42

3 Sequential Advancements of DNA Profiling: An Overview of Complete Arena

Kriti Nigam, Ankit Srivastava, Subhasish Sahoo, I. P. Dubey, I. P. Tripathi, and Pankaj Shrivastava

Abstract

Discovery of DNA profiling technology has led the forensic investigations to another level of confidence. This technique is among the utmost discoveries of twentieth century that has revolutionized the criminal justice system. This chapter briefly recapitulates the sequential progressions made in the discipline of forensic DNA fingerprinting which aids the justice system in multiple ways by making it far more efficient in comparison to the existing conventional techniques. Right from the discovery of this substantial technique, current capillary electrophoresis based methods using autosomal STRs along with lineage markers (Y STRs, X-STRs, mtDNA) are covered here along with an insight to the latest advancements including the next generation sequencing (NGS).

Keywords

DNA profiling · STRs · DNA markers · NGS

K. Nigam · A. Srivastava
Dr. APJ Abdul Kalam Institute of Forensic Science and Criminology, Bundelkhand University, Jhansi, UP, India

S. Sahoo
DNA Profiling Unit, State Forensic Science Laboratory, Bhubaneswar, Odisha, India

I. P. Dubey
DNA Profiling Unit, State Forensic Science Laboratory, Bhubaneswar, Odisha, India

Faculty of Science and Environment, MGCGV, Satna, Madhya Pradesh, India

I. P. Tripathi
Faculty of Science and Environment, MGCGV, Satna, Madhya Pradesh, India

P. Shrivastava (✉)
DNA Fingerprinting Unit, State Forensic Science Laboratory, Department of Home (Police), Government of Madhya Pradesh, Sagar, Madhya Pradesh, India

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_3

3.1 Introduction

The most magnificent breakthrough of late twentieth century that revolutionized scientific community and gave an exceptional vigor to criminal justice system is DNA profiling. As name is suggesting, DNA profile is the profile of a person generated by using its DNA. This is pretty much similar to the barcodes printed on different products available in a super market that barcode reader reads to give details about that particular product. DNA typing or DNA fingerprinting are the terms that are synonymously used to refer this technique. The credit of this vital discovery goes to Sir Alec Jeffreys, a geneticist from University of Leicester, UK. During September 1984, while studying on tentative ways to sort out immigration and paternity controversies by signifying the genetic linkage among different persons, Sir Jeffreys employed restriction fragment length polymorphism (RFLP) for analyzing DNA. He noticed that DNA has certain repetitive sequences, commonly known as minisatellite or variable number of tandem repeats (VNTRs) that are perpetually present in all human beings. However, their length varies among individuals (Jeffreys et al. 1985a). He utilized these variations in establishing the identity of a person. He soon realized that these variations are unique to each individual except for identical twins and thus, could be used to affirmatively ascertain the identity and individuality of a person (Jeffreys et al. 1985b). Eventually this ground breaking discovery of Dr. Alec Jeffreys emerges out as a potential tool for exonerating the innocent and incarcerating the guilty. Present chapter gives a chronological over view of different DNA profiling systems.

3.2 Restriction Fragment Length Polymorphism (RFLP)

RFLP is a pioneer technique of DNA fingerprinting. It employs a molecular method for DNA analysis i.e., it permits identification of different individuals on the basis of exclusive patterns generated by single restriction enzyme nicking DNA at peculiar sites commonly referred to as restriction endonuclease recognition sites. Polymorphic nature of genetic codes is the key factor behind this technique. Nicked DNA is then separated into different fragments using gel electrophoresis depending on relative sizes of generated fragments. Since DNA carries an overall negative charge, different fragments have a tendency to move toward positive pole. Smaller fragments will obviously move faster. Gel with these bands is then denatured by placing it in sodium hydroxide (NaOH) solution. The single-stranded DNA so obtained is transferred into a nylon or nitrocellulose sheet by capillary blotting, named after its inventor as southern blotting (Southern 1975). This sheet is then fixed by autoclaving. This single-stranded fragmented DNA on nylon or nitrocellulose sheet is then allowed to base pair with the labeled RFLP probes. These probes may be labeled with radioactive or chemiluminescent tag (Klevan et al. 1995). This process is commonly referred to as hybridization. When such hybridized DNA is exposed to X-rays, it produces distinct bands on X-ray film. The autogram so generated is the unique genetic signature of an individual with an exception of

monozygotic twins who have same autogram. Shortly after its development, in 1986, the technique of DNA profiling was first utilized in solving a criminal case. Colin Pitchfork, a U.K. resident, was convicted of a double rape and murder because his DNA profile matched with DNA found at both crime scenes (Jobling and Gill 2004). Through RFLP, paternity can be excluded with a cumulative probability greater than 99.9% with the use of as few as four probes (Jeffrys et al. 1987) likewise 3–5 probes can provide an individualized fingerprint in forensic testing (Alghanim and Almirall 2003). Sequences of VNTR serve as the basis for fingerprinting. For many years RFLP technique served criminal justice system by resolving paternity disputes, ascertaining criminal identity, studying evolution and migration of wildlife, studying breeding patterns in animal populations and the detection, etc. However, requirement to large quantity of good quality sample, being expensive, tedious and time consuming, were certain limitations associated with this technique. This leads to the need of discovery of more efficient and advance techniques of DNA profiling.

3.3 Polymerase Chain Reaction (PCR)

Before discussing various PCR-based DNA profiling techniques, it is highly desirable here to discuss the basic conception of PCR first. This technique became a response to various limitations of existing RFLP technique. Even small and degraded DNA samples which deemed to be unfit for RFLP could be analyzed using this technique. PCR is an excellent technique for replicating single-stranded DNA from a template with the help of synthetic primers and a polymerase enzyme (Kleppe et al. 1971; Panet and Khorana 1974). Credit of this exceptional discovery goes to Kary Mullis who, in 1983 while working with Cetus Corporation, gave a very simple idea of making millions of replicates of desired DNA fragment with high fidelity (Mullis and Faloona 1987; Mullis et al. 1986). In the year 1993, Kary Mullis was awarded Nobel Prize in chemistry for his invention. PCR amplification is a chain reaction of series of three steps repeated in a sequential manner.

3.3.1 Denaturation

This is the first step of PCR amplification in which helix of DNA unwinds to separate the two strands of DNA by raising the temperature up to 98 °C. Now each strand acts as a template for synthesizing new strand. Ingredients of this step are a primer (like probes of RFLP), base pairs and a heat stable DNA polymerase usually Taq polymerase.

3.3.2 Annealing

During this step, reaction mixture's temperature is decreased to 50–65 °C, this allows annealing between the single-stranded DNA templates and the primers.

Primers have been named so because they mark the initial location for the synthesis of new strand of DNA.

3.3.3 Elongation

In this step, DNA polymerase came into play. Temperature of reaction mixture is again raised to approximately 75–80 °C and Taq polymerase add new nucleotides to the primer complementary to parent strand.

In order to ascertain whether the desired DNA fragment has been amplified or not (also known as amplicon or amplimer), agarose gel electrophoresis may be performed. Results are then compared with molecular weight markers which contain fragments of DNA of identified size.

3.4 PCR-Based Techniques

3.4.1 HLA DQA1

HLA stands for human leukocyte antigens—a group of proteins, located on the outer membrane of all the nucleated cells of our body. Genes of these antigens are positioned on sixth chromosome. Major histocompatibility complex (MHC) is the generic name of the genetic province to which HLA loci belong (Hugh et al. 1984). The MHC contains genes (including HLA) that are liable for immune responses of body. HLA antigens perform vital task in self-recognition and discriminating self from non-self. HLA antigens thus, crucially strategize defense mechanism against foreign substances. Every individual possess a unique set of HLA proteins inherited from his parents. The first typing method that was used to detect HLA alleles was—the hybridization of sequence specific oligonucleotide (SSO) probes with PCR-amplified DNA. Researchers have suggested various alternative SSO typing methods, chiefly varying on the length, sequence of probes, and their methods of detection. Initially, P^{32} SSO probes were used to hybridize with the desired region of HLA-DQA gene (Saiki et al. 1986a), but shortly afterwards biotin emerge out as an alternative of conventional P^{32} label (Saiki et al. 1987). The PCR-SSO technique was later applied to other allelic sequences also like DP (Bugawan et al. 1988), DQ (Horn et al. 1988; Morel et al. 1990) and DR (Erlich et al. 1989; Baxter-Lowe et al. 1989).

Although SSO technique is appropriate for investigation of large sample sizes but it is not a method of choice for analysis of small number of samples. Reverse dot blotting technique emerged out as a solution of this problem. Procedurally, it employs same PCR amplification procedure and SSO probes, but these probes are adhered to a solid support medium. The sample so amplified is labeled and hybridized on that supporting panel. A single hybridization and stringency wash permits the determination of polymorphic sequences of the selected sample. Conjugate streptavidin horseradish peroxidase is used to detect positive reaction by having

a colored soluble substrate (Buga et al. 1990). Alternatively, chemiluminescence could also be employed for visualization (Buyse et al. 1993).

HLA-DQA was the foremost commercially available PCR-based kit for DNA fingerprinting. This kit had a potential to differentiate six different alleles i.e., 1.1, 1.2, 1.3, 2, 3, and 4. Initially HLA-DQA strips contain nine SSA probes. Perkin-Elmer later introduced a more sophisticated kit DQA1 with 11 SSO probes.

3.4.2 Amplified Fragment Length Polymorphism (AFLP)

AFLP is a fingerprinting technique based on PCR, which was explained for the first time by Vos et al. (1995). Then after, many protocols have been declared from time to time suggesting various modifications, all including following three basic steps:

(1) restriction endonucleases are employed in digesting genomic DNA, forming DNA fragments along with its ligation to double-stranded adaptors with known adaptor sequences, comprising of two nucleotides; (2) these DNA fragments are specifically amplified using primers complementary to adaptor sequences and the unknown genomic DNA; and (3) labeled fragments are separated by electrophoresis followed by silver staining (Fry et al. 2009). Fragmented samples are then subjected to automated analysis. Locus D1S80, with repeat size of 16 bp, is a popular choice for DNA fingerprinting (Thymann et al. 1993). However, other available choices include YNZ22, Apolipoprotein-B, and Collagen 2A1 (De Guglielmo et al. 1994). AFLP profiling can discriminate between illicitly grown marijuana and hemp (Coyle et al. 2003; Datwyler and Weiblen 2006; Hakki et al. 2003), identify illegal hallucinogenic fungi (Coyle et al. 2001; Lee et al. 2000; Linacre et al. 2002), legally protected owls species and their hybrids (Haig et al. 2004). This automation is the biggest advantage of this technique allowing a trouble free comparison of DNA samples in an economic manner.

3.4.3 Short Tandem Repeats (STRs)

In spite of the triumph of PCR-based AFLPs, in 1990s there was a switch to PCR of short tandem repeats (STRs); a technique that utilizes too small repeat units, with a length of just 2–7 bp. STRs are named as microsatellites or simple sequence repeats. Since their discovery, they are often referred as 'gold standard' in personal identification in forensic context. STRs alternatively known as microsatellites are short repeats of DNA sequences having repetitive units ranging from 1 to 6 bp (Chambers and MacAvoy 2000; Tautz 1993), entire length ranging up to 100 nucleotides. STRs are usually located near centromere of chromosome. Another category of these sequence repetitions is not so long i.e., it consists of 2–6 bp repeats only—popularly known as minisatellites. This small length of repeat unit makes them capable of easier amplification. Additionally, they also have a relatively higher resistance to the problems associated with degraded and contaminated DNA. Nucleotide sequence mutation might be the probable reason of these variations. STRs are present in both

pro and eukaryotes, including human beings. They are somewhat evenly scattered all through human genome, consisting about 3% part of the entire genomic structure. But within chromosomes, they are not so uniformly distributed (Koreth 1996). Exploration of noncoding regions revealed that majority of STRs is found in those regions which do not code for any protein. Only about 8% of STRs are positioned in the coding regions (Ellegren 2000). Length of such repeated units vary among individuals which is their most significant aspect from forensic point of view. Not only their length but their densities also vary among chromosomes. Nineteenth chromosome is found to have highest density of STRs in humans (Subramanian 2003). An individual can be either homozygous i.e., with same number of repeat units or heterozygous i.e., with different number of repeat units at a certain locus. Tetranucleotide repeats, abundant in Adenine (A) are used in forensic genotyping (Goodwin et al. 2011; Nadir 1996). DNA Commission of the International Society of Forensic Genetics (ISFG) in 1993 assigned nomenclature to STR loci and its allelic variants. The STR loci are named in a peculiar manner, for example, D1S80, where D refers to DNA, 1 signifies the chromosome number where on STR locus is located, S refers to STR, and 80 is its exclusive identifier signifying position of the repeat i.e., 80th on the 1st chromosome. Name of STRs are given on the basis of base composition of the repeat unit (in parenthesis). After that the number of times of its repetition is mentioned in subscript, e.g., $(GTA)_4$, a trinucleotide STR having sequence GTA is repeated 4 times. The first STR was discovered in the year 1991. In 1994 only four loci namely, F13A1, THO1, Vwa and FES/FPS were available (Kimpton et al. 1994; Lygo et al. 1994) but within a year this number got extend to seven (Gill et al. 1997; Kimpton et al. 1996; Sparkes et al. 1996a, b). By the end of 2000, number of available STR multiplex increased to 16 including amelogenin—a sex determining marker (Krenke et al. 2002). Today more than 20 STR multiple systems are available commercially for DNA typing. STR is found to be more sensitive in comparison to conventional single-locus RFLP methods and more discriminating when compared to other PCR-based profiling techniques, like HLA-DQA1 (Saiki et al. 1986b). Since then several researchers have been conducted researches in this field studying different populations for developing different standardized protocols across the globe (Butler 2005). Main advantages of the STRs markers are that by multiplexing, they can test more than ten STR loci in fast, simple, and simultaneous way (Morretti et al. 2001). Classification of STRs is made on the basis of length of their repeat units i.e., mononucleotides, dinucleotides, trinucleotides, and so on. However, tetranucleotides are utilized most due to the reason of minor stutter products probability, amplicons having one repeat less than true allele (Romeika and Yan 2013). In multiplex analysis kits, tetra- and pentanucleotide systems are also incorporated due to their ability to offer results with an augmented exclusion index. Multiplexing aids simultaneous analysis of several different loci. This does not merely save time, but also saves the evidences by utilizing a lesser sample size. Current trends of DNA profiling use multi-allelic STR markers which not only have structural comparability with original minisatellites but have much shorter repeats which makes them easier to amplify using PCR. The Federal Bureau of Investigations (FBI) in 1997 established a

database named CODIS (Combined DNA Index System) that integrated amelogenin sex loci and 13 autosomal loci in the process of development of STR Typing System (Sullivan et al. 1993). These 13 STR loci set have efficient distinguishing abilities having an average probability of random match one in a quadrillion (1×10^{-15}), which further extended to 20 markers (Budowle et al. 1998) while the standard set of Europe consisted of 12 STR markers (Gill et al. 2006). These loci are located on different chromosomes in form of highly polymorphic noncoding regions. Together with amelogenin, new multiplexes have been introduced in a single reaction within past few years that can amplify even more than 16 loci (Dixon et al. 2005). STR profiling has numerous applications including parentage testing, identification of disaster victims, rape felon's identification, and many more (Yoshida et al. 2011). Butler et al. offered a summary of each STR marker (24 forensic autosomal STR loci) including their classification, length of repeat unit, chromosomal location, region of STR repeats, etc. (Butler and Hill 2012).

The application of STR multiplexing method to the biological evidences has changed the entire scenario of our justice system dealing with different heinous crimes. Multiplex PCR with fluorescently labeled primers has become a vital method adopted for short tandem repeat's amplification having applicability in testing individual identity. STRs have popularly become the preferred markers for various applications of human identification (HID) due to their unique and polymorphic nature (Butler 2006). Measurement of length of different variable alleles is the foundation of individualization by STR profiling (Giardina et al. 2011). Due to its higher accuracy rate, STR Marker technology has become applicable in routine criminal investigations. In fact, this is one among the most responsive and extensively acknowledged techniques of scientific community. In the investigative procedure of various offences like manslaughter, assassination, rape, kidnapping, etc., STR profiling is a frequent technique today. But the capillary electrophoresis (CE) based DNA fragment size detection technique is not capable of distinguishing between alleles of different sequences but of similar lengths in complex paternity cases based on STR mutation. The upcoming NGS technology can discriminate these alleles and also helps in identifying mixed samples (Berglund et al. 2011).

3.5 Lineage Markers

Autosomal STRs are greatly used by forensic science laboratories for human identification as well as for human kinship determination. However, in some complex cases based on kinship relationships, the results of autosomal STRs may be inconclusive. Consequently, the analysis of STRs located on sex chromosomes came into play (Roewer 2003).

3.5.1 Autosomal Single Nucleotide Polymorphisms (SNP) Typing

SNPs have a lower heterozygosis in comparison to STRs. Where STRs needs a template size of 300 bp, the size of template is just 50 bp in case of SNP typing. Thus SNPs have become significant tools in degraded samples analysis. SNP typing played a crucial role in identification of victims of World Trade Center disaster which took place in 2001 (Brenner and Weir 2003; Marchi 2004). The European Network of Forensic Science Institutes (ENFSI) and the US FBI Scientific Working Group on DNA Methods (SWGDAM) are functioning on recommendations concerning standardization of SNPs and its use in degraded biological samples (Gill et al. 2004).

3.5.2 Y-STR Typing

Y-STRs are short tandem repeats located on Y chromosome and thus, are male specific. Polymorphic nature of Y chromosome plays a crucial role in differentiating unrelated males as well as proving lineage in paternal generations (Coble et al. 2009). Forensic significance of these Y-STRs lies in the fact that besides being male specific, they have a low mutation rate and have specific population allele distribution (Kayser et al. 2001, 2005, 2007; Kayser 2003). Y-STR is a brilliant way to recognize male DNA which often gets mixed with female DNA in offences involving sexual assaults (Kayser 2017; Roewer 2009; Prinz et al. 1997). In such instances, autosomal STR typing is often complicated or sometimes even unfeasible to give conclusive results. Detection capabilities of Y-STR for male DNA mixed with female DNA samples is also far above any discrepancy i.e., 1:2000 (Ballantyne and Kayser 2013). Presently existing commercial Y-STR m are much vigorous and even male mixed samples (Gang rape cases), samples contaminated with different biological fluids, aged and degraded samples could be analyzed successfully (Brenner 2010). Not only fixing the guilty, Y-STR typing is also important in exonerating the innocent in false accusations. Despite of numerous advantages, Y-STR possesses some limitations as well i.e., it represents a single locus and it requires large databases for study (Willuweit et al. 2011; Ballentyne et al. 2010). Inability of exclusion of paternally related suspects lowers its credential in comparison to autosomal STRs. However, the possible answer of this limitation emerges out in terms of rapidly mutating Y-STRs (RM Y-STRs). These STRs have high rates of mutation and thus, can potentially differentiate paternally related suspects (Ballentyne et al. 2012; Roewer et al. 1992). If we go back in history of Y-STRs, it was first described by Rower and Epplen in 1922 (Davis et al. 2013). Since then several advancements have taken place. Various Y-STR kits are now available commercially. However, two most commonly used ones are PowerPlex® Y23 (Promega Corporation) and Yfiler® Plus PCR Amplification Kit (Thermo Fisher Scientific, USA). PowerPlex® Y23 employs 17 traditional Y-STR loci along with six new loci including DYS481, DYS533, DYS549, DYS570, DYS576, and DYS643, thereby enhancing its discriminating power (Jain et al. 2016; Thompson et al. 2013).

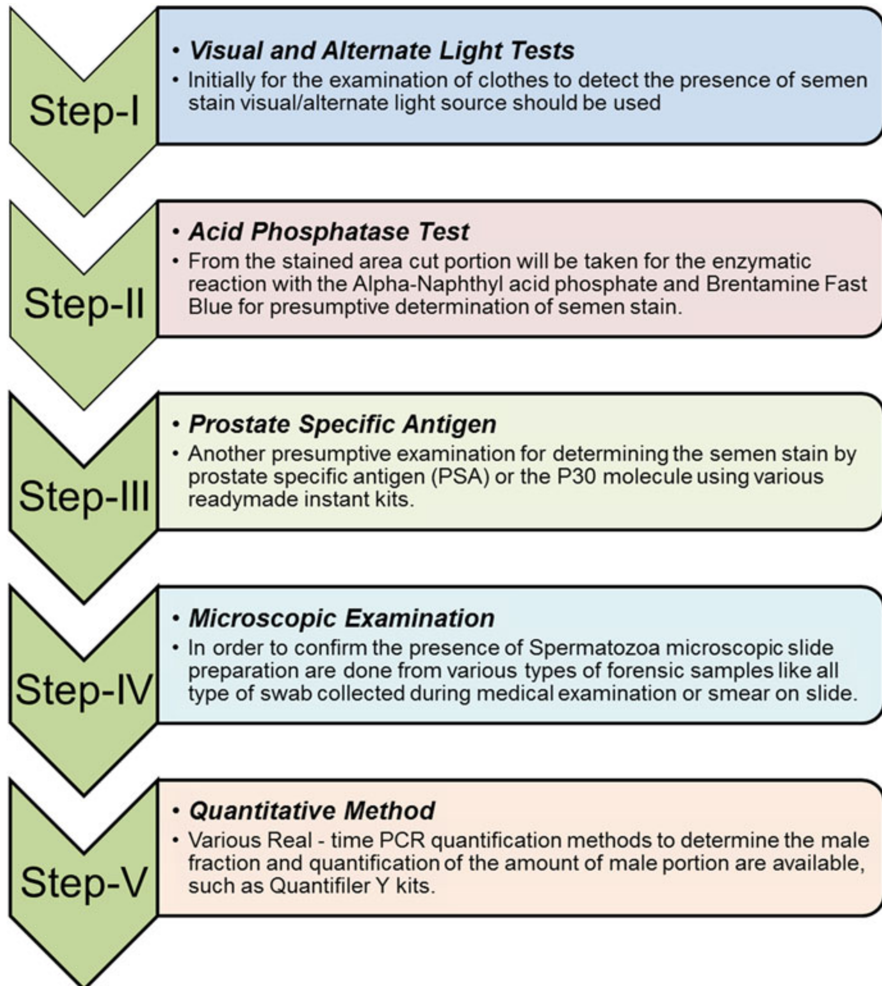


Fig. 3.1 Steps involved in Y screening for determining the male fractions

On the other hand, Yfiler® Plus PCR Amplification Kit allows multiplex amplification of 27 Y-STRs including seven RM Y-STRs (Fig. 3.1).

3.5.3 Mitochondrial DNA Typing

Cytoplasm of almost all eukaryotic cells contains mitochondrial DNA (mtDNA)—a separate genome apart from nuclear genome (Hameed et al. 2015). mtDNA a 16-kb circular molecule present in multiple copies. mtDNA is maternally inherited and thus, could easily prove linkage between maternally related people. It codes for

enzymes and proteins required for proper functioning of mitochondria. It codes for 13 proteins, 2 rRNA, and 22 tRNA. Mitochondrial genome imparts its forensic significance due to its stability and multiple copies. mtDNA in mammals has two different points of origin of replication. The point of origin of heavy chain is usually guanine rich and is located within D-loop (Displacement Loop). However, in contrast to this, point of origin of light chain is usually cytosine rich and is located nearly opposite to D-loop. The current region of forensic interest is D-loop (Pai et al. 1997). Approximately 1100 bp constitute the noncoding hypervariable regions of D-loop. This region is named so because of its higher rate of mutation in comparison to other regions of mtDNA (Quintans et al. 2004). This region is further subdivided into HV1, HV2, and HV3 regions with position range from 16,024 to 16,365, 73 to 340, and 438 to 574, respectively. From forensic point of view HV3 is rarely targeted, commonly HV1 and HV2 are examined. mtDNA obtained from calcified tissues and hair could be a potential tool of personal identification especially in cases of mass disasters where evidences obtained are often environmentally challenged (Kavlick et al. 2011). Despite all aforementioned advantages, the biggest challenge associated with the analysis of mtDNA is the interpretation of heteroplasmy. It refers to existence of different mtDNA haplotypes within a person (Bar et al. 2000). However, many researchers consider this phenomenon as a rare event (Payne et al. 2013). Various degraded samples can be analyzed by mitochondrial DNA Analysis like the hairs without root, teeth, skeletonized remains and various types of biological material which are degraded environmentally which cannot be processed by nuclear DNA Analysis. From very less quantity of samples, DNA can be detected and thus, maternal lineage can be established.

3.5.3.1 Preparation of the Sample

The first step in the sample preparation is to clean the sample properly in order to remove contaminants. Bone and teeth samples are need to be preprocessed by sanding or scrapping to remove the dirt or other unwanted debris attached to it. Similarly, hair samples need to be cleaned by sonication for removal of any dust particles or microbial growths. Hair or hair shaft region is used for mitochondrial analysis.

3.5.3.2 Extraction of Sample

The samples are extracted using C_6H_5OH/CCl_3 or other basic chemicals in order to break up the DNA from different cofactors, proteins, and ions. Further it is purified to get pure DNA sample. Various commercial extraction kits are also available for extraction of DNA in very less time. Then the samples are quantified using various quantification methods and amplified by subjecting to polymerase chain reaction using HVR-I and HVR-II primers.

3.5.3.3 Sequencing of mtDNA

PCR-based mtDNA typing through automated sequencing has proven to be a legitimate, robust, and dependable means of forensic analysis. Applied Biosystem's SOLiD sequencing by ligation is very famous and practiced by various laboratories

(Sheshanna et al. 2014). This method is utilized for entire mtDNA sequencing coupled with targeted re-sequencing. Method incorporates preparation of sample, emulsion PCR and preparation of substrate, basic chemistry of ligation along with imaging and data analysis. Solid sequencing is an encoding system consisting of two bases having an accuracy rate of 99.94% (Kircher and Kelso 2010), adequate for abundant samples to identify SNPs. For last decade, next generation sequencing (NGS) by Illumina has become widely acceptable. NGS provides a unified platform to prevail over the extensive range of complicated samples encountered. Genetic analysis of evidences of mass fatality with NGS let scientists to take out vital informations even from compromised samples. It is possible to get clue of mutations and SNPs from the whole genome sequencing where several suspects and victims are involved. The sequencing can be performed on compromised samples such as preserved tissues; blood exposed to different temperatures, samples of hair, etc. In comparison to other techniques tested in laboratories, multiplex mini sequencing of mtDNA provides elevated rate of success (Kinra 2006).

3.5.4 X-STR Typing

Autosomal STRs have been used utilized for forensic services long before the Y-STR and X-STR markers. But these X- and Y-STR markers offer the potential to provide additional information which autosomal STRs fail to accomplish. The first major breakthrough in the field of X chromosomal markers was done by Race and Sanger, when they detected Xga blood group (Tippett and Ellis 1998). Females have a homologous pair of X chromosome. The use of X-STRs is highly proficient for determination of kinship especially between father and daughter by increasing the power of discrimination. In contrast to this, in cases related to questioned mother–daughter relationships, X chromosome markers fails to provide any specific information, just like autosomal STRs. In case of mixed stains especially in rape cases, X-STRs can positively identify female DNA (Shin et al. 2005; Szibor et al. 2000). In reference of X-STR, in the first multiplex study nine loci were analyzed (in three different multiplexes) including duplex PCR (DXS6789 and DXS6795), Triplex PCR (DXS7133, DXS9895, and DXS9898), and Quadruplex PCR (GATA164A09, DXS6803, DXS8378, and DXS7132) (Son et al. 2002). The greater the number of markers, higher will be degree of discrimination. Currently more than 40 X-STR markers have been identified for forensic purposes (Jedrzejczyk et al. 2010; Machado and Medina-Acosta 2009; Szibor 2007; Szibor et al. 2006).

3.6 Next Generation Sequencing

Since 2005, next generation sequencing (NGS) most commonly referred to, massively parallel sequencing (MPS) has substantially changed the horizon of genomic research. NGS is a ‘non-Sanger sequencing’ method that can simultaneously sequence billions of DNA molecules to minimize the fragment cloning (Yang

et al. 2014; Aly and Sabri 2015; Mostafa et al. 2015). This sequencing technique works on loop array sequencing that can simultaneously analyze huge data. Several commercial systems have been introduced based on this sequencing process. This vital technology has developed in following three sequencing generations:

1. **First generation sequencing**: The first pyro-sequencing based system was '454-Genome Sequencing System', introduced by Roche in 2005. This system primarily detects pyrophosphate which gets released after incorporation of each nucleotide in a new synthetic DNA strand. However, this system is obsolete now.
2. **Second generation sequencing**: In 2006, Solexa released Illumina, which used the technology of sequencing by synthesis. In 2007, 'SOLiD' a combination of two base encoding system and oligonucleotides ligation based 'second generation sequencing' system was introduced. In 2010, Ion Personal Genome Machine (PGM) MiSeq were released by Illumina and Torrent. PGM used semiconductor DNA technology while MiSeq is a versatile instrument performing all the steps starting from cluster generation, amplification, and sequencing up to data analysis.
3. **Third generation sequencing**: This technology is based on Single Molecule Real Time (SMRT) DNA sequencing system for determining the base composition of the single DNA molecule. SMRT sequencing basically relay on sequencing-by-synthesis approach; an SMRT chip is used that contains numerous mode-zero wave-guides. To this SMART chip, DNA polymerase molecules are attached for the synthesis of desired DNA fragments. DNA is labeled with fluorescent markers generating an illuminating signal which is captured by a sensor. SMRT technology can attain an average read lengths of 5500–8500 bp and it is also capable of directly detecting certain epigenetic modifications like 4-mC (methyl Cytosine), 5-mC, and 6-mA (methyl Adenine) (Meldrum et al. 2011; Schadt et al. 2010; Murray et al. 2012).

STR analysis is mainly carried out via separation of DNA through capillary electrophoresis (CE) based on size of DNA (Butler 2012). Although capillary electrophoresis fails due to detect internal sequence variation in STR alleles. This STR allele sequence variation is currently considered to be very crucial in different cases having traces or compromised DNA samples. Some significant techniques such as mass spectroscopy and next generation sequencing (NGS) have presently been used in forensics to learn and recognize the possible the variations in internal sequence of STR alleles (Pitterl et al. 2010; Planz et al. 2012; Rockenbauer et al. 2014a). Analysis of variation in STR sequence is these days becoming even more imperative as this will enhance the discrimination of individuals in mixed DNA samples.

NGS is the key of revealing and exactly identifying STR allele variations. The CE-based STR assay may identify the alleles as per their relative size in comparison to a sequenced allelic ladder, albeit variations in internal sequence may exist (Gettings et al. 2015). An in-depth evaluation of sequence variation of STR alleles can be made by using by NGS technology. However, such result should be

compared with the PCR-CE-length-based genotype. This further assists in studying the STR allele's rates of mutation and enhancing the understanding of rate of mutation of STR for a specific allele. An additional demerit of CE-based STR analysis is the variation in allele peak w.r.t. size of allelic ladder particularly if there are insertions or deletions (indels) in the contiguous regions of that allele which is not frequent but have a potential of its presence. For a particular marker, PCR-CE based assay can be performed within a day. Still the biggest benefit of NGS is its cumulative STR analysis and NGS assay despite of relatively longer time that it takes (Børsting and Morling 2015). Additionally NGS is also capable of analyzing DNA samples which is hard with CE-based analysis. Out of all available NGS methods, sequencing by ligations has least rate of error, and the paramount platform for various genetic applications in context to criminal justice system (Børsting and Morling 2015). Thus, progression of NGS integrated with STRs, is the need of the hour since STRs alone constitute about 15% of human genome (Gettings et al. 2015; Ayres et al. 2002). Out of 24 autosomal STR loci, nine loci showed a boost in alleles exceeding 30% when sequenced with NGS platform as compared to PCR-CE-length-based genotype (Gettings et al. 2015). However rest of 15 STR loci showed less repeat region variation, for instance, loci like D5S818, D7S820 and D13S317. But this less variation in the contiguous region is also helpful in understanding the role of different mutational events in evolution process. NGS sequencing reveals the true STR loci variation, as new alleles have been detected by sequencing the simple STRs which is vital in enhancing the statistical strength of analysis.

3.6.1 Advantages of NGS Technology

Several limitations of present CE-based analysis are there which prompted the forensic fraternity to discover the utility of NGS technology in criminal justice system. Limitations of existing system includes loss of crucial genomic information from compromised DNA samples, low-resolution genotyping of current markers, low-resolution mixture and mtDNA analysis and inability to analyze manifold genetic polymorphisms using a single workflow in a single reaction. NGS has revolutionized the existing genomic research by exceptionally improving the time, speed, cost, accuracy, and sequence length (Berglund et al. 2011; Dalsgaard et al. 2013; Fordyce et al. 2011, 2015; Gelardi et al. 2014; Phillips et al. 2014; Rockenbauer et al. 2014b; Scheible et al. 2014). The supremacy of NGS technology has its applications in different genetic arenas, like sex and mitochondrial chromosomes, autosomes, etc. NGS has impending applications in numerous aspects of research, including forensic microbiological, plant and animal analyses, construction of DNA database, phenotypic and ancestry inference, studies of monozygotic twins and species and body fluid identification. DNA sequencing is no longer a tedious task.

3.6.2 Application in Forensic Science

Due to its utility in forensic investigations, DNA analysis has been rendered as an important tool in criminal justice system. Forensic DNA analysis is often encountered with limited, contaminated and highly degraded samples, and the need of the hour is reproducibility, higher precision as well as consideration of time and cost factors.

3.6.2.1 STR Sequence Variation Detection

STR analysis is the most essential and commonly employed technique in criminal justice system. It has several advantages including low DNA template requirement, speedy and accurate allele determination, fluorescence-based detection and multiplex amplification, employment of the plentiful genomic element and digitized results. Present, STR-based databases have been established by more than 60 countries, and these databases persist to grow quickly. For instance, over 27 million entries are there in the forensic database of China (Ministry of Public Security-China 2012). Statistically, there is higher probability of an accidental correlation between unrelated individuals if analysis is based solely on 13 CODIS STR markers of routine use i.e., FGA, CSF1PO, TPOX, THO1, VWA, D5S818, D3S1358, D8S1179, D7S820, D16S539, D13S317, D21S11, and D18S51 or 15 markers (13 CODIS loci plus D19S433 and D2S1338). Assimilation of more STR markers is thus recommended to avoid this unfavorable situation. However, owing to technical faults of fluorescent-based CE sequencers which are presently in use, detection of more STR markers at the same time, would not be an easy task. Conventional CE-based STR typing using CE is unable to distinguish identical alleles of varied sequences. As a result, this technique fails to resolve cases of STR mutations in complex paternity disputes. Mixed DNA samples pose additional challenge in forensic DNA investigations. Such samples have low rates of detection and thus, are not so helpful in forensic investigations. In the beginning, NGS technology was not considered fit for STR analysis due to much shorter read length. However, with technological advancements; continual efforts have been made to increase the average read length. Many researchers have now employed NGS technology for STR analysis as it is able to easily distinguish similar length alleles, digital read count could considerably smooth the process of complex paternity analysis and the recognition of compromised DNA samples. For instance, a pioneer research has been conducted by Zajac and his colleagues. In their study, by 454 Genome Sequencing System, they used trinucleotide threading (TnT) approach to analyze three CODIS STR loci, D18S51, TPOX, and CSF1PO (Zajac et al. 2009). After this, Irwin et al., coupled with multiplex identifier technology, analyzed 13 CODIS STR loci for single source samples using 454 GS Junior system (Irwin et al. 2011). Bornman et al. went a step ahead and demonstrated that the AMEL gene and 13 CODIS STR loci can be accurately identified using high-throughput sequencing technology even for mixed samples (Bornman et al. 2012). Warshauer et al. developed software named STRait Razor, able to analyze NGS data for 44 STRs, including 21 Y and 23 autosomal chromosome STRs (Warshauer et al. 2013). Van

Neste et al. established a reference allele database using Illumina's MiSeq system to detect mixed and single source DNA samples; they found that genotyping results of most locus were reliable and stable (Van Neste et al. 2014). NGS technology thus, can considerably assist the analysis of complex and compromised DNA samples and consequently can significantly enhance the cost-efficacy and competence of cases of forensic interest.

3.6.2.2 Single Nucleotide Polymorphisms (SNP)

The 'whole-genome sequencing' based on NGS enables the exploitation of SNPs in large scale for criminal investigation with higher accuracy. Multiple important morphological characters for instance skin, hair, and eye color have been predicted with 80–90% accuracy rate by this method. SNP technique also overcomes the stuttering artifacts and it has comparatively less chance of mutation that enables it to prove the kinship (Berglund et al. 2011).

3.6.2.3 Uniparental Markers (Lineage-Based Genetic Markers)

Mitochondrial DNA (MtDNA)

mtDNA has proved its utility as a potential forensic tool due to its characteristics like small size, maternal inheritance, multiple copies, lack of recombination, and high mutation rate, in variety of cases involving low amounts of DNA is recovered and where maternal lineage is under investigation. Current forensic mtDNA analysis usually detects only polymorphisms within a hypervariable region. However, additional polymorphic loci are required for using mtDNA as a genetic haplotype marker and to enhance the discriminating capacity of identification. Consequently, NGS technology has great prospective in whole mitochondrial sequences analysis. Binladen et al. utilized a technique of coding of primer and created 256 tagged primers to be used in multiple parallel sequencing, which allowed sequencing of 256 samples in one run (Binladen et al. 2007). Gunnarsdottir et al. simultaneously sequenced complete mitochondrial genomes of 109 Filipino individuals using NGS technology (Gunnarsdottir et al. 2011). Heteroplasmy of human mtDNA is a very common event and is often encountered in different cells of an individual (Cao et al. 2006). Forensic mitochondrial analysis is often get affected by mtDNA heteroplasmy. The advantages of detecting heteroplasmy at the whole mitochondrial genome level by using NGS (Li et al. 2010) includes high sensitivity and accuracy, low cost, simple operation, and high throughput (Tang and Huang 2010). A different study used 454 GS Junior system to simultaneously examine; a Y chromosome STR locus (DYS389I/II), an Autosomal STR locus (D18S51) and multiple mitochondrial hypervariable regions and results confirmed that a mixing ratio as minimal as 1:250 of two DNA sources is detectable. Authors further assert that a mixing ratio as low as 1:1000 might be detectable by increasing the sequencing coverage (Holland et al. 2011). Sixty-four complete mitochondrial genome sequences were examined in order to evaluate the haplotypes defined via NGS technology at complete mitochondrial genome level with conventional Sanger sequencing. The results manifested differences in <0.02% of nucleotides by employing these methods. Distinction was

noticed in or around homopolymeric stretches, as these areas are more susceptible to sequencing flaws (Parson et al. 2013). Mikkelsen et al. described that if the results have careful visual inspection, in homopolymers, by using the 454 NGS method, up to six bases, 95% of the reads could be correctly sequenced (Mikkelsen et al. 2014). Earlier-unreported heteroplasmy in GM9947A component of the National Institute of Standards and Technology human mtDNA SRM-2392 standard reference was revealed. The recent developments in the complete mtDNA sequencing based on NGS technology overcome the limitation of only noncoding region sequencing. By virtue of NGS new mtDNA phylogeny and haplogroups have been discovered and important ancestry-based biogeographical markers are also found. NGS also solved the problem of mtDNA heteroplasmy due to mixed samples from the same person (Melton et al. 2012).

Y-STR

Use of Y-STRs explicitly resolves inferential issues of paternal relationships among male individuals and to determine male DNA mixed with high female background. NGS technology was used to compare more than 10,000,000 nucleotides of the Y chromosome of two males who have same ancestry before 13 generations were compared (Xue et al. 2009). Four genetic distinctions were revealed which suggested that Y chromosome sequencing could be used to resolve issues related to discrimination between mixed male samples offspring of the same parents. Van Geystelen et al. used Y chromosome SNPs to develop AMY-tree and fruitfully confirmed the variations among 118 unrelated males belonging to 109 diverse geographical locations (Van Geystelen et al. 2013a, b). From this research, AMY-tree emerged out as a potential tool for determining Y chromosome pedigrees and identifying unknown Y-SNPs belonging to diverse geographical locations. NGS technology is very useful in the Y chromosome sequencing of mixed DNA samples of male perpetrators belong to the same parent to distinguish them individually.

3.6.2.4 Epigenetic Markers

Several recent studies suggest various applicabilities of epigenetic markers in the field of forensic science. Epigenetic markers can be used to predict tissue type (Frumkin et al. 2011), to distinguish monozygotic (MZ) twins (Li et al. 2013), and to determine the age of a DNA donor with precision (Bocklandt et al. 2011a). Epigenetic markers are capable of identifying the specific body fluid source of DNA, the age of the source person, and DNA fabrication (Lee et al. 2012; Bocklandt et al. 2011b; Courts and Madea 2011). NGS technology-based epigenetic approaches comprise complete-genome methylation beadchips, bisulfite sequencing (Grunau et al. 2001), methylated DNA immunoprecipitation sequencing (Weber et al. 2005) and reduced representation bisulfate sequencing (Meissner et al. 2005). Since these sequencing methods require DNA samples in large amount; their capability to utilize minute DNA samples will be critical for the accomplishment of this technique. However, traces of DNA (say about 100 pg) have been effectively analyzed through genome-wide amplification of a bisulfite-modified DNA template, subsequently pyro-sequenced quantitative methylation detection (Paliwal et al.

2010). Another encouraging study was performed with trace blood spot samples using bisulfate genomic DNA sequencing (Xu et al. 2012).

3.6.2.5 MicroRNA (miRNA)

MicroRNAs are a set of minute endogenous RNA molecules having length range of 18–24 nucleotides. Due to their relative resilience to deterioration, tiny size and extremely tissue-divergent or tissue-specific expression, they are considered to be appropriate for forensic identification of body fluids, postmortem interval (PMI) analysis and species identification (Courts and Madea 2010). For body fluid identification due to its tissue-specific expression, it is an ideal biomarker. Introduction of miRNAs in the field of forensic sciences is a recent event. Mostly RT-PCR and biochip technology are used for the analysis of miRNA. Hanson et al., in the year 2009 initiated the use of miRNA profiling in forensic science and demonstrated that from forensic samples 452 miRNAs were genotyped by means of quantitative PCR method (Hanson et al. 2009). In a different study, on a microarray, the expression levels of 718 miRNAs in menstrual blood, saliva, venous blood, semen and vaginal secretions were profiled (Zubakov et al. 2010). Among these, 14 differentially expressed miRNAs were recognized as probable candidates for identification of body fluid. Millions of miRNA sequences can be analyzed quickly by using NGS technology, to identify miRNA expression in diverse disease states as well as organ- and developmental stage-specific expression, thus acting as a promising instrument for forensic analysis. For miRNA profiling, NGS is highly sensitive (Courts and Madea 2010; Wang et al. 2012).

3.7 Conclusion

DNA profiling has brought a boom in forensic science, but it was not just an overnight miracle. It is a bonafide result of the tireless efforts of the researchers of this field over last several decades that DNA profiling technique has emerged and established itself as an infallible technique for personal identification. This technique has changed lives of many people, not only those who have been awarded time to time in making significant contribution to this field, but also to those people who were waiting for justice. Advancements are being made in this field at a higher pace. NGS has undoubtedly given birth to the opportunity of extremely responsive and high-throughput analysis. However, more work is required to fulfill this objective, especially to overcome the problems with error rate, low-template library preparation, issues with NGS data processing and mining and type estimations. There is a need of generation of guidelines for NGS's applications in field of criminal justice system. With unremitting translational efforts of forensic scientists and technical advancements of NGS technology, it is hopefully believed that those days are not far away when NGS technology would probably become an effortlessly available customary practice in forensic science.

References

Alghanim HJ, Almirall JR (2003) Development of microsatellite markers in Cannabis sativa for DNA typing and genetic relatedness analyses. Anal Bioanal Chem 376:1225–1233

Aly SM, Sabri DM (2015) Next generation sequencing (NGS): a golden tool in forensic toolkit. Archiv Forensic Med Criminol 65(4):260–271

Ayres KL, Chaseling J, Balding DJ (2002) Implications for DNA identification arising from an analysis of Australian forensic databases. Forensic Sci Int 129:90–98

Ballantyne KN, Kayser M (2013) Additional Y-STRs in forensic: why, which and when. In: Jaiprakash S, Ray L (eds) Forensic DNA typing (current practices and emerging technologies). CRC Press, Boca Raton, FL

Ballentyne KN, Goedbloed M, Fang R et al (2010) Mutability of Y-chromosome microsatellites: rates, characteristics, molecular basis and forensic application. Am J Hum Genet 82:341–353

Ballentyne KN, Keerl V, Wollstein A et al (2012) A new future of forensic Y-chromosome analysis: rapidly mutating Y-STRs for differentiating male relatives and paternal lineages. Forensic Sci Int Genet 6:208–218

Bar W et al (2000) DNA commission of the International society of forensic genetics: guidelines for mitochondrial DNA typing. Int J Legal Med 113(4):193–196

Baxter-Lowe LA, Hunter B, Casper JT, Gorski J (1989) HLA gene amplification and hybridization analysis of polymorphism. J Clin Invest 84:613–618

Berglund EC, Kiialainen A, Syvanen AC (2011) Next generation sequencing technologies and applications for human genetic history and forensics. Investig Genet 2:23–37

Binladen J, Gilbert MT, Bollback JP, Panitz F, Bendixen C, Nielsen R et al (2007) The use of coded PCR primers enables high throughput sequencing of multiple homolog amplification products by 454 parallel sequencing. PLoS One 2:e197

Bocklandt S, Lin W, Sehl ME, Sanchez FJ, Sinsheimer JS, Horvath S et al (2011a) Epigenetic predictor of age. PLoS One 6:e14821

Bocklandt S, Lin W, Sehl ME, Sánchez FJ, Sinsheimer JS, Horvath S, Vilain E (2011b) Epigenetic predictor of age. PLoS One 6:e14821

Bornman DM, Hester ME, Schuetter JM, Kasoji MD, Minard-Smith A, Barden CA et al (2012) Short-read, high-throughput sequencing technology for STR genotyping. Biotechniques:1–6

Børsting C, Morling N (2015) Next generation sequencing and its applications in forensic genetics. Forensic Sci Int Genet 18:78–89

Brenner CH (2010) Fundamental problems of forensic mathematics - the evidential value of a rare haplotype. Forensic Sci Int Genet 4:281–291

Brenner CH, Weir BS (2003) Issues and strategies in the identification of World Trade Centre victims. Theor Popul Biol 63:173–178

Budowle B, Moretti TR, Niezgoda SJ, Brown BL (1998) CODIS and PCR-based short tandem repeat loci: law enforcement tools. In: Proceedings of the Second European Symposium on Human Identification. Promega Corporation, Madison, WI, pp 73–88

Buga W, An TL, Begovich AB, Erlich HA (1990) Rapid HLA-DPB typing using enzymatically amplified DNA and non-radioactive sequence-specific oligonucleotide probes. Immunogenetics 32:231–241

Bugawan TL, Horn GT, Long CM, Mickelson E, Hansen JA, Ferrara GB, Angelini G, Erlich HA (1988) Analysis of HLA-DP alleleic sequence polymorphism using the in vitro enzymatic DNA amplification of DPA and DP-loci. J Immunol 141:4024–4030

Butler JM (2005) Forensic DNA typing: biology, technology, and genetics of STR markers, 2nd edn. Elsevier Academic Press, New York, NY

Butler JM (2006) Genetics and genomics of core STR loci used in human identity testing. J Forensic Sci 51(2):253–265

Butler JM (2012) Advanced topics in forensic DNA typing: methodology. Elsevier Inc, Waltham, MA

Butler JM, Hill CR (2012) Biology and genetics of new autosomal STR loci useful for forensic DNA analysis. Forensic Sci Rev 24:15

Buyse I, Decorte R, Baens M, Cuppens H, Semana G, Emonds MP et al (1993) Rapid DNA typing of class II HLA antigens using the polymerase chain reaction and reverse dot blot hybridization. Tissue Antigens 41:114

Cao Y, Wan LH, Gu LG, Huang YX, Xiu CX, Hu SH et al (2006) Heteroplasmy in human mtDNA control region. Fa Yi Xue Za Zhi 22:190–192

Chambers GK, MacAvoy ES (2000) Microsatellites: consensus and controversy. Comp Biochem Physiol B Biochem Mol Biol 126:455–476

Coble MD, Loreille OM, Wadhams MJ et al (2009) Mystery solved: the identification of the two missing Romanov children using DNA analysis. PLoS One 4(3):e4838

Courts C, Madea B (2010) Micro-RNA – a potential for forensic science? Forensic Sci Int 203:106–111

Courts C, Madea B (2011) Specific micro-RNA signatures for the detection of saliva and blood in forensic body-fluid identification. J Forensic Sci 56:1464–1470

Coyle HM, Ladd C, Palmbach T, Lee HC (2001) The green revolution: botanical contributions to forensics and drug enforcement. Croat Med J 42:340–345

Coyle HM, Palmbach T, Juliano N, Ladd C, Lee HC (2003) An overview of DNA methods for the identification and individualization of marijuana. Croat Med J 44:315–321

Dalsgaard S, Rockenbauer E, Gelardi C, Borsting C, Fordyse SL, Morling N (2013) Characterization of mutations and sequence variations in complex STR loci by second generation sequencing. Forensic Sci Int Genet Suppl Ser 4:e218–e219

Datwyler SL, Weiblen GD (2006) Genetic variation in hemp and marijuana (Cannabis sativa L.) according to amplified fragment length polymorphisms. J For Sci 51:371–375

Davis C, Ge J, Sprecher C, Chidambaram A, Thompson JM et al (2013) Prototype PowerPlex1 Y23 system: a concordance study. Forensic Sci Int 7:204–208

De Guglielmo MA, Rader JM, Bever RA (1994) Evaluation of Amp-FLP markers and summary of PCR-based forensic casework. In: Bär W, Fiori A, Rossi U (eds) Advances in forensic haemogenetics, vol 5. Springer, Berlin

Dixon LA et al (2005) Validation of a 21 locus autosomal SNP multiplex for forensic identification purposes. Forensic Sci Int 154:62–77

Ellegren H (2000) Heterogeneous mutation processes in human microsatellite DNA sequences. Nat Genet 24:400–402

Erlich HA, Scharf SJ, Long CM, Horn GT (1989) Analysis of isotypic and allotypic sequence variation in the HLA-DRb region using the in vitro enzymatic amplification of specific DNA segments. In: Dupont B (ed) Immunobiology of HLA. Springer, New York, NY, pp 181–187

Fordyce SL, Avial-Arcos MC, Rockenbauer E et al (2011) High-throughput sequencing of core STR loci for forensic genetic investigations using the Roche genome sequencer FLX platform. Biotechniques 51:127–133

Fordyce SL, Mogensen HS, Borsting C, Lagac'e RE et al (2015) Second generation sequencing of forensic STRs using the Ion Torrent™ HID STR 10-plex and the ion PGMTM. Forensic Sci Int Genet 14:132–140

Frumkin D, Wasserstrom A, Budowle B, Davidson A (2011) DNA methylation-based forensic tissue identification. Forensic Sci Int Genet 5:517–524

Fry NK, Savelkoul PH, Visca P (2009) Amplified fragment-length polymorphism analysis. Methods Mol Biol 551:89–104

Gelardi C, Rockenbauer E, Dalsgaard S, Borsting C, Morling N (2014) Second generation sequencing of three complex STRs D3S1358, D21S11 and D12S391 in Danes and a proposal for nomenclature of sequenced STR alleles. Forensic Sci Int Genet 12:38–41

Gettings KB, Aponte RA, Vallone PM, Butler JM (2015) STR allele sequence variation: current knowledge and future issues. Forensic Sci Int 18:118–130

Giardina E, Spinella A, Novelli G (2011) Past, present and future of forensic DNA typing. Nanomedicine (London) 6:257–270

Gill P, Sparkes R, Kimpton C (1997) Development of guidelines to designate alleles using an STR multiplex system. Forensic Sci Int 89:185–197

Gill P, Werrett DJ, Budowle B, Guerrieri R (2004) An assessment of whether SNPs will replace STRs in national DNA databases joint considerations of the DNA working group of the European Network of Forensic Science Institutes (ENFSI) and the Scientific Working Group on DNA analysis methods (SWGDAM). Sci Justice 44:51–53

Gill P, Fereday L, Morling N, Schneider PM (2006) The evolution of DNA databases - recommendations for new European STR loci. Forensic Sci Int 156:242–244

Goodwin W, Linacre A, Hadi S (2011) An introduction to forensic genetics. Wiley Blackwell, Hoboken, NJ

Grunau C, Clark SJ, Rosenthal A (2001) Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. Nucleic Acids Res 29:E65–E65

Gunnarsdottir ED, Li M, Bauchet M, Finstermeier K, Stoneking M (2011) High-throughput sequencing of complete human mtDNA genomes from the Philippines. Genome Res 21:1–11

Haig SM, Mullins TD, Forsman ED, Trail PW, Wennerberg L (2004) Genetic identification of spotted owls, barred owls, and their hybrids: legal implications of hybrid identity. Conserv Biol 18:1347–1357

Hakki EE, Uz E, Sag A, Atasoy S, Akkaya MS (2003) DNA fingerprinting of Cannabis sativa L. accessions using RAPD and AFLP markers. Forensic Sci Int 136:31

Hameed IH, Jebor MA, Kareem MA (2015) Forensic analysis of mitochondrial DNA hypervariable region HVII (encompassing nucleotide positions 37 to 340) and HVIII (encompassing nucleotide positions 438 to 574) and evaluation of the importance of these variable positions for forensic genetic purposes. Afr J Biotechnol 14(5):365–374

Hanson EK, Lubenow H, Ballantyne J (2009) Identification of forensically relevant body fluids using a panel of differentially expressed microRNAs. Anal Biochem 387:303–314

Holland MM, McQuillan MR, O'Hanlon KA (2011) Second generation sequencing allows for mtDNA mixture deconvolution and high resolution detection of heteroplasmy. Croat Med J 52:299–313

Horn GT, Buga Wan TL, Long CM, Erlich HA (1988) HLA class II antigens and susceptibility to insulin-dependent diabetes mellitus. Proc Natl Acad Sci U S A 85:6012–6016

Hugh H, Fudenberg JRL, Wang A-C, Ferrara GB (1984) Basic immunogenetics, 3rd edn. Oxford University Press, London

Irwin J, Just R, Scheible M, Loreille O (2011) Assessing the potential of next generation sequencing technologies for missing persons identification efforts. Forensic Sci Int Genet Suppl Ser 3:447–448

Jain T, Shrivastava P, Bansal DD, Dash HR, Trivedi VB (2016) PowerPlex Y23 system: a fast, sensitive and reliable Y-STR multiplex system for forensic and population genetic purpose. J Mol Biomark Diagon 7:3

Jedrzejczyk M, Jacewicz R, Kozdraj A, Szram S, Berent J (2010) Application of X-STR loci in forensic genetics. Problem Forensic Sci 82:141–150

Jeffreys AJ, Wilson V, Thein SL (1985a) Hypervariable 'minisatellite' regions in human DNA. Nature 314:67–73

Jeffreys AJ, Wilson V, Thein SL (1985b) Individual-specific 'fingerprints' of human DNA. Nature 316:76–79

Jeffrys AJ, Brookfield JFY, Semeonff R (1987) Positive identification of a immigration test case using human DNA fingerprints. Nature 317:818–819

Jobling M, Gill P (2004) Encoded evidence: DNA in forensic analysis. eNat Rev 5:739–752

Kavlick MF, Lawrence HS, Merritt RT, Fisher C, Isenberg A, Robertson JM, Budowle B (2011) Quantification of human mitochondrial DNA using synthesized DNA strands. J Forensic Sci 56 (6):1457–1463

Kayser M (2003) The human Y chromosome: introduction into genetics and applications. Forensic Sci Rev 15:78–90

Kayser M (2017) Forensic use of Y-chromosome DNA: a general overview. Hum Genet 136:621–635

Kayser M, Krawczak M, Excoffier L et al (2001) An extensive analysis of Y-chromosomal microsatellite haplotypes in globally dispersed human populations. Am J Hum Genet 68:990–1018

Kayser M, Lao O, Anslinger K et al (2005) Significant genetic differentiation between Poland and Germany follows present day political borders, as revealed by Y-chromosome analysis. Hum Genet 117:428–443

Kayser M, Vermeulen M, Knoblauch H, Schuster H, Krawczak M, Roewer L (2007) Relating two deep rooted pedigrees from Central Germany by high resolution Y-STR haplotyping. Forensic Sci Int Genet 1:125–128

Kimpton C, Fisher D, Watson S, Adams M, Urquhart A, Lygo J et al (1994) Evaluation of an automated DNA profiling system employing multiplex amplification of four tetrameric STR loci. Int J Legal Med 106:302–311

Kimpton CP, Oldroyd NJ, Watson SK, Frazier RR, Johnson PE, Millican ES et al (1996) Validation of highly discriminating multiplex short tandem repeat amplification systems for individual identification. Electrophoresis 17:11283–11293

Kinra SLP (2006) The use of mitochondrial DNA and short tandem repeat typing in the identification of air crash victims. Ind J Aerospace Med 50:54–65

Kircher M, Kelso J (2010) High-throughput DNA sequencing-concepts and limitations. Bioessays 32(6):524–536

Kleppe K, Ohtsuka E, Kleppe R et al (1971) Repair replications of short synthetic DNA's as catalyzed by DNA polymerases. J Mol Biol 56(2):341–361

Klevan L, Horton L, Carlson DP, Eisenberg AJ (1995) Chemiluminescent detection of DNA probes in forensic analysis. Electrophoresis 16(9):1553–1558

Koreth J (1996) Microsatellites and PCR genomic analysis. J Pathol 178:239–248

Krenke BE, Tereba A, Anderson SJ, Buel E, Culhane S, Finis CJ et al (2002) Validation of a 16 locus fluorescent multiplex system. J Forensic Sci 47:773–785

Lee JCI, Cole M, Linacre A (2000) Identification of hallucinogenic fungi from the genera Psilocybe and Panaeolus by amplified fragment length polymorphism. Electrophoresis 21:1484–1487

Lee HY, Park MJ, Choi A, An JH, Yang WI, Shin KJ (2012) Potential forensic application of DNA methylation profiling to body fluid identification. Int J Legal Med 126:55–62

Li M, Schonberg A, Schaefer M, Chroeder R, Nasidze I, Stoneking M (2010) Detecting heteroplasmy from high-throughput sequencing of complete human mitochondrial DNA genomes. Am J Hum Genet 87:237–249

Li C, Zhao S, Zhang N, Zhang S, Hou Y (2013) Differences of DNA methylation profiles between monozygotic twins' blood samples. Mol Biol Rep 40:5275–5280

Linacre A, Cole M, Lee JCI (2002) Identifying the presence of 'magic mushrooms' by DNA profiling. Sci Justice 42:50–54

Lygo JE, Johnson PE, Holdaway DJ, Woodroffe S, Whitaker JP, Clayton TM et al (1994) The validation of short tandem repeat (STR) loci for use in forensic casework. Int J Legal Med 107:77–89

Machado FB, Medina-Acosta E (2009) Genetic map of human X-linked microsatellites used in forensic practice. Forensic Sci Int Genet 3:202–204

Marchi E (2004) Methods developed to identify victims of the World Trade Centre disaster. Am Lab 36:30–36

Meissner A, Gnirke A, Bell GW, Ramsahoye B, Lander ES, Jaenisch R (2005) Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. Nucleic Acids Res 33:5868–5877

Meldrum C, Doyle MA, Tothill RW (2011) Next generation sequencing for cancer diagnostics: a practical perspective. Clin Biochem Rev 32:177–195

Melton T, Holland C, Holland M (2012) Forensic mitochondrial DNA analysis: current practice and future potential. Forensic Sci Rev 24:101–122

Mikkelsen M, Frank-Hansen R, Hansen AJ, Morling N (2014) Massively parallel pyrosequencing of the mitochondrial genome with the 454 methodology in forensic genetics. Forensic Sci Int Genet 12:30–37

Ministry of Public Security-China (2012) Proceedings of the 4th National Symposium on Forensic DNA Inspection Technology & 2012 International Symposium on New Advances in Forensic Genetics, Fuzhou, China, December 9–12, 2012

Morel C, Arlen F, Jeannet M, Mach B, Tiercy JM (1990) Complete analysis of HLA-DQB1 polymorphism and DR-DQ linkage disequilibrium by oligonucleotide typing. Hum Immunol 29:64–77

Morretti TR et al (2001) Validation of short tandem repeats (STRs) for forensic usage: performance testing of fluorescent multiplex STR systems and analysis of authentic and simulated forensic samples. J Forensic Sci 46:647–660

Mostafa EM, Sabri DM, Aly SM (2015) Overviews of “next-generation sequencing”. Res Rep Forensic Med Sci 2015:1–5

Mullis KB, Faloona F (1987) Specific synthesis of DNA in vitro via a polymerase chain reaction. Methods Enzymol 155:335–350

Mullis KB, Faloona F, Scharf SJ, Saiki RK, Horn GT, Erlich HA (1986) Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Quant Biol 51:263–273

Murray IA, Clark TA, Morgan RD, Boitano M, Anton BP, Luong K et al (2012) The methylomes of six bacteria. Nucleic Acids Res 40:11450–11462

Nadir E (1996) Microsatellite spreading in the human genome: evolutionary mechanisms and structural implications. Proc Natl Acad Sci U S A 93:6470–6475

Pai CY, Chou SL, Tang TK, Wei YH, Yang CH (1997) Haplotyping of mitochondrial DNA in the D-loop region by PCR: forensic application. J Formosan Assoc 96:73–82

Paliwal A, Vaissiere T, Herceg Z (2010) Quantitative detection of DNA methylation states in minute amounts of DNA from body fluids. Methods 52:242–247

Panet A, Khorana HG (1974) The linkage of deoxyribopolynucleotide templates to cellulose and its use in their replication. J Biol Chem 249(16):5213–5221

Parson W, Strobl C, Huber G, Zimmermann B, Gomes SM, Souto L et al (2013) Evaluation of next generation mtGenome sequencing using the Ion Torrent Personal Genome Machine (PGM). Forensic Sci Int Genet 7:543–549

Payne BAI et al (2013) Universal heteroplasmy of human mitochondrial DNA. Hum Mol Genet 22 (2):384–390

Phillips C, Gelabert-Besada M, Fernandez-Formoso L, Garcia-Magarinos M, Santos C, Fondevila M, Ballard D, Court DS, Carracedo A, Lareu MV (2014) New turns from old STRs: enhancing the capabilities of forensic short tanden repeat analysis. Electrophoresis 35:3173–3187

Pitterl F, Schmidt K, Huber G, Zimmermann B, Delport R, Amory S, Ludes B, Oberacher H, Parson W (2010) Increasing the discrimination power of forensic STR testing by employing high-performance mass spectrometry, as illustrated in indigenous south African and central Asian populations. Int J Legal Med 124:551–558

Planz JV, Sannes-Lowery KA, Duncan DD, Manalili S, Budowle B, Chakraborty R, Hofstadler SA, Hall TA (2012) Automated analysis of sequence polymorphism in STR alleles by PCR and direct electrospray ionization mass spectrometry. Forensic Sci Int Genet 6:594–606

Prinz M, Boll K, Baum HJ, Shaler B (1997) Multiplexing of Y chromosome specific STRs and performance for mixed samples. Forensic Sci Int 85:209–218

Quintans B, Alvarez-Iglesias V, Salas A, Phillips C, Lareu MV, Carracedo A (2004) Typing of mitochondrial DNA coding region SNPs of forensic and anthropological interest using SNaP-shot minisequencing. Forensic Sci Int 140(2):251–257

Rockenbauer E, Hansen S, Mikkelsen M, Børsting C, Morling N (2014a) Characterization of mutations and sequence variants in the D21S11 locus by next generation sequencing. Forensic Sci Int Genet 8:68–72

Rockenbauer E, Hansen S, Mikkelsen M, Borsting C, Morling N (2014b) Characterization of mutations and sequence variants in the D21S11 locus by next generation sequencing. Forensic Sci Int Genet 8:68–72
Roewer L (2003) The use of the Y chromosome in forensic genetics - current practices and future perspectives. Int Congr Ser 1239:279–280
Roewer L (2009) Y chromosome STR typing in crime casework. Forensic Sci Med Pathol 5:77–84
Roewer L, Arnemann J, Spurr NK, Grzeschik KH, Epplen JT (1992) Simple repeat sequences on the human Y chromosome are equally polymorphic as their autosomal counterparts. Hum Genet 89(4):389–394
Romeika JM, Yan F (2013) Recent advances in forensic DNA analysis. J Forensic Res S12:001. https://doi.org/10.4172/2157-7145.S12-001
Saiki RK, Buga Wan TL, Horn GT, Mullis KB, Erlich HA (1986a) Analysis of enzymatically amplified-globin and HLA-DQA DNA with allele-specific oligonucleotide probes. Nature 324:163–166
Saiki RK, Bugawan TL, Horn GT, Mullis KB, Erlich HA (1986b) Analysis of enzymatically amplified beta-globin and HLA-DQ alpha DNA with allele-specific oligonucleotide probes. Nature 324:163–166
Saiki R, Chang CA, Levenson CH, Boehm CD, Kazazian RH, Erlich HA (1987) Rapid genetic analysis of enzymatically-amplified DNA with non-radioactive allele-specific oligo-nucleotide probes. Am J Hum Genet 41:A237
Schadt EE, Turner S, Kasarskis A (2010) A window into third generation sequencing. Hum Mol Genet 19:2
Scheible M, Loreille O, Just R, Irwin J (2014) Short tandem repeat typing on the 454 platform: strategies and considerations for targeted sequencing of common forensic markers. Forensic Sci Int Genet 12:107–109
Sheshanna SH, Hegde U, Srinivasaiyer M, Balaraj BM (2014) Review research paper, mitochondrial DNA: a reliable tool in forensic odontology. J Ind Acad Forensic Med 36(4):407–410
Shin SH, Yu JS, Park SW, Min GS, Chung KW (2005) Genetic analysis of 18 X-linked short tandem repeat markers in Korean population. Forensic Sci Int 147:35–41
Son JY, Lee YS, Choung CM, Lee SD (2002) Polymorphism of nine X chromosomal STR loci in Koreans. Int J Legal Med 116:317–321
Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503–527
Sparkes R, Kimpton C, Gilbard S, Carne P, Andersen J, Oldroyd N et al (1996a) The validation of a 7- locus multiplex STR test for use in forensic casework. (II), artefacts, casework studies and success rates. Int J Legal Med 109:195–204
Sparkes R, Kimpton C, Watson S, Oldroyd N, Clayton T, Barnett L et al (1996b) The validation of a 7- locus multiplex STR test for use in forensic casework. (I), Mixtures, ageing, degradation and species studies. Int J Legal Med 109:186–194
Subramanian S (2003) Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. Genome Biol 4:R13
Sullivan KM, Mannucci A, Kimpton CP, Gill P (1993) A rapid and quantitative DNA sex test: fluorescence based PCR analysis of X–Y homologous gene amelogenin. Biotechniques 15:636–641
Szibor R (2007) X-chromosomal markers: past, present and future. Forensic Sci Int Genet 1:93–99
Szibor R, Lautsch S, Plate I, Beck N (2000) Population data on the X chromosome short tandem repeat locus HumHPRTB in two regions of Germany. J Forensic Sci 45:231–233
Szibor R, Hering S, Edelmann J (2006) A new Web site compiling forensic chromosome X research is now online. Int J Legal Med 120:252–254
Tang S, Huang T (2010) Characterization of mitochondrial DNA heteroplasmy using a parallel sequencing system. Biotechniques 48:287–296
Tautz D (1993) Notes on the defunction and nomenclature of tandemly repetitive DNA sequences. In: Pena SDJ (ed) DNA fingerprinting: state of the science. Birkhauser, Basel, pp 21–28

Thompson JM, Ewing M, Frank W, Pogmiller J, Nolde C et al (2013) Developmental validation of the PowerPlex Y23 system: a single multiplex Y-STR system for casework and database samples. Forensic Sci Int Genet 7(2):240–250

Thymann M, Nellemann LJ, Masumba G, Irgens-Moller L, Morling N (1993) Analysis of the locus D1S80 by amplified fragment length polymorphism technique (AMP-FLP): frequency distribution in Danes. Intra and inter laboratory reproducibility of the technique. Forensic Sci Int 60 (1-2):47–56

Tippett P, Ellis NA (1998) The Xg blood group system: a review. Transfus Med Rev 12(4):233–257

Van Geystelen A, Decorte R, Larmuseau MH (2013a) Updating the Y-chromosomal phylogenetic tree for forensic applications based on whole genome SNPs. Forensic Sci Int Genet 7:573–580

Van Geystelen A, Decorte R, Larmuseau MH (2013b) AMY-tree: an algorithm to use whole genome SNP calling for Y chromosomal phylogenetic applications. BMC Genomics 14:101

Van Neste C, Vandewoestyne M, Van Criekinge W, Deforce D, Van Nieuwerburgh F (2014) My-Forensic-Loci-queries (MyFLq) framework for analysis of forensic STR data generated by massive parallel sequencing. Forensic Sci Int Genet 9:1–8

Vos P, Hogers R, Bleeker M et al (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Wang Z, Luo H, Pan X, Liao M, Hou Y (2012) A model for data analysis of microRNA expression in forensic body fluid identification. Forensic Sci Int Genet 6:419–423

Warshauer DH, Lin D, Hari K, Jain R, Davis C, LaRue B et al (2013) STRait Razor: a length-based forensic STR allele-calling tool for use with second generation sequencing data. Forensic Sci Int Genet 7:409–417

Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL et al (2005) Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat Genet 37:853–862

Willuweit S, Caliebe A, Andersen MM, Roewer L (2011) Y-STR frequency surveying method: a critical reappraisal. Forensic Sci Int Genet 5:84–90

Xu H, Zhao Y, Liu Z, Zhu W, Zhou Y, Zhao Z (2012) Bisulfite genomic sequencing of DNA from dried blood spot microvolume samples. Forensic Sci Int Genet 6:306–309

Xue Y, Wang Q, Long Q, Ng BL, Swerdlow H, Burton J et al (2009) Human Y chromosome base-substitution mutation rate measured by direct sequencing in a deep-rooting pedigree. Curr Biol 19:1453–1457

Yang Y, Xie B, Yan J (2014) Application of next-generation sequencing technology in forensic science. Genom Proteom Bioinformatics 12:190–197

Yoshida K, Yayama K, Hatanaka A, Tamaki K (2011) Efficacy of extended kinship analysis utilizing commercial STR kit in establishing personal identification. Legal Med 13:12–15

Zajac P, Oberg C, Ahmadian A (2009) Analysis of short tandem repeats by parallel DNA threading. PLoS One 4:e7823

Zubakov D, Boersma AW, Choi Y, van Kuijk PF, Wiemer EA, Kayser M (2010) MicroRNA markers for forensic body fluid identification obtained from microarray screening and quantitative RTPCR confirmation. Int J Legal Med 124:217–226

Forensic DNA Evidence: From Crime Scene to Conviction

4

Ashish Badiye, Neeti Kapoor, and Pankaj Shrivastava

Abstract

DNA analysis is the preferred and often considered to be one of the most reliable method in forensics and criminal investigation. Forensic DNA typing is utilised in a wide variety of cases for proving the guilt as well as to prove innocence. From lifting the exhibits at the crime scene or recovering the pieces of evidence from the victim or suspect, their transportation, laboratory analysis, and reporting while maintaining the chain of custody are the tasks which are taken care of by the investigating agency and forensic DNA analyst. All these steps should be free from contamination and cross-contamination. This chapter throws light on some of the critical aspects of forensic DNA analysis, along with a few real case studies to showcase the importance of forensic DNA examination. With practice, precaution, and precision, from the inception (origin of the DNA evidence) to the end (reporting and deposition in the court), the optimum potential of DNA evidence from the crime scene to conviction is discussed.

Keywords

DNA fingerprinting · Chain of custody · DNA sources · DNA technology · DNA case study · Biological evidence

A. Badiye (✉) · N. Kapoor
Department of Forensic Science, Government Institute of Forensic Science, Nagpur, Maharashtra, India

P. Shrivastava
DNA Fingerprinting Unit, State Forensic Science Laboratory, Department of Home (Police), Government of Madhya Pradesh, Sagar, Madhya Pradesh, India

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_4

4.1 Introduction

DNA is the most sensitive and reliable evidence for the assessment of personal identity and has now become a backbone of the criminal justice system. DNA analysis is the 'gold standard' of forensic science and the preferred method of personal identification (Lynch 2003), given its exclusivity (with the exception of identical twins) and usefulness in proving the guilt or innocence especially in sexual assault cases. The DNA evidence which is collected from the crime scene can identify and/or eliminate the person from the existing pool of suspects in an investigation making the investigation scientific and thus more reliable. With the continual advancement in the DNA know-how, the technology has become faster, cheaper and more accessible. With the help of DNA databases (like Combined DNA Index System (CODIS) and now expanded CODIS), it is possible to trace or connect the same suspect locally, nationally, or internationally, watch list of investigating/watch list of investigating agencies (Ashcroft et al. 2002).

DNA is also known as the 'blueprint of life'. The information within the DNA 'blueprint' is coded by the series of the four different nitrogenous bases, adenine, guanine, thymine and cytosine, on the sugar-phosphate backbone (Fig. 4.1). DNA is the sequence of the 'nucleotide base pairs' that link the two strands of the double helix which is composed of sugar and phosphate (Goodwin et al. 2007; Kaye 2007). With the known exception of RBCs, DNA is found in all the cells of human and

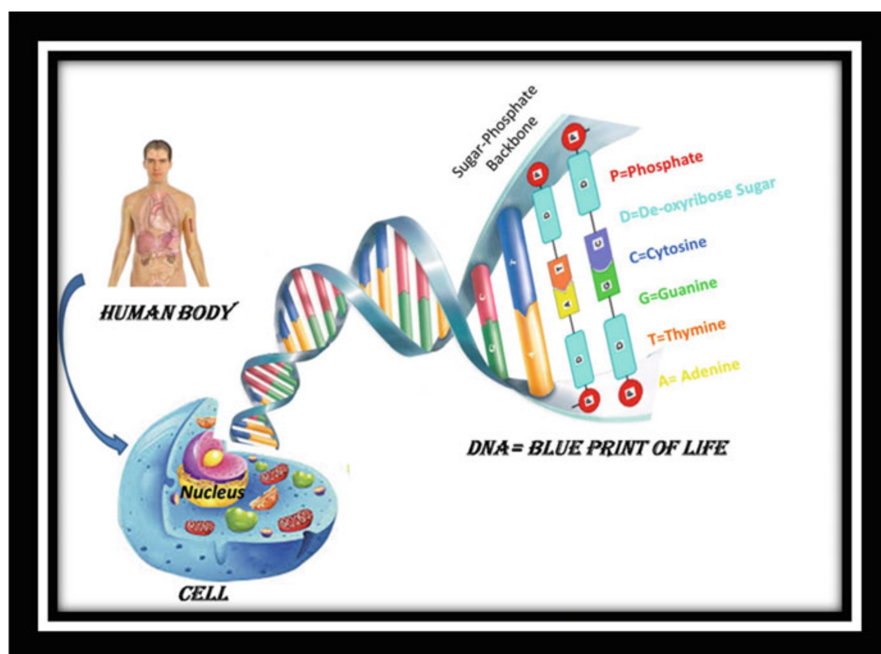


Fig. 4.1 DNA the 'Blueprint of life'

Fig. 4.2 Some of the biological evidences as the potential source of DNA

animal biological origin. It also remains unchanged throughout life. Every individual's genome possesses a vast quantity of DNA that is a likely target for DNA profiling even from a sample which is very much degraded.

Fingerprints were and still are said to be an 'infallible' means of individualisation and identification (Kapoor and Badiye 2015a, b, c). However, it has undergone quite some criticism and scrutiny in the scientific community as well as the court of law in the recent past (de Villiers 2014). Also, owing to the increasing awareness among the criminals, they often wear gloves while committing a crime, ensuring not to leave behind fingerprints that could lead the investigators to them (Harris 2001). Additionally, the skillful lifting of the chance prints from the scene of occurrence is another challenge. But, in several cases, the suspects unknowingly and invariably leave other traces (biological/trace evidences) like blood, semen, saliva, skin, sweat, hair, among others (Fig. 4.2). These pieces of evidence can then be utilised to determine the DNA profile of the person.

In the last two decades, there have been several attempts to further work on the lines of Locards' Principle of Exchange (Locard 1930), which states that 'every touch/contact leaves a trace' and hence, DNA of the individual would be found on any surface that was touched/contacted (Phipps and Petricevic 2007; Daly et al. 2012; Polley et al. 2006; Allen et al. 2008). Therefore, another advancement in the present-day DNA technique is 'Touch DNA'. Even if a person wore gloves during the crime, and the gloves are dumped and recovered at someplace near the crime scene, the inner side of the gloves would be sufficient to reveal the DNA profile of the individual (Goray et al. 2019).

Although the DNA evidence found may be aged or often degraded, but the damage still usually would not prevent the DNA analysis. With modern progressions, it is now possible to obtain a complete DNA profile even from highly degraded samples. Continual advancements in multiplex kits including several works on keeping more and more markers in size less than 250 bp in a single multiplex, use of improved master mix and Taq polymerase have helped in forensic DNA analysis (Shrivastava et al. 2016).

4.2 DNA Evidence and Its Applications

DNA as evidence has potential use in several fields (Fig. 4.3) including criminal, civil, immigration and other cases. DNA in the form of biological evidence may be encountered by the investigators in various type of criminal cases such as (but not limited to) murder, rape, gang-rape, sexual assault, sexual abuse, child abuse, unnatural sexual offences (like buccal coitus, sodomy, bestiality, etc.), sexual perversions (sadism, fetishism, necrophilia, etc.), theft, robbery, mass-graves, drowning, mutilation of the dead body, necrophagy, black magic related crimes, disfigured/dismembered/burnt bodies, and wildlife crimes.

DNA as evidence may also be useful in cases of accidental deaths, mass-disasters, exhumation cases, archaeological cases, etc. Additionally, it can also be utilised in civil cases such as maternity-paternity dispute cases, immigration cases, adultery, recognition of legal heirs, in the cases involving the exchange of babies in the hospital or baby care centre, and identification of the missing person (after elapsed time). Old, cold and unsolved cases have also been solved with the help of DNA evidence. The DNA databases of various agencies across the world facilitates law enforcement by aiding in investigations by successfully identifying suspects and related sequential/serial crimes committed in one or more jurisdictions.

In the earliest known case, DNA as evidence was used in the rape cum murder case of a 15-year-old Dawn Ashworth in Leicestershire, England in July 1986 (Hodge 2017; Arnaud 2017). According to the discovery of Dr Alec Jeffreys, the DNA patterns of individuals were dissimilar and accounted for the individualisation of the person (Jeffreys et al. 1985a, b, c). In the Ashworth murder case, the police officer took the help of Dr Alec Jeffreys' technique. He used the DNA recognition technique by analysing the blood, and saliva samples of available tracked suspect Richard Buckland, who had already confessed to Ashworth's murder. But the result

Fig. 4.3 The DNA evidence and its applications

of the matching of DNA of suspect with the sample collected from the crime scene was found to be a non-match. Further, 4000 samples of men aged between 17 and 34 years from the Leicestershire area were examined. The result stated that the DNA of these men did not match with the suspect sample. Meanwhile, someone overheard a man boasting about how he had been paid to give a sample in someone else's place. In 1987, the police officer found that local baker Colin Pitchfork had denied giving a sample for assessing DNA. When Pitchfork's DNA was examined, it matched with the crime scene samples. Pitchfork was detained on Sept 19, 1987, convicted and sentenced to life in prison a year later (Hodge 2017; Arnaud 2017).

Since many years, DNA has turned out to be one of the most powerful tools, to evaluate the identity of suspects and victims, convict the blameworthy, and vindicate the innocent. The DNA profiling technology has so advanced and developed now that degraded samples, archaeological samples, and even world war victim's samples can be analysed (Alghafri et al. 2018). Even a slight touch can establish linkage of someone at a crime scene. With the availability of many techniques for individualisation by DNA tests, the courts demand further promising result with greater accuracy (Dash et al. 2020). Few courts require for experts to not only deliver

the DNA profiles but also to give the quantitative assessment of how uncommon the identifying features among the particular groups and subgroups are (National Research Council (US) Committee on DNA Forensic Science 1996).

There are some cases that pose challenging conditions for establishing personal identity. Like in cases of drowning, the body after being submerged in water is often all swollen up and is in a highly decomposed state. In cases of burning/charring of the body, all/most of the features permitting personal identification are lost. In cases of the crushed face, or a mutilated body or severing off of the head from the torso also prevents identification of the person. Fingerprints of innocent are often not present on any records to permit comparison and identification. In cases of skeletonised remains or recovery of skeletons (complete or partial), it becomes problematic to estimate the sex of the person. In all these seemingly 'difficult' cases, DNA-based methods have proved to be better in providing accurate results. Especially in comparison with the anthropological methods of sex determination which relies on the anthropometric landmarks of skull, mandible and other bones (but is often quite confusing in Indian context wherein the population is non-homogenous), DNA-based methods more often than not offer quick and accurate results.

4.3 DNA Evidence at the Scene of Crime

The crime scene is a place where the crime has occurred, or the evidence related to a crime are found. In a broader sense, it will also include the victim/survivor, the assailant, dumping areas (where the dead body or other related evidence may have been thrown or dumped or disposed of), the weapon of offence, the clothes of all those involved and surfaces or object touched/used. The presence of DNA usually proves the presence of a person on the spot or his/her involvement in the crime. While mere presence or absence of DNA evidence may not be sufficient to rule-in or rule-out an individual as the perpetrator, but it does shift the burden of proof on such a person.

It is essential to understand the intricacies of DNA evidence encountered at the place of occurrence of a crime or on a person involved in a crime. The crime scene is to be secured to prevent unauthorised access. Crowding of the scene is to be avoided/minimised at all costs. The pieces of evidence should be searched using an appropriate method(s) depending upon the crime scene. Labelling of the location of the evidence must be followed by the documentation of the crime scene (using any one or combination of these methods, i.e. notes taking/sketching/photography/videography). Once this has been done, the next step of foremost importance would be the recognition of evidence or the source from which such biological material or evidence may be obtained that can be a potential source of DNA. The evidence can occur in varied forms (solid, liquid, semisolid, dried), at different locations (primary crime scene, secondary crime scene, tertiary crime scene, etc.) and in different conditions (pure, mixed, contaminated, diluted, etc.). The various sources of DNA from some of the biological evidences at the scene are shown in Fig. 4.2.

These may be found on any of the victim and/or assailant and the objects associated with them such as the weapon of offence, fingernails of the victim/survivor, cigarette stub, clothes and garments of the victim and/or assailant and/or witnesses, used condoms, bedsheets, or surfaces or objects touched or contacted by them.

The crime scene investigation is a complex and meticulous process. For successful conviction, it essential that the evidence is appropriately recognised, collected, preserved, packed, stored, forwarded to the Forensic Science Laboratory, appropriately analysed in the lab and reported. All these steps should be free from contamination and cross-contamination. A proper chain of custody should be maintained throughout the process (Badiye et al. 2019). Additionally, taking care of the quality control issues at every step is also imperative. The expert performing the DNA analysis should be proficient and well versed with the techniques and other relevant details so as to face and withstand the scrutiny of the legal and judicial system.

Forensic Science is the field that acts as a bridge between the investigation and trial. Its knowledge facilitates the crime scene analyst to collect the physical evidence scientifically, the laboratory analyst/scientific officer to appropriately analyse the evidence and present the findings in the form of a scientific report in the court of law.

The collection method used should be appropriate and as per the need of the sample (Dash et al. 2020). Gloves must be used to handle all the exhibits or potential sources of evidence. Fresh gloves must be used for touching different objects/evidence stained with biological fluids. To reduce the risk of contamination, the crime scene analyst must invariably use PPE (personal protective equipment, i.e. coveralls, gloves, shoe-covers, headgear/headcover, mask, eyewear/safety goggles). The sterile and disposable materials and supplies must be used for collection. The blood sample, semen, saliva or other stains collected from the crime scene must be invariably air-dried at room temperature and stored in paper bags rather than directly in plastic to avoid degradation and prevent microbial contamination. Adequate quantities must be collected so as to ensure multiple testing (if required). Shreds of evidence, even if of the same type, if obtained from different places/sources at the crime scene, must be packed separately. For collecting liquid samples, appropriate preservatives (e.g. EDTA vials for liquid blood samples) must be used. If appropriate preservatives are not available for liquid samples, they must be stored at a cold temperature (at 4 °C or a freezer at −20 °C to reduce microorganisms' growth rate and to avoid DNA degradation) and immediately transferred to the lab (Magalhães et al. 2015). For dried stains, swabbing with the help of sterile swab ((saline) wet and/or dry swabbing) may be used. In sexual assault cases, the fingernail scrapping of the victim as well as the suspect(s) is of immense importance. The doctors examining the victim/survivor of rape/sexual assault must keep an empathetic attitude and must follow the appropriate guidelines of examination and must not end-up further distressing the survivor. The appropriate method of management of biological evidence for DNA analysis should be used (Kleypas and Badiye 2020; Magalhães et al. 2015; Lee and Ladd 2001; Baxter 2015).

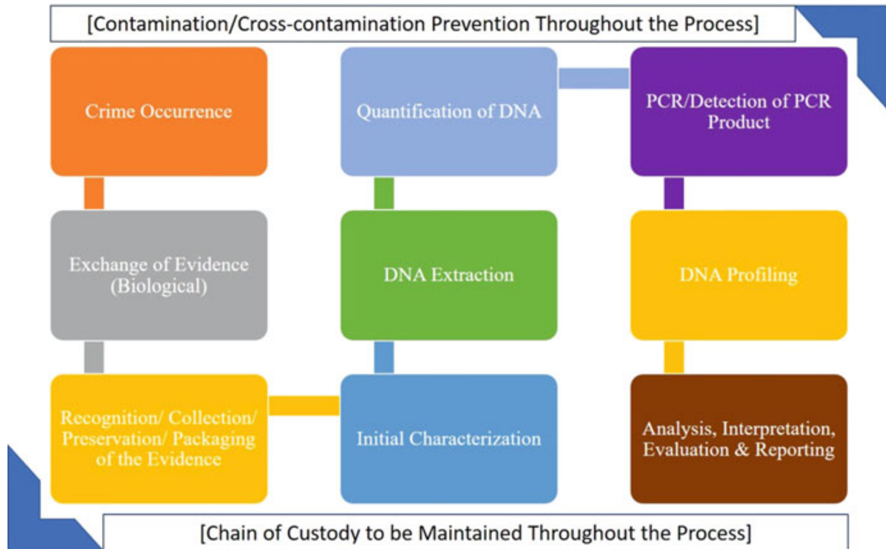


Fig. 4.4 DNA evidence from the crime scene to courtroom

The packaging of the evidence should be done in tamper-proof and/or tamper-evident bags. The labelling should be appropriate with the signature of witnesses at the appropriate place. The transportation to the laboratory should be secured. At the laboratory, the sample would be received and sent to the appropriate section/division for further processing and examination. After initial characterisation, the extraction of the DNA would be done using an appropriate method (depending upon the type of sample and availability of resources). It would be followed by the quantification of DNA, amplification by PCR (polymerase Chain Reaction), detection of PCR product, obtaining the DNA profile followed by the analysis, interpretation, evaluation and finally reporting (Fig. 4.4).

On analysis of the DNA evidence, there is a possibility of obtaining either of the following results (Lynch 2003):

- **Inclusion**: If the DNA profile of the suspect matches with the DNA profile obtained from the exhibit, the results are considered to be an 'inclusion' or 'non-exclusion', i.e. the suspect is included (and hence not excluded) as a possible source of the contributor to the DNA exhibit (Lynch 2003).
- **Exclusion**: If the DNA profile of the suspect does not match with the DNA profile obtained from the exhibit, the results are considered to be an 'exclusion' or 'non-inclusion' i.e. the suspect may be excluded (and hence not included) as a possible contributor to DNA exhibit (Lynch 2003).
- **Inconclusive Result**: As a third possibility, the results may be 'inconclusive'. There may be several reasons for the same, including, degraded or contaminated samples, mixed samples or samples not adequately preserved (Lynch 2003).

4.4 Solved Criminal Cases by Using DNA Technology

We discuss herein some fascinating cases that involved the use of DNA as the primary evidence and have led to solving the mystery of these crimes. The first case study was of a case of sexual assault (Sect. 4.4.1). The second case study was of a case of sexual assault followed by the murder of an 8-year-old girl child. The face of the girl child was disfigured by crushing with a heavy stone. For confirming the identity of the girl, paternity testing was also conducted (Sect. 4.4.2). The third case study involved a case wherein a girl was gang-raped by three men but trying to be 'ahead' of the investigating agencies, all three of them wore condoms. The outside, as well as the inside of the condoms, were tested and profiled. The investigating team, as well as the forensic science laboratory, worked in tandem to effectively solve this case without the loss of evidence (Sect. 4.4.3). The fourth case study depicts the advantage of the use of DNA technology over the traditional anthropological assessments used to determine the sex of the skeletal remains found at a crime scene (Sect. 4.4.4). The fifth case study involves a case of rape on a minor. The vaginal swab of the girl child did not reveal the presence of male DNA, but penile swab of the suspect showed the presence of the girl child's DNA (Sect. 4.4.5).

4.4.1 Case 1: A Case of Sexual Assault Confirmed by Using DNA Technology

In a case of sexual assault blood sample of the victim, the blood sample of the suspect, undergarment of the victim and vaginal smear slide of the victim were received for forensic DNA examination. DNA was extracted from the reference blood sample of the victim and suspect, using an automated DNA extraction system AutoMateExpress™ (ThermoFisher Scientific) as per the protocol of the manufacturer. The differential extraction procedure was followed for obtaining male and female DNA from undergarment and vaginal smear slide of the victim. Obtained DNA was quantified using ABI 7500 Real-Time PCR (Thermo Fisher Scientific) using the manufacturer recommended protocol. Five hundred picogram of DNA was amplified using ABI Global Filer™ multiplex kit as per recommendations of the manufacturer. The amplified DNA fragments were run separated using Genetic Analyser ABI 3500XL as per recommendations of the manufacturer, and the obtained data were analysed using GeneMapper® ID-X software v1.5. Mixed autosomal STR profile was detected on the source of both the articles of the victim. All the alleles in the DNA profile of the victim and accused were observed in the mixed DNA profile observed on the source of the victim. Thus, the obtained results confirmed the transfer of male's DNA on the source objects of the victim (Table 4.1).

Table 4.1 Mixed autosomal STR DNA profile obtained in a case of sexual assault

Genetic markers	Article A Undergarment of victim	Article B Vaginal slide of victim	Article C Blood sample of the victim	Article E Blood sample of accused
D3S1358	15,16,18	15,16,18	15,18	16
vWA	17,18	17,18	17	18
D16S539	9,11,12,13	9,11,12,13	12,13	9,11
CSF1PO	11,12,13	11,12,13	11,13	11,12
TPOX	8,9,10,11	8,9,10,11	8,11	9,10
D8S1179	8,12,13,16	8,12,13,16	8,16	12,13
D21S11	30,33.2	30,33.2	30	30,33.2
D18S51	13,16	13,16	16	13,16
D2S441	10,11	10,11	10	10,11
D19S433	13,13.2	13,13.2	13	13,13.2
TH01	8,9	8,9	9	8
FGA	19,21,23,26	19,21,23,26	19,21	23,26
D22S1045	11,12,15	11,12,15	11,12	12,15
D5S818	10,11,12	10,11,12	12	10,11
D13S317	9,10,12,13	9,10,12,13	9,13	10,12
D7S820	8,11	8,11	8	8,11
SE33	20,27.2,28.2	20,27.2,28.2	27.2,28.2	20,27.2
D10S1248	14,15,16	14,15,16	15,16	14
D1S1656	11,15,16,17	11,15,16,17	11,16	15,17
D12S391	18,19,20,21	18,19,20,21	20,21	18,19
D2S1338	23,25	23,25	23,25	23
AMELOGENIN	XY	XY	XX	XY
DYS391[a]	11	11	–	11
Y INDEL[a]	2	2	–	2

[a]Markers indicative of male DNA

4.4.2 Case 2: A Case of Sexual Assault Followed by the Murder of an 8-Year-Old Girl Child Confirmed by Using DNA Technology

In 2013, an 8-year girl was found brutally murdered by crushing her face with a heavy stone after cutting her neck. With few clues, a suspect was arrested, and the high-level investigation team sent the case exhibits for biological examination to Regional Forensic Science Laboratory, Bhopal. The biological and serological examination suggested the presence of spermatozoa and also the similar blood group on the seized exhibits.

After this confirmation, the case was referred for DNA fingerprinting examination. As per the directions given by DNA Fingerprinting Unit, State Forensic Science Laboratory, Sagar, Madhya Pradesh, India, the samples from parents of the deceased were also sent to the laboratory for confirming the identity of the girl child. This was directed because the head was crushed, and the girl was identified by her dress (frock). The exhibits on which male mixed DNA was expected were processed using

the Phenol Chloroform Isoamyl Alcohol (PCI) differential DNA extraction method, and the teeth of girl and the reference samples from the parents were processed by Phenol Chloroform Isoamyl Alcohol (PCI) (Köchl et al. 2005) organic DNA extraction method. The obtained DNA was amplified with the then-available 15 autosomal marker Ampf*l*STR® Identifiler plus PCR Amplification Multiplex kit (ABI/Now, Thermo Fisher Scientific). The amplicons were run on Genetic Analyser ABI 3100XL, and the obtained data were analysed using GeneMapper® ID software v3.2. A mixed autosomal STR profile was detected with the source of both the articles of the victim. All the alleles in the DNA profile of the victim and accused were observed in the mixed DNA profile obtained from the source of the victim. Thus, the obtained results confirmed the transfer of the deceased's DNA on the source clothes of the accused, and weapon of offence (stone and trowel (a mason tool) (Table 4.2). The results also confirmed the presence of the DNA of the accused on the clothes (Frock, Stain 3) of the deceased (Table 4.2). The biological relationship of the girl child (deceased) is confirmed by matching the DNA profile of parents (Table 4.3).

4.4.3 Case 3: A Gang-Rape Case Solved by Using Present-Day DNA Technology

A lady was sexually assaulted by three persons, who used condoms that were recovered from the crime scene. The investigating officer contacted the experts at DNA division and very wisely seized all the three condoms, tied those and sent them to the DNA Unit, State FSL, Sagar. The laboratory carried out the DNA examination. In this gang-rape case, the blood sample of the victim, three suspects, along with a vaginal slide of the victim, were sent for DNA examination. DNA was extracted from the reference blood sample of victims and suspects, using Phenol Chloroform Isoamyl Alcohol (PCI) organic DNA extraction method. Outer and inner portions of all three condoms were processed separately. DNA from condoms and vaginal smear slide were extracted by phenol-chloroform extraction. Obtained DNA was quantified using ABI 7000 Real-Time PCR (Thermo Fisher Scientific) and Quantifiler™ Duo DNA Quantification Kit (Thermo Fisher Scientific) using the manufacturer's protocol. One nanogram of DNA was amplified using PowerPlex® 16HS Multiplex kit (Promega Corporation) as per the recommendations of the manufacturer. The amplified DNA fragments were separated using Genetic Analyser ABI 3100 as per recommendations of the manufacturer, and the obtained data were analysed using GeneMapper software v3.2. The obtained genotyping results are presented in Tables 4.4 and 4.5.

The obtained results, in this case, using autosomal STR, confirmed that all three condoms were used with the same lady. The profile obtained from the blood sample and vaginal slide of the victim and DNA profile observed from the outer surface of all three condoms were the same, confirming that all three condoms were used with the same lady (Table 4.4). However, three different DNA profiles were observed from the inner part of all three condoms (Table 4.5), confirming the involvement of

Table 4.2 Autosomal STR DNA profile obtained in a case of sexual assault followed by the murder of the 8-year-old girl child

Genetic markers	Vaginal smear slide and swab of deceased	Teeth and muscle attached with a piece of Jaw	Frock of deceased stain 1 and 2	Stones and trowel mason tool	Frock of deceased stain 3	Clothes of accused	Blood sample of accused
D8S1179	10,15	10,15	10,15	10,15	13,14	10,15	13,14
D21S11	28,31	28,31	28,31	28,31	29,31.2	28,31	29,31.2
D7S820	8,10	8,10	8,10	8,10	8,10	8,10	8,10
CSF1PO	11,12	11,12	11,12	11,12	11,12	11,12	11,12
D3S1358	17,18	17,18	17,18	17,18	16,17	17,18	16,17
THO1	6,9	6,9	6,9	6,9	7,9	6,9	7,9
D13S317	11,12	11,12	11,12	11,12	8,8	11,12	8,8
D16S539	11,13	11,13	11,13	11,13	9,13	11,13	9,13
D2S1338	21,23	21,23	21,23	21,23	20,20	21,23	20,20
D19S433	14,14	14,14	14,14	14,14	13,14	14,14	13,14
vWA	18,19	18,19	18,19	18,19	14,19	18,19	14,19
TPOX	8,11	8,11	8,11	8,11	8,8	8,11	8,8
D18S51	17,19	17,19	17,19	17,19	14,15	17,19	14,15
D5S818	11,11	11,11	11,11	11,11	11,13	11,11	11,13
FGA	20,22	20,22	20,22	20,22	22,24	20,22	22,24
AMELOGENIN	XX	XX	XX	XX	XY	XX	XY

Table 4.3 Autosomal STR DNA profile in the case of the brutal murder of a girl child confirming parentage

Genetic markers	Blood sample of Father	Teeth and muscle attached with a piece of jaw	Blood sample of Mother
D8S1179	14,15	10,15	10,14
D21S11	31,31.2	28,31	28
D7S820	10	8,10	8,13
CSF1PO	11,12	11,12	12
D3S1358	16,17	17,18	17,18
THO1	6,8	6,9	9
D13S317	10,12	11,12	8,11
D16S539	8,11	11,13	11,13
D2S1338	20,21	21,23	23,24
D19S433	11.2,14	14,14	14,16
vWA	17,18	18,19	19
TPOX	9,11	8,11	8
D18S51	12,19	17,19	15,17
D5S818	11,12	11,11	9,11
FGA	22	20,22	20,23
AMELOGENIN	XY	XX	XX

Table 4.4 DNA profile obtained from the outer surface of the three condoms and blood sample and vaginal slide of the victim

Genetic markers	DNA profile obtained from outer portion of condom 1, 2 and 3	DNA profile obtained from the blood sample of the victim	DNA profile obtained from vaginal slide sample of the victim
D3S1358	14,17	14,17	14,17
TH01	6	6	6
D21S11	29,30	29,30	29,30
D18S51	11,16	11,16	11,16
Penta E	14,16	14,16	14,16
D5S818	10,11	10,11	10,11
D13S317	11,12	11,12	11,12
D7S820	9,12	9,12	9,12
D16S539	12	12	12
CSF1PO	11	11	11
Penta D	10,11	10,11	10,11
vWA	17	17	17
D8S1179	12,13	12,13	12,13
TPOX	11	11	11
FGA	21,23	21,23	21,23
AMELOGENIN	XX	XX	XX

Table 4.5 DNA profile obtained from the inner surface of the three condoms

Genetic markers	The inner portion of condom 1	The inner portion of condom 2	The inner portion of condom 3	Suspect 1	Suspect 2	Suspect 3
D3S1358	15,17	15,16	16,16	15,16	12,14	15,17
TH01	9,9	7,9.3	9,9	7,9.3	7,10	9,9
D21S11	29,34.2	32.2,34.2	30,32.2	32.2,34.2	24,27	29,34.2
D18S51	14,21	14,16	14,15	14,16	9,11	14,21
Penta E	7,13	7,19	10,12	7,19	8,11	7,13
D5S818	12,13	11,13	11,11	11,13	12,12	12,13
D13S317	11,12	10,11	8,11	10,11	8,12	11,12
D7S820	8,8	11,11	7,10	11,11	13,14	8,8
D16S539	10,11	11,12	11,13	11,12	13,15	10,11
CSF1PO	11,11	9,13	10,12	9,13	12,12	11,11
Penta D	10,14	11,13	10,12	11,13	11,14	10,14
vWA	14,17	15,18	17,19	15,18	19,21	14,17
D8S1179	14,15	10,16	15,17	10,16	17,17	14,15
TPOX	8,11	9,11	9,11	9,11	12,13	8,11
FGA	19,24	20,23	24,24	20,23	21,23	19,24
AMELOGENIN	XY	XY	X,Y	X,Y	X,Y	X,Y

three distinct persons. The two autosomal STR profiles obtained from two different inner portions of condoms were found to match with the profiles of the two suspects (Table 4.5).

4.4.4 Case 4: A Case of Identification of a Mutilated Decomposed Body

A mutilated body in an advanced stage of decomposition was recovered from a drain. The medical doctor that conducted the autopsy preserved the skull (with missing frontal bone and teeth intact) and long bone (femur) for DNA-based identification. DNA fingerprinting analysis was carried out by the Forensic Science Laboratory. The blood samples of the presumed father and mother were also sent to the lab, along with their written informed consent. The DNA laboratory used an organic extraction method for all the samples. It was followed by the amplification of 15 autosomal STR markers by a multiplex polymerase chain reaction. The PCR product was detected by Genetic Analyser 3100 and analysed by GeneMapper v3.2. Half of the alleles of each genetic marker analysed for presumed father as well as the mother's blood sample were found to be present in the profile generated from the bone and skull of the deceased (Table 4.6). In this process, the identification of the deceased individual was established to be the biological son of the putative father and mother.

Table 4.6 Autosomal STR DNA profile obtained from the blood samples of presumed father and mother and the unidentified skull and femur bone using PowerPlex$^{\hat{U}}$ 16HS Multiplex kit (Promega Corporation)

Genetic markers	Blood sample of the presumed father	Unidentified skull/femur	Blood sample of the presumed mother
D3S1358	15,18	17,18	15,17
TH01	8,9	9,9	9,9
D21S11	31.2,32.2	31,31.2	31,32.2
D18S51	14,17	16,17	14,16
Penta E	11,20	11,18	15,18
D5S818	11,11	11,13	11,13
D13S317	11,11	11,11	8,11
D7S820	10,11	10,11	8,10
D16S539	11,11	11,11	11,11
CSF1PO	12,12	11,12	10,11
Penta D	10,10	9,10	9,10
vWA	14,15	14,17	17,18
D8S1179	14,15	15,15	13,15
TPOX	8,11	11,11	8,11
FGA	23,24	22,23	20,22
AMELOGENIN	X,Y	X,Y	X,X

4.4.5 Case 5: Child Sexual Abuse Case

In the sexual cases involving minors (age of girl child less than 10 years), it is challenging to prove the guilt as the vagina of the girl concerned is not prepared at this age to bear the force generated during the attempt of an adult penetration. As a consequence, the vagina of the girl gets ruptured and results in severe bleeding. With the experience of practically handling cases, as forensic DNA expert, the authors are of the opinion that these cases are usually linked by observing the immediate transfer of victim's blood after the incidence. In these cases, penile swabs and the undergarments of accused/suspect have been observed to be better objects in proving the guilt by using forensic DNA technology instead of vaginal slides and swabs. In a case of sexual assault against a girl child aged 5 years, vaginal swab and slide of the victim, clothes of victim, undergarment and a penile swab of the suspect along with referral blood sample of the suspect were received for DNA examination. DNA was extracted from the reference blood sample of the suspect by Automated DNA extraction system (12 GC, PSS, Japan). DNA from vaginal swab and slide of the victim, clothes of victim, undergarment and a penile swab of the suspect was extracted using Phenol Chloroform Isoamyl Alcohol (PCIA) organic DNA extraction method. The final obtained DNA was concentrated using 30 KD Microcon concentration device (Millipore). The final DNA was eluted to 20 μL in Tris EDTA buffer. Differential extraction method (Butler 2012) was adopted in extracting DNA from vaginal swab and slide of the victim. Obtained DNA was quantified using ABI

7000 Real-Time PCR (ThermoFisherScientific) and Quantifiler™ Duo DNA Quantification Kit (ThermoFisher Scientific) using the manufacturer's protocol. Presence of male DNA could not be observed by RT PCR analysis. One nanogram of DNA was amplified using Amp*fl*STR identifiler plus multiplex kit (ABI/Thermo Fisher Scientific) and the then-available Amp*fl*STR Y filer kit as per the recommendations of the manufacturer, except the half-reaction volume was used. The amplified DNA fragments were separated using Genetic Analyser ABI 3100 as per recommendations of the manufacturer, and the obtained data were analysed using GeneMapper software v3.2. No male DNA could be detected by the use of Y-STRs multiplex kit. The obtained genotyping results with autosomal STR typing are presented in Table 4.7. Presence of the female DNA profile was observed from the source of vaginal swab and slide of the victim was observed on the source of penis and undergarment of the suspect, thus establishing the body fluid transfer conclusively using forensic DNA technology.

4.5 Conclusion

DNA forensics will continue to progress. Modern DNA forensic methods are authoritative and subtle. As the sensitivity of DNA techniques increases, so does the risks of contamination and secondary transfer of DNA material. The issues of contaminations during every step from sampling, analysis, and finally to report,

Table 4.7 DNA profile observed in the case using Amp*fl*STR® identifiler plus kit

Genetic markers	Article A Vaginal slide of the victim	Article B Vaginal swab of the victim	Article F Penile swab of the suspect	Article H Undergarment of the suspect
D8S1179	10,13	10,13	10,13	10,13
D21S11	29,33.2	29,33.2	29,33.2	29,33.2
D7S820	11,11	11,11	11,11	11,11
CSF1PO	9,13	9,13	9,13	9,13
D3S1358	17,18	17,18	17,18	17,18
THO1	7,7	7,7	7,7	7,7
D13S317	8,12	8,12	8,12	8,12
D16S539	10,11	10,11	10,11	10,11
D2S1338	20,21	20,21	20,21	20,21
D19S433	13,16	13,16	13,16	13,16
vWA	13,19	13,19	13,19	13,19
TPOX	8,11	8,11	8,11	8,11
D18S51	15,16	15,16	15,16	15,16
D5S818	11,11	11,11	11,11	11,11
FGA	24,25	24,25	24,25	24,25
AMELOGENIN	XX	XX	XX	XX

proper sampling and proficiency of scientists dealing with DNA should be taken care of. With the increasing reliance on the technique, it is rare to see any court reject the DNA evidence on the basis on the invalidity of the fundamental scientific theory. However, with practice, precaution and precision, from the inception to the end, we would be able to utilise the optimum potential of DNA evidence from the crime scene to conviction.

References

Alghafri R, Pajnič IZ, Zupanc T, Balažic J, Shrivastava P (2018) Rapidly mutating Y-STR analyses of compromised forensic samples. Int J Legal Med 132:397. https://doi.org/10.1007/s00414-017-1600-z

Allen RW, Pogemiller J, Joslin J, Gulick M, Pritchard J (2008) Identification through typing of DNA recovered from touch transfer evidence: parameters affecting yield of recovered human DNA. J Forensic Identif 58(1):33–41

Arnaud CH (2017) Thirty years of DNA forensics: How DNA has revolutionized criminal investigations. Chem Eng News 95:16. Available from: https://cen.acs.org/articles/95/i37/Thirty-years-DNA-forensics-DNA.html

Ashcroft J, Daniels DJ, Hart SV (2002) Using DNA to solve cold cases. U.S. Department of justice office of justice programs. Available from: https://www.ncjrs.gov/pdffiles1/nij/194197.pdf

Badiye A, Kapoor N, Menezes RG (2019) Chain of Custody (Chain of Evidence). In: StatPearls [Internet]. StatPearls Publishing, Treasure Island (FL). Available from: http://www.ncbi.nlm.nih.gov/books/NBK551677/

Baxter E Jr (2015) Complete crime scene investigation handbook. CRC Press, Boca Raton, FL

Butler JM (2012) DNA extraction methods. In: Advanced topics in forensic DNA typing: methodology. Elsevier, Waltham, MA

Daly DJ, Murphy C, McDermott SD (2012) The transfer of touch DNA from hands to glass, fabric and wood. Forensic Sci Int Genet 6:41–46

Dash HR, Shrivastava P, Das S (2020) Principles and practices of DNA analysis: a laboratory manual for forensic DNA typing. Springer, New York, NY

Goodwin W, Linacre A, Hadi S (2007) An introduction to forensic genetics. John Wiley & Sons, Ltd, New York, NY

Goray M et al (2019) DNA transfer: DNA acquired by gloves during casework examinations. Forensic Sci Int Genet 38:167. https://doi.org/10.1016/j.fsigen.2018.10.018

Harris W (2001) How DNA evidence works. HowStuffWorks.com. Available from: https://science.howstuffworks.com/life/genetic/dna-evidence.htm. Accessed 13 April 2020

Hodge C (2017) The shifting science of DNA in the courtroom. CNN Health. Available from: https://edition.cnn.com/2017/06/09/health/dna-technology-forensic-evidence/index.html

Jeffreys AJ et al (1985a) Hypervariable minisatellite regions in human DNA. Nature 314:67–73

Jeffreys AJ et al (1985b) Positive identification of an immigration test-case using human DNA fingerprints. Nature 317:818–819

Jeffreys AJ et al (1985c) Individual-specific fingerprints of human DNA. Nature 316:76–79

Kaye H (2007) The science of DNA identification: from the laboratory to the courtroom (and beyond). Minnesota J Law Sci Technol 8:4. Available from: https://scholarship.law.umn.edu/mjlst/vol8/iss2/4/

Kapoor N, Badiye A (2015a) Sex differences in the thumbprint ridge density in a central Indian population. Egypt J Forensic Sci 5:23–29. https://doi.org/10.1016/j.ejfs.2014.05.001

Kapoor N, Badiye A (2015b) An analysis of whorl patterns for determination of hand. J Forensic Leg Med 32:42–46. https://doi.org/10.1016/j.jflm.2015.02.015

Kapoor N, Badiye A (2015c) Digital dermatoglyphics: a study on muslim population from India. Egypt J Forensic Sci 5:90–95. https://doi.org/10.1016/j.ejfs.2014.08.001

Kleypas DA, Badiye A (2020) Evidence collection. In: StatPearls [Internet]. StatPearls Publishing, Treasure Island (FL). Available from: https://www.ncbi.nlm.nih.gov/books/NBK441852/

Köchl S, Niederstätter H, Parson W (2005) DNA extraction and quantitation of forensic samples using the phenol-chloroform method and real-time PCR. In: Forensic DNA typing protocols. Springer, New York, NY, pp 13–29

Lee HC, Ladd C (2001) Preservation and collection of biological evidence. Croat Med J 42 (3):225–228

Locard E (1930) Te analysis of dust traces. Part 1. Am J Police Sci 1:276–298

Lynch M (2003) God's signature: DNA profiling, the new gold standard in forensic science. Endeavour 27(2):93–97

Magalhães T, Dinis-Oliveira RJ, Silva B, Corte-Real F, Nuno VD (2015) Biological evidence management for DNA analysis in cases of sexual assault. Sci World J 2015:365674. https://doi.org/10.1155/2015/365674. PMID: 26587562; PMCID: PMC4637504

National Research Council (US) Committee on DNA Forensic Science (1996) 6. DNA Evidence in the Legal System. In: An update. the evaluation of forensic DNA evidence. National Academies Press, Washington, DC. Available from: https://www.ncbi.nlm.nih.gov/books/NBK232607/

Phipps M, Petricevic S (2007) The tendency of individuals to transfer DNA to handled items. Forensic Sci Int 168:162–168

Polley D, Mickiewicz P, Vaughn M, Miller T, Warburton R, Komonski D, Kantautas C, Reid B, Frappier R, Newman J (2006) An investigation of DNA recovery from firearms and cartridge cases. Can Soc Forensic Sci J 39(4):217–228

Shrivastava P, Jain T, Trivedi VB (2016) DNA fingerprinting: a substantial and imperative aid to forensic investigation. Eur J Forensic Sci 3(3):23–30

de Villiers WP (2014) Fingerprint comparison evidence has been under sustained attack in the United States of America for the last number of years: Is the critique with regard to reliability sufficiently penetrating to warrant the exclusion of this valuable evidence? Int J Law Crime Justice 42:54. https://doi.org/10.1016/j.ijlcj.2014.01.004

RNA- and DNA-Based Identification of Body Fluids

5

Arjun Rao, Manisha Rana, Ashish Pradhan, and Moumita Sinha

Abstract

The most common form of evidence to forensic investigators is the body fluids collected at crime scenes. They ascertain the suspect or victim, they provide valuable DNA evidences and can play a vital role in acquit an innocent individual. The determination of a specific bodily fluid is predominantly the initial step as the body fluid composition is very relevant to the further investigation process. The ability to identify and report an unexplained stain at the scene of crime without waiting for the laboratory results is another very important phase in the forensic body fluids analysis. Many forms of detection methods for body fluids have been known for over a century, such as alternative light source, immunological tests, spectroscopic techniques, chemical methods, catalytic tests, and microscopic methods. Although these modern forms of detection of body fluids are often definitive, these are done at a time with only one body fluid. Currently the usage of molecular genetic based approaches using DNA methylation detection or RNA-based profiling methods has recently conquered to replace the traditional body fluids identification methods.

Keywords

Blood · Saliva · Semen · Forensic · RNA · DNA

5.1 Introduction

Individual's involvement in a crime can be detected and identified by the body fluids traced at the crime scene event. These types of evidences found at the scene of crime are among the most ubiquitous in nature especially to forensic investigation

A. Rao (✉) · M. Rana · A. Pradhan · M. Sinha
Department of Forensic Science, Guru Ghasidas Vishwavidyalaya, Bilaspur, CG, India

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_5

purposes. Body fluids identification and detection is a vital part in forensic practices; however, many biological fluids are invisibly present in the crime scene or present in very small quantity due to which the easily identification is not possible. The presence of biological fluids not only provides type and origin of the samples but also provides clues in reconstructing crime scene and linkages between innocent and actual criminalists.

Biological fluids like blood (including menstrual stains), semen, saliva, urine and vaginal stains, skin, and tissues are the prime interest of forensic scientists in various crime cases. In a crime scene approach the identity of these biological fluids are not straightforward, due to their invisible presence in very minute quantities. For the identification of sequence of events at the crime scene, these biological fluids can be used as crucial indicators. Such as blood stains are a strong indicator of murder, some form of assault, or struggle, whereas upon the detection of semen or vaginal stains able to specify physical and or sexual assault.

After potential identification of body fluids at scene of crime, a specific form of light is required to visualize these stains or the spot chemical examination is required. The biological identification of these stains previously based on some limited chemical methods which were often inadequate in sensitivity and specificity, whereas the confirmatory tests based on some microcrystal tests for the identification of body fluids in forensic-related cases. The techniques apply were based on chemical tests, enzyme catalytic reactions, immunological tests, spectroscopic based tests, and microscopic examinations. The presumptive/preliminary tests which are also used for screening the biological fluids even at the crime scene or at the forensic laboratories setup; some of them are, Phenolphthalein (Kastle–Meyer test), and Luminol test (for minute traces) for blood identification, Phadebas test for saliva identification, Jaffe's test for urine identification, Acid Phosphatase test for semen identification. The other tests are based on microscopic examination (confirmatory tests) which will convincingly identify the presence and the nature of such biological fluids. These gold standard evidences contain valuable genetic information to identify a suspect or victim as well as either acquit an innocent individual. Technology has rapidly evolved to reveal the individual identity of an unidentified stain at the scene of crime scene, without much waiting for the laboratory results to arrive is a crucial breakthrough in the advancement of body fluids analysis in forensic scenario. An evident of the last decade is the progressive development of the laser technology and novel light detectors which have either gave birth to spectroscopic methods and/or molecular-based approaches for the detection of biological fluids.

In the field of body fluids detection, the latest developments must culminated in a range of novel methods and approaches which allows (1) first time identification of some body fluids, (2) experiments enabling nondestructive research at the scene of crime for more study via efficient DNA profiling, and (3) research lab-based protein and RNA techniques allowing for the requirements.

5.2 The Drawbacks and Benefits of Chemical Tests

For several years, preliminary chemical studies have been utilized and are still playing a significant role to identify a region of significance for DNA study and forensic investigation. Ample of reviews and reports for their accuracy and precision are available (Virkler and Lednev 2009; Gaensslen 1983; Tobe et al. 2007). These chemical measures are not unique to humans and usually are administered sequentially where there may be a mixed body fluid. Most depend on body fluids enzyme properties, and many reagents are sample harmful and/or obstruct downstream procedures.

Stains on dark surfaces or non-visible stains face challenges to detect and similarly been deciphered with alternative light sources using the auto fluorescence provided by some body fluids (Kobus et al. 2002; Vandenberg and van Oorshot 2006). The effectiveness of such approaches may be influenced by differences between body fluids and various surfaces, and application to these light sources may induce DNA damage in the stain. The super reactive luminol method which is used in the darker areas to identify the blood stains, but with a broad variety of chemicals, it provides false positives and devalues any stain that might be needed for further study (Webb et al. 2006; Quinones et al. 2006).

5.3 Limitations and Drawbacks of Antibody-Antigens Studies

The reasonable degree of accuracy and sensitivity has been found while using immuno chromatographic and enzyme-linked immunosorbent assay (ELISA) for the detection of body fluids (Akutsu et al. 2012a, b; Simich et al. 1999; Casey and Price 2010). Such measures determine the identity of the antigen in question instead of the antigen behavior. Environmental conditions may influence antigen–antibody interaction contributing to false positive or negative outcomes, and the elevated-dose hook effect affects both of these studies. These experiments are not intended to identify staining areas or in a systematic fashion, so a part of the sample is extracted so solubilized in both of these methods prior to the study.

5.4 Blood

Chemical blood tests are susceptible but lacks precision and tend to rely onto the catalytically activity of the hemoglobin groups and the oxidant, including plant peroxidases, which can produce false positive results (Virkler and Lednev 2009; Tobe et al. 2007; Cox 1991). Assays include Leucomalachite Green, named as less sensitive but also more specified tests; Tetramethylbenzidine (TMB), Ortho-tolidine, and Phenolphthalein (Kastle–Meyer test), some form of high specificity sensitive testes for the identification of blood (Gaensslen 1983; Webb et al. 2006; Cox 1991). Confirmatory yet often cumbersome testing requires a microscopy of red and white blood cells and the formation of crystals with hemochromogenic crystals

(Takayama) or hematinic crystals (Teichman) that indicate the blood existence (Gaensslen 1983).

Some tests are based on the monoclonal antihuman antihemoglobin antibodies which shows cross reaction with primates blood and mustelidase, presumably triggered by a specific sequence of amino acid in hemoglobin alpha chain, in the ABAcard® Hematrace® (Misencik and Laux 2007; Johnston et al. 2003). RSID™ blood tests (Independent Forensics, Hillside, IL, U.S.) screening test detects the rich and precisely red blood cell membrane of glycophorin A protein without any cross reactions (Schweers et al. 2008).

5.5 Saliva

An α-amylase-based detection The Phadebas® test is the regular standardized test for the identification of presence of saliva. Due to the small traces of the α-amylase enzyme are known to found in different other body fluids such as breast milk, blood, semen, urine, fecal urine, and other mammals as salivary and pancreatic sources this test is not confirmatory for the human saliva identification (Kipps and Whitehead 1975; Auvdel 1986; Keating and Higgs 1994). The SALIgAE® test from Abacus Diagnostics provides an substitute test for saliva based on calorimetric approach, which is known to be more sensitive target reagent than the Phadebas® test (West Hills, CA, USA) (Pang and Cheung 2008; Liang and Roy 2014).

False positive values have been traced in breast milk, rat saliva, clean urine and fecal/feces and semen samples. The RSID™ test for saliva is focused on anti-amylase antibodies of antihuman salivary (Old et al. 2009). The above test has been found to be more flexible and more accurate than the two tests Phadebas® and SALIgAE® tests, respectively (Casey and Price 2010; Pang and Cheung 2008).

5.6 Semen

In the lack of microscopically detected spermatozoa, semen usually takes the presumptive examination that detects phosphatase of seminal acids, an enzyme isolated from the prostate gland, although that not directly specifically connected with the fluid of seminal substances (Gaensslen 1983; Kind 1957). While no longer commonly used, confirmatory crystal tests for semen detection are present, along with the Florence test based on choline crystal formation (Gaensslen 1983).

Semen is detected for a certain time using glycoprotein, the prostate-based antigen (Kallikrein 3 and PSA or P30) (Gaensslen 1983; Graves 1995). The widely used immunological tests nowadays which includes semiquant PSA test SERATEC, p30 ABA card® and PSA testing Biosign®, while some of these test kits shows false positive reactions against vaginal stains, urine, breast milk, postmortem semen samples (Graves 1995; Yu and Diamandis 1995; Lunetta and Sippel 2009).

Semenogelin is often used to identify semen via immunochromatographic tests (Sato et al. 2004; Pang and Cheung 2007; Old et al. 2012). The RSID™-semen test

has been noticed less susceptible while comparing with the other test kits like ABAcard®, P30 and SERATEC® PSA, and some test results are occupied with inaccurate results (Boward and Wilson 2013).

5.7 Vaginal Secretions and Menstrual Blood

The differentiating feature between menstrual blood and vaginal secretions is found to be very difficult in classification. The identification of the vaginal wall glycogenated squamous epithelial cells through Lugol's staining, the endometrial cells microscopic identification and tracing the lactate dehydrogenase isoenzymes 4 and 5 are not the definitive methods for tracing the vaginal cells (Jones and Leon 2004; Stombaugh and Kearney 1987).

Immunochromatography examinations are known as potential menstrual blood checks for D-dimers, fibrin which is soluble and degradation drug medically identified for thrombosis diagnosis (Baker et al. 2011; Holtkoetter et al. 2015). One different method ELISA is utilized which focuses on estrogen receptor α and fibrinogen to differentiate the peripheral blood from menstrual blood, whereas none of the other body fluids showed any cross reactions (Gray et al. 2012).

5.8 Urine

Urine stain localization is complicated because it is usually translucent, light, and ubiquitous. Usually, the presumptive tests are focused on urea, urease and uric acid analysis. Such tests are not unique since sweat and other urea-substances often respond favorably (Gaensslen 1983; Huang et al. 2002). Creatinine checks were also used for urine detection (Gaensslen 1983). Tamm–Horsfall Glycoprotein (TMP) protein has also been found in animals' urine and has already been documented and included in the RSID™ urine study of humans. TMP tends to be sufficient as a special urine test, but the existence of vaginal fluid may complicate the test results and it could be difficult to read the test if the blood is in the specimen (Akutsu et al. 2012a, b).

5.9 Sweat

No realistic sweat recognition test is available to date. While DNA is often recuperated and profiled in clothing areas which may contain sweat, there is little study. ELISA-based testing of G-81 and dermicidin has developed but have not been commonly used to detect sweat-specific proteins (Sagawa et al. 2003; Sakurada et al. 2010).

5.10 Discovery of New (Protein) Markers

The identification and characterization of common markers are the basis for ELISA and immunochromatographic studies. 2-DHPLC, MS, and QMS were all inured to establish protein profiles typical of all six primary forensically important body fluids i.e., blood, menstrual blood, saliva, semen, vaginal substance, and skin and also acquire new candidate protein markers for instance osteopontin and uromodulin to identify urine (Legg et al. 2014). Statherin in Saliva and Semenogelin 1 and 2 in Seminal fluids are used for the examination of mRNA as markers.

5.11 Identifications Based on RNA and DNA Technologies

Ribo Nucleic Acid (RNA) and Deoxyribo Nucleic Acid (DNA) are being used progressively for so many new forensic applications, including: body fluid detection, assessing postmortem and age decay of RNA; the estimation of injury age through the analysis of reactive gene expression changes; and determining the cause of death (Bauer et al. 2003; Anderson et al. 2010; Cecchi 2010; Palagummi et al. 2014). These approaches including the utilization of messenger RNA, micro RNA, other forms DNA methylation, and microbial characterization of each and everybody fluid.

The list of potent markers in various body fluids identified using RNA- and DNA-based methods are given in Table 5.1.

5.12 mRNA-Based Methods

Most of the body fluids contain several types of cells, each expressing an individual mRNA transcript pattern. The layout and carrying out of mRNA profiling in research associated to forensic science is focused on using these multicellular transcriptomes. mRNA in body fluids deposited on a range of surfaces is now generally understood to be stable and can be collected in an appropriate quantity and consistency for analysis of various samples (Bauer et al. 1999; Setzer et al. 2008; Visser et al. 2011; Kohlmeier and Schneider 2012; Fox et al. 2014; Alvarez et al. 2004).

A benefit of mRNA profiling is that the recovery of RNA from tissues, with a variety of specific RNA extraction procedures, can be incorporated into a standard workflow of DNA profiling.

Generalizations between different commercial RNA removal methods have revealed different yields and diagnose of DNA and RNA, of no better solution than the other; collective studies have confirmed this finding. The enormous amount of transcript and consistency by means of substitute markers designed for the identical fluid with unlike susceptibilities in the same body are attributed to the capability to detect a messenger RNA transcript that is of significance (Zubakov et al. 2008; Haas et al. 2014, 2015, 2011; Van den Berge et al. 2014; Fleming and

Table 5.1 List of potential markers in various body fluids with their chromosomal location

Body fluid	Markers	Official symbol	Chromosomal location
Blood	Ankyrin 1	ANK 1	8p11.21
	Glycophorin A	GYPA	4q31.21
	Beta-spectrin	SPTBN1	2p16.2
	Alpha hemoglobulin	HBA	16p13.3
	Beta hemoglobulin	HBB	11p15.4
	Porphobilinogen deaminase	HMBS	11q23.3
	Amino-levulinate synthase 2	ALAS2	Xp11.21
Seminal fluid	Protamine 1	PRM 1	16p13.13
	Protamine 2	PRM 2	16p13.13
Saliva	Histatin	HTN	4q13.3
	Statherin	STATH	4q13.3
Vaginal fluid	Beta-defensin 1	DEFB1	8p23.1
	Mucin 4	MUC4	3q29
	Cytochrome P450	CYP2B7P1	19q13.2
	Myozenin 1	MYOZ1	10q22.2
Menstrual blood	Matrix metallopeptidase 7	MMP7	11q22.2
	Matrix metallopeptidase 11	MMP11	22q11.23
Epidermal cells	Late cornified envelope 1 C	LCE1C	1q21.3
	Loricrin	LOR	1q21.3

Harbison 2010a, b; Lindenbergh et al. 2012, 2013a, b). In presumptive tests where these comparisons are made, mRNA profiling is comparable (Haas et al. 2011).

The creation of a robust mRNA quantification method will boost RNA profiling; excessive template leads to target amplification and an increased incidence of "nonspecific artifact" (Haas et al. 2011). The present steps are applied to the entire nucleic acid or Ribo Nucleic Acid which happen to be not human specific and quantified mostly by utilizing techniques like UV-Visible spectrometry, fluorometry, Nanodrop ND-1000, Agilent 2100 Bioanalyzer System, or Quant-iTMRiboGreen® RNA kits (Visser et al. 2011; Grabmuller et al. 2015; Fleming and Harbison 2010a, b; Juusola and Ballantyne 2005).

A flexible method of analysing low-abundance mRNA extracted from minute amount of samples is by reverse transcriptase polymerase chain reaction (RT-PCR) (Kohlmeier and Schneider 2012; Zubakov et al. 2008; Fleming and Harbison 2010a, b; Hanson et al. 2012; Haas et al. 2009; Bustin 2000; Juusola and Ballantyne 2007). The most reasonable approach in forensic work is the endpoint reverse transcriptase PCR in conjunction with capillary electrophoresis, which allows the concurrent finding of multiple body fluids and thereby minimizes sample utility and background impacts. Juusola and Ballantyne first introduced this approach in 2005, followed by others (Fleming and Harbison 2010a, b; Lindenbergh et al. 2012, 2013a, b; Juusola and Ballantyne 2005; Richard et al. 2012; Roeder and Haas 2013). The quantitative RT-PCR (qRT-PCR) measures the variation in between the focused transcript and a household gene called the ΔCT (Zubakov et al. 2008;

Bustin 2000; Juusola and Ballantyne 2007; Nussbaumer et al. 2006). The benefit of this very reactive method is the possibility of setting numerical thresholds. The existing dyes limitation, which limits the sum of markers that will be able to reach in a solo reaction, represents a disadvantage. In order to address this constraint, enhancements were suggested, such as a high-resolution melting study (Hanson and Ballantyne 2013a, b).

Findings between RT-PCR and qRT-PCR indicate that perhaps the detected abundance of mRNA transcript to a large or small DNA portion (Van den Berge et al. 2014; Harteveld et al. 2013) can still not be allocated with any of these methods. New techniques have recently been developed such as isothermal amplification mediated by real-time RT loops, offering the same sensitivity and precision but then much simpler and faster studies (Su et al. 2015). The housekeeping genes constitutively expressed make available a reference point and determine the efficiency of a reaction by means of capillary electrophoresis approaches, and they are important in quantitative techniques (Bustin 2000; Moreno et al. 2012).

The desired housekeeping genes expressing themselves across all tissues, need not be different substantially between people and is not greatly influenced by physiological/pathologic problems. A variety of housekeeping genes, including ACTB, GAPDH, B2M, S15, UCE, TEF, UBC, 18S RNA and G6PD, are being used in forensic research (Zubakov et al. 2008; Haas et al. 2014; Fleming and Harbison 2010a, b; Lindenbergh et al. 2012, 2013a, b; Juusola and Ballantyne 2007; Roeder and Haas 2013; Moreno et al. 2012; Jakubowska et al. 2013; Donfack and Wiley 2015).

Even between similar samples, RT-PCR can show significant variation. This may be caused by many factors including different secondary structures of RNA transcripts, stochastic variations in samples, and inhibition and consistency of RNA. Housekeeping genes is no exemption and there is overall conformity that the quantity of transcripts will differ from individual to individual in different fluid (Fleming and Harbison 2010a, b; Roeder and Haas 2013; Bustin et al. 2005; Park et al. 2007) e.g., the prevalence of the household genes relative toward the body fluid as specific genes in buccal cells and semen is quite small and there is probably hardly any appropriate housekeeping gene meant for every single body fluid. The messenger RNA (mRNA) markers are chosen from a candidate gene method for which the gene function is normally known (Fleming and Harbison 2010a, b; Lindenbergh et al. 2012, 2013a, b; Juusola and Ballantyne 2005). Instead, a discovery technique such as a comparative multicellular transcriptome analysis based on the microarray or RNA sequencing (transcriptome) analysis can be employed (Zubakov et al. 2008; Park et al. 2013a, b; Hanson and Ballantyne 2013a, b). Candidate markers dependent on the specific role of a body fluid, such as hemoglobin, are often more prone to be precisely expressed.

Furthermost, commonly suggested blood RNA markers may typically classified proteins involved including the erythrocyte membrane (such as glycophorin A, beta-spectrin, and ankyrin 1) and the proteins allied with hgb and hemic biosynthesis (such as α and β hemoglobins, porphobilinogen deaminase and (ALAS 2) amino-levulinate synthase 2). In the afore-mentioned immunological tests, some proteins

(e.g., Glycophorin A) are being used. The tissues, for instance menstrual blood and saliva, can be classified as being nonspecific due to trace levels of blood markers in them (Kohlmeier and Schneider 2012; Zubakov et al. 2008; Fleming and Harbison 2010a, b; Lindenbergh et al. 2012; Juusola and Ballantyne 2005; Haas et al. 2009; Richard et al. 2012; Roeder and Haas 2013; Jakubowska et al. 2013; Donfack and Wiley 2015; Park et al. 2013a, b).

Protamines—PRM1 and PRM2—are most frequently used for sperm identification, and transglutaminases 4 and semenogelines 1 and 2 for seminal fluids include seminal fluid and spermatozoa markers. Semenogelin is a substratum used by some for PSA/P30/kallikrein 3. Specific markers also show little, or no, cross-reactivity (Kohlmeier and Schneider 2012; Fleming and Harbison 2010a, b; Lindenbergh et al. 2012; Juusola and Ballantyne 2005; Haas et al. 2009; Richard et al. 2012; Roeder and Haas 2013; Nussbaumer et al. 2006; Donfack and Wiley 2015; Park et al. 2013a, b). Among the four peptide groups, statherin and histatin are preferred as Ribo Nucleic Acid markers for saliva, divided by the salivary glands into saliva. A significant proportion of the overall salivary peptides contain three proline-rich proteins. Some keratins were also used (Zubakov et al. 2008; Fleming and Harbison 2010a, b; Lindenbergh et al. 2012; Juusola and Ballantyne 2005, 2007; Haas et al. 2009; Richard et al. 2012; Roeder and Haas 2013; Donfack and Wiley 2015; Park et al. 2013a, b; Sakurada et al. 2011). In factors such as the daily routine and personal circumstances of the person are affected by both the amount and content of such peptides and the complete length and partly damaged mRNA transcripts observed in saliva and many authors also reported that there can be a wide variability, along with small amounts of RNA in samples of saliva (Fleming and Harbison 2010a, b; Haas et al. 2009). Statherin allows to be observed at high levels in nasal secretions and at low levels in vaginal secretions often.

A specific problem is the differentiation of stratified squamous cells in the skin, vagina, and mouth, since they are strongly connected structurally. Because of the similar functionality (protection and secretion) of these cell types, it is difficult to find observable distinctions between them, especially in non-keratinized oral and vaginal cells.

The fluid in which the menstrual fluid curves, cervix, endometrium, and fallopian tubes, and blood differ according to age and health conditions comprises the vaginal fluid. Early vaginal marker candidates, human beta-defense 1 and mucine 4, are indeed reliably represented in vaginal secretions, they were also found to be in nasal secretions, semen, and saliva. These can be used as mucosal markers better (Sijen 2015; Haas et al. 2014; Lindenbergh et al. 2012; Juusola and Ballantyne 2005; Richard et al. 2012; Roeder and Haas 2013; Donfack and Wiley 2015; Hanson and Ballantyne 2013a, b; Cossu et al. 2009). Recently, two other candidates' genes MYOZ1 and CYP2B7P1, equally appearing responsive markers were found using transcriptome profiling. CYP2B7P1 has not been identified with any cross-reactivity, although a skeletal muscle protein, MYOZ1 localized in the tongue, was noticed in saliva (Hanson et al. 2012). Most authors have been used to classify vaginal material by microflora, which is defined in the following document (Giampaoli et al. 2014).

Menstrual blood is often a diverse fluid comprised of different quantities of flowing blood, discharges from vagina, infectious species, and menstrual cycle cells. The refurbishments of the endometrium while menstruation offer ideal identification candidate markers. MMP7 and MMP11 are the most significant groups studied (Haas et al. 2014; Fleming and Harbison 2010a, b; Lindenbergh et al. 2012, 2013a, b; Juusola and Ballantyne 2005, 2007; Richard et al. 2012; Roeder and Haas 2013; Jakubowska et al. 2013). The transcription of MMP genes during the entire menstrual cycle was found to differ.

Distinctions concerning the epithelial (vaginal and mouth) and the epidermal cells are established in the detection of skin cells on a sample from which a DNA profile can be derived. For the epidermal cell recognition, a variety of cytokeratin families are suggested. Among these LCE1C and LOR showed that skin samples containing weak and inefficient identification of the other markers and housekeeping genes have the most stable detectable abundance of transcription, which is possibly representing the very low level of mRNA in the cell (Hanson et al. 2012; Schulz et al. 2010). Many of these markers appeared significantly in vaginal secretions (Hanson et al. 2012; Simons and Vintiner 2012). Regardless of the methodology used, laboratory confirmation documenting standards are needed. The suggested PCR end point reporting approaches differ. The following are weighted rating schemes focused on presence (peakingapex), dearth and sensitivity of several markers; consensus amplification; multifacet scale strategies and the use of controls. The related testing methods are used for quantitative PCR methods, yet they have trouble with mixed sample interpretation. The analysis of DNA is becoming widespread in the forensic sciences analysis and the Massive Parallel Sequencing or Next Generation Sequencing (MPS) technique has only been recently extended to Ribo Nucleic Acid (RNA) analysis. Assay of messenger RNAs in bodily fluids utilizing MPS/NGS for menstrual blood, saliva, and vaginal secretions and the simultaneous sequence of Deoxyribo Nucleic Acid and RNA out of the identical sample (Lin et al. 2015; Zubakov et al. 2015).

5.13 NanoString®

NanoString® nCotunter is a tool for recording and counting of the various mRNA transcripts, which can use color-coded molecular bar codes to measure the expression of up to 800 mRNA applicants with one reaction (Malkov et al. 2009; Brumbaugh et al. 2011). In recent times, NanoString® used as a tool for distinguishing body fluid with the utilization of messenger RNA in criminal research. In the first test, 18 mRNAs unique to body fluid and two endogenous controls were used. Throughout the study, a total RNA was used and the counts of housekeeping gene GAPDH were normalized. The use of body fluid-based markers for blood and semen were correctly identified, but the vaginal marker and the mRNA for saliva were not reliable (Park et al. 2013a, b). In a study a wide variety of different specimens, including total RRNA extracted from body fluids, were analysed using direct cell lysates with 23 messenger RNA markers and ten

housekeeping genes (Danaher et al. 2015). Samples out of a solo source of semen, blood, menstrual blood, skin, and vaginal secretions, every part of which showed gene expression specific to body fluid using an algorithm and maximum likelihood calculations. Once again, the use of this technique made saliva samples troublesome.

5.14 miRNA Profiling Approaches

There are a wide range of different messenger RNA markers have been projected being sensitive and specific in support of body fluids detection in criminal cases and have also proven their stability in the different stored samples over longer period of time (Zubakov et al. 2009; Setzer et al. 2008). The recently introduced micro RNA (miRNA) markers have explored to be alternate tool for body fluids detection in criminal cases (Zubakov et al. 2010). Non-coding RNA molecules are miRNAs that are of 18–25 nucleotides length which regulates expression of gene at the post transcription levels and is also reported that several miRNAs have tissue-specific expression patterns (Bartel 2004; Jost et al. 2011). Additionally, the very small strands of miRNA markers make them very less susceptible to dreadful conditions by environmental factors, which offer advantage biomarkers for body fluids identification than mRNA markers (Hanson et al. 2009).

The foremost microRNA profiling to criminalistics was reported by Hanson et al. in 2009, and assessed the miRNA expression in biological fluids relevant to forensic approach. Using qRT-PCR analysis he explored the expression of 452 human micro RNAs in saliva, blood, semen, menstrual blood and vaginal fluids. They have characterized nine microRNA markers viz. miR451, miR135b, miR16, miR10b, miR205, miR124a, miR658, miR412, and miR372—these markers are expressed distinctively in genetic experiments related to forensic cases, and have also established their competences to detect body fluids utilizing as minute as 50 pg of whole RNA. Zubakov et al. in 2010 have also reported the expression degrees of 718 microRNAs in biological fluids addressed to forensic utilizing genome-based microarrays and recognized 14 distinctive markers for possible body fluid identification. In this authentication test using qRT-PCR, merely blood and semen exclusive microRNA candidates indicated similar expression degrees, whereas fewer similarities were identified intended for vaginal secretions, menstrual blood, and saliva. Another study which was designed similarly but including skin cells was also reported. In this study they identified 14 new markers which revealed suitable body fluid exclusive expression arrays and also reported that microRNA-4761-5p can distinct skin cells and menstrual blood; microRNA-137 can make a distinction between cells of skin, peripheral blood and saliva; microRNA-4473 can individualize semen and skin cells; microRNA-585-3p can separate vaginal secretion from skin cells (Luo et al. 2015).

5.15 DNA Methylation Profiling Methods

The discovery of entire genome studies specified that the genetic material called DNA relays tissue-distinctive methylation arrays, the possibilities of tissue-specificity DNA methylation has been intended for examination of different body fluids identification in forensic fields (Frumkin et al. 2011; Lee et al. 2012). DNA methylation patterns took place at the CpG dinucleotides-5′ position of pyrimidine ring of cytosine, is largely supposed to express the genes of particular body fluids (Miranda and Jones 2007). The DNA methylation arrays observed will be depends upon the different types of cells and chromosome portions called tissue-distinctive Differentially Methylated Regions (tDMRS) are identified to spectacle unique DNA methylation profiles. Accordingly, the recognition of DNA methylation at a CpG location of Differentially Methylated Regions (tDMR) allows for the tissue or else cell-specific kind of samples of DNA (Byun et al. 2009; Ohgane et al. 2008).

Around 2011, Frumkin et al. described 15 genomic loci to identify saliva, semen, blood, skin, epidermis, vaginal secretions and urine. Within this research they applied methylation sensitive restriction enzyme PCR (MSRE-PCR) designed using an assay for certain markers which are composed up of methylation sensitive restriction enzyme. These types of restriction enzymes digest the samples of DNA subsequently multiplex PCR reactions of specified loci of genomic DNA with fluorescently marked primers, followed by capillary electrophoresis of amplified products using programmed signal detection method. These particularly designed assays can be easily incorporated in forensic science laboratories like other techniques of DNA typing viz. SniPS (SNPs), microsatellites or Short Tandem Repeats (STRs) which would positively identify the source tissues. Similarly, Wasserstrom et al. (2013) reported with the advance methodology by build-up a kit, DNA source identifier (DSI)-Semen™, which was designed to substitute the microscopic investigation of sperm cells for identification of semen in forensic DNA samples. This kit was designed on the basis to detect the semen-distinctive DNA methylation arrays at five genomic loci utilizing MSRE-PCR. The source kit was ratified with 135 samples of different body fluids and 33 case evidences from forensic laboratory, which was verified to be vigorous and consistent by presenting a positive outcome for samples of semen as minute as 31 pg of DNA template (Madi et al. 2012). In another research carried out by Lee et al. studied the tissue-distinctive Differentially Methylated Regions possibility for forensic body fluids detection by bisulfite sequencing approach, which regulates the methylation of DNA via detecting the nucleotide base alteration as a result of treatment of sodium bisulfite. Treatment Bisulfite does not affect any methylated cytosine instead alters free or else unmethylated cytosine of CpG into uracil, which converts into thymine during the PCR reaction. By means of this methodology, they generated five tDRMs profiles in combined DNA samples from saliva, blood, menstrual blood, semen, and vaginal fluids. The tDRMs for USP49 and DACT1 existed recognized as an exclusive marker for semen—by revealing semen-distinctive hypomethylation and PFN3 tDMR was preferred to characterize the vaginal fluids. In a recent approach proposed by Madi et al. illustrated the tissue-restricted DNA methylation in biological body

fluids which are forensically relevant including saliva, blood, epithelial cells, and semen. They identified a limited genomic locus with bisulfite modified method and pyrosequencing technique to observe that the methylation arrays at the FGF7 and ZC3H12D loci can distinguish the semen sample from more genetic samples although locus—C20 and BCAs4 can distinguish the saliva and blood samples as of other biological fluids separately. The outcomes as well specify the DNA methylation-based methods are invaluable tool for the detection of biological fluids in criminal cases.

5.16 DNA-Compatible Cell-Specific Identification

Another method for identifying biological fluid ensues classification of specific cells by marking the cells in combination with a microscopy at a protein, DNA, and RNA level. In the case of immunohistochemistry, epidermal cells have been identified and cytokeratin's are used to differentiate between vaginal and mouth mucosal cells (Schulz et al. 2010; Paterson et al. 2006).

As compared directly to epidermal cells, cells of mucosal origin may be distinguished since each cytokeratin has a low expression of the other type of cell. In a different method, immunofluorescence has been suggested for the on-site recognition of human blood with fluorescently tagged antihuman antibodies, nuclear (CD45, myeloperoxidase, histone 1), and (glycophorin A) while several washing steps may find this method unworkable for forensic work (Thorogate et al. 2008). XY Fluorescent In Situ Hybridisation (FISH) combined with Laser Microdissection (LMD) and DNA profiling has been shown to be an efficient way of extracting the data from specific cell groups marked with fluorescent labels, i.e., chromosomes X and Y (Lynch et al. 2015).

ESR1 (Estrogen Receptor 1) with fluorescently labeled monoclonal antibodies has been shown to be able to differentiate between vaginal and buccal epithelial cells with this marker if messenger RNA profiling is out of the question. This displays the expression of messenger RNA and tissue-specific protein are not always interconnected (Fleming et al. 2013).

RNA suspension FISH are used to classify and detect epithelial cells using keratin 10 with the fluorescent label LNA sample; subsequently, the epithelial cells were isolated by LMD and profiled with DNA (Williams et al. 2014). While these labeling methods that tend to be accurate, the longer time taken to specifically gather labeled cells by LMD illustrates, instead for general applications, these methods are presumably restricted to particular cases.

5.17 Conclusion

Many forensic laboratories are still in practice with the chemical methods, antigen–antibody reaction, catalytic reactions, while the current DNA profiling methods are based on STR profiling which allows personal identity of sample donor. A recent

advance in forensics has advocated the usage of other types of evidences which add more information to the collected evidences. Particularly in recent times several markers have existed for body fluids detection in forensic community. mRNA markers usage in recent times have been utmost meticulously explored with the number of different markers which is sufficient for the body fluids identification in forensic perspective. The currently used methods have possible approaches for the quick detection of biological fluids, but then again in majority of the cases tissue-specific identification is difficult due to cross overreaction with other biological body fluids present at the crime scene. In this scenario mRNA marker appears as high sensitivity and tissue specificity for forensic analysis with a new method to over-power the boundaries of established methods. Similarly, micro RNA methods have explored to be alternative tool for body fluids identification in forensics. microRNA profiling has the capabilities to identify or to differentiate the different body fluids traced at the crime scene. In recent times, profiling through DNA methylation was recommended as a novel means in criminal cases involving body fluids detection. Like messenger and micro RNA profiling, DNA methylation profiling as well appeared highly sensitive and specific, for the parallel assessment of several markers particular for different tissues in a definite multiplex reaction. In the coming future, forensic examinations have to progress with a huge deal as the encroachments in molecular biology studies and forensic genetics.

References

Akutsu T, Watanabe K, Sakurada K (2012a) Specificity, sensitivity, and operability of RSID™--urine for forensic identification of urine: comparison with ELISA for Tamm-Horsfall protein. J Forensic Sci 57:1570–1573

Akutsu T, Motani H, Watanabe K, Iwase H, Sakurada K (2012b) Detection of bacterial 16S ribosomal RNA genes for forensic identification of vaginal fluid. Legal Med 14:160–162

Alvarez M, Juusola J, Ballantyne J (2004) An mRNA and DNA co-isolation method for forensic casework samples. Anal Biochem 335:289–298

Anderson SE, Hobbs GR, Bishop CP (2010) Multivariate analysis for estimating the age of a bloodstain. J Forensic Sci 56:186–193

Auvdel ML (1986) Amylase levels in semen and saliva stains. J Forensic Sci 31:426–431

Baker DJ, Grimes EA, Hopwood AJ (2011) D-dimer assays for the identification of menstrual blood. Forensic Sci Int 212:210–214

Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116:281–297

Bauer M, Kraus A, Patzelt D (1999) Detection of epithelial cells in dried bloodstains by reverse transcriptase polymerase chain reaction. J Forensic Sci 44:1232–1236

Bauer M, Gramlich I, Polzin S, Patzelt D (2003) Quantification of mRNA degradation as possible indicator of postmortem interval – a pilot study. Legal Med 5:220–227

Boward ES, Wilson SL (2013) A comparison of ABAcard® p30 and RSID™-semen test kits for forensic semen identification. J Forensic Legal Med 20:1126–1130

Brumbaugh CD, Kim HJ, Giovacchini M, Pourmand N (2011) NanoStriDE: normalization and differential expression analysis of NanoString nCounter data. BMC Bioinformatics 12:479

Bustin SA (2000) Absolute quantification of mRNA using real time reverse transcription polymer-ase chain reaction assays. J Mol Endocrinol 25:169–193

Bustin SA, Benes V, Nolan T, Pfaffl MW (2005) Quantitative real time RT PCR – a perspective. J Mol Endocrinol 34:597–601

Byun HM, Siegmund KD, Pan F, Weisenberger DJ, Kanel G, Laird PW, Yang AS (2009) Epigenetic profiling of somatic tissues from human autopsy specimens identifies tissue- and individual-specific DNA methylation patterns. Hum Mol Genet 18:4808–4817
Casey DG, Price J (2010) The sensitivity and specificity of the RSID™ saliva kit for the detection of human salivary amylase in the forensic science laboratory, Dublin, Ireland. Forensic Sci Int 194:67–71
Cecchi R (2010) Estimating wound age: looking into the future. Int J Legal Med 124:523–536
Cossu C, Germann U, Kratzer A, Baer W, Haas C (2009) How specific are the vaginal secretion mRNA markers HBD1 and MUC4. Forensic Sci Int Genet Suppl Ser 2:536–537
Cox M (1991) A study of the sensitivity and specificity of four presumptive tests for blood. J Forensic Sci 36:1503–1511
Danaher P, White RL, Hanson EK, Ballantyne J (2015) Facile semi-automated forensic body fluid identification by multiplex solution hybridization of NanoString® barcode probes to specific mRNA targets. Forensic Sci Int Genet 14:18–30
Donfack J, Wiley A (2015) Mass spectrometry based cDNA profiling as a potential tool for body fluid identification. Forensic Sci Int Genet 16:112–120
Fleming RI, Harbison SA (2010a) The development of an mRNA multiplex RT-PCR assay for the definitive identification of body fluids. Forensic Sci Int Genet 4:244–256
Fleming RI, Harbison SA (2010b) The use of bacteria for the identification of vaginal secretions. Forensic Sci Int Genet 4:311–315
Fleming RI, Harbison SA, Lin MH (2013) New RNA methods for the identification of body fluids and cell types. Forensic Sci Int Genet Suppl Ser 4:e87–e88
Fox A, Gittos M, Harbison SA, Fleming R, Wivell R (2014) Exploring the recovery and detection of messenger RNA and DNA from enhanced fingermarks in blood. Sci Justice 54:192–198
Frumkin D, Wasserstrom A, Budowle B, Davidson A (2011) DNA methylation-based forensic tissue identification. Forensic Sci Int Genet 5:517–524
Gaensslen RE (1983) Sourcebook in forensic serology, immunology and biochemistry. US Department of Justice, Washington, DC
Giampaoli S, Alessandrini F, Berti A et al (2014) Forensic interlaboratory exercise of the For FLUID kit for vaginal fluids identification. J Forensic Legal Med 21:60–63
Grabmuller M, Madea B, Courts C (2015) Comparative evaluation of different extraction and quantification methods for forensic RNA analysis. Forensic Sci Int Genet 16:195–202
Graves HCB (1995) Non prostatic sources of protein-specific antigen: a steroid hormone dependent phenomenon? Clin Chem 41:7–9
Gray D, Frascione N, Daniel B (2012) Development of an immunoassay for the differentiation of menstrual blood from peripheral blood. Forensic Sci Int 220:12–18
Haas C, Klesser B, Maake C, Bar W, Kratzer A (2009) mRNA profiling for body fluid identification by reverse transcription endpoint PCR and real time PCR. Forensic Sci Int Genet 3:80–88
Haas C, Hanson E, Kratzer A, Bar W, Ballantyne J (2011) Selection of highly specific and sensitive mRNA biomarkers for the identification of blood. Forensic Sci Int Genet 5:449–458
Haas C, Hanson E, Anjos MJ et al (2014) RNA/DNA co-analysis from human menstrual blood and vaginal secretion stains: results of a fourth and fifth collaborative EDNAP exercise. Forensic Sci Int Genet 8:203–212
Haas C, Hanson E, Banemann R et al (2015) RNA/DNA co-analysis from human skin and contact traces – results of a sixth collaborative EDNAP exercise. Forensic Sci Int Genet 16:139–147
Hanson EK, Ballantyne J (2013a) Multiplex high resolution melt (HRM) messenger RNA profiling analyses for body fluid identification. Forensic Sci Int Genet Suppl Ser 4:e125–e126
Hanson EK, Ballantyne J (2013b) Highly specific mRNA biomarkers for the identification of vaginal secretions in sexual assault investigations. Sci Justice 53:14–22
Hanson EK, Lubenow H, Ballantyne J (2009) Identification of forensically relevant body fluids using a panel of differentially expressed microRNAs. Anal Biochem 387:303–314
Hanson E, Haas C, Jucker R, Ballantyne J (2012) Specific and sensitive mRNA biomarkers for the identification of skin in “tough” DNA evidence. Forensic Sci Int Genet 6:548–558

Harteveld J, Lindenbergh A, Sijen T (2013) RNA cell typing and DNA profiling of mixed samples: can cell types and donors be associated? Sci Justice 53:261–269
Holtkoetter H, Dierig L, Schuerenkamp M, Sibbing U, Pfeiffer H, Vennemann M (2015) Validation of an immunochromatographic D-dimer test to presumptively identify menstrual fluid in forensic exhibits. Int J Legal Med 129:37–41
Huang CT, Chen ML, Huang LL, Mao IF (2002) Uric acid and urea in human sweat. Chin J Phys 45:109–115
Jakubowska J, Maciejewska A, Pawlowski R, Bielawski KP (2013) mRNA profiling for vaginal fluid and menstrual blood identification. Forensic Sci Int Genet 7:272–278
Johnston S, Newman J, Frappier R (2003) Validation study of the Abacus Diagnostics ABAcard® Hematrace® membrane test for the forensic identification of human blood. Can Soc Forensic Sci J 36:173–183
Jones EL Jr, Leon JA (2004) Lugol's test re-examined again: buccal cells. J Forensic Sci 49:64–67
Jost D, Nowojewski A, Levine E (2011) Small RNA biology is systems biology. BMB Rep 44:11–21
Juusola J, Ballantyne J (2005) Multiplex mRNA profiling for the identification of body fluids. Forensic Sci Int 152:1–12
Juusola J, Ballantyne J (2007) mRNA profiling for body fluid identification by multiplex quantitative RT-PCR. J Forensic Sci 52:1252–1262
Keating SM, Higgs DF (1994) The detection of amylase on swabs from sexual assault cases. J Forensic Sci Soc 34:89–93
Kind SS (1957) The use of the acid phosphatase test in searching for seminal stains. J Crim Law Criminol Police Sci 47:597–600
Kipps AE, Whitehead PN (1975) The significance of amylase in forensic investigations of body fluids. Forensic Sci 6:137–144
Kobus HJ, Silenieks E, Scharnberg J (2002) Improving the effectiveness of fluorescence for the detection of seminal stains on fabrics. J Forensic Sci 47:819–823
Kohlmeier F, Schneider PM (2012) Successful RNA profiling of 23 year old blood stains. Forensic Sci Int Genet 6:274–276
Lee HY, Park MJ, Choi A, An JH, Yang WI, Shin KJ (2012) Potential forensic application of DNA methylation profiling to body fluid identification. Int J Legal Med 126:55–62
Legg KM, Powell R, Reisdorph N, Reisdorph R, Danielson PB (2014) Discovery of highly specific protein markers for the identification of biological stains. Electrophoresis 35:3069–3078
Liang T, Roy R (2014) Ultraviolet-visible spectrophotometry (UV-VIS) and SALIgAE® qualitative and semi-quantitative tools for the analysis of salivary amylase. J Forensic Res 5:247
Lin MH, Jones DF, Fleming R (2015) Transcriptomic analysis of degraded forensic body fluids. Forensic Sci Int Genet 17:35–42
Lindenbergh A, De Pagter M, Ramdayal G et al (2012) A multiplex (m)RNA profiling system for the forensic identification of body fluids and contact traces. Forensic Sci Int Genet 6:565–577
Lindenbergh A, Van den Berge M, Oostra RJ et al (2013a) Development of a mRNA profiling multiplex from the inference of organ tissues. Int J Legal Med 127:891–900
Lindenbergh A, Maaskant P, Sijen T (2013b) Implementation of RNA profiling in forensic casework. Forensic Sci Int Genet 7:159–166
Lunetta P, Sippel H (2009) Positive prostate-specific antigen (PSA) reaction in post mortem rectal swabs: a cautionary note. J Forensic Legal Med 16:397–399
Luo XY, Li ZL, Peng D, Wang L, Zhang L, Liang WB (2015) MicroRNA markers for forensic body fluid identification obtained from miRCURYTM LNA array. Forensic Sci Int Genet Suppl Ser 5:e630–e632
Lynch L, Gamblin A, Vintiner S, Simons J (2015) STR profiling of epithelial cells identified by X/Y FISH labelling and laser microdissection using standard and elevated PCR conditions. Forensic Sci Int Genet 16:1–7

Madi T, Balamurugan K, Bombardi R, Duncan G, McCord B (2012) The determination of tissue-specific DNA methylation patterns in forensic biofluids using bisulfite modification and pyrosequencing. Electrophoresis 33:1736–1745

Malkov VA, Serikawa KA, Balantac N et al (2009) Multiplexed measurements of gene signatures in different analytes using the Nanostring nCounter™ assay system. BMC Res Notes 2:80

Miranda TB, Jones PA (2007) DNA methylation: the nuts and bolts of repression. J Cell Physiol 213:384–390

Misencik A, Laux DL (2007) Validation study of the SeratecHemDirect hemoglobin assay for the forensic identification of human blood. MAFS Newsl 36:18–26

Moreno LI, Tate CM, Knott EL et al (2012) Determination of an effective housekeeping gene for the quantification of mRNA for forensic applications. J Forensic Sci 57:1051–1058

Nussbaumer C, Gharehbaghi-Schnell E, Korschineck I (2006) Messenger RNA profiling: a novel method for body fluid identification by real time PCR. Forensic Sci Int 157:181–186

Ohgane J, Yagi S, Shiota K (2008) Epigenetics: the DNA methylation profile of tissue dependent and differentially methylated regions in cells. Placenta 29:S29–S35

Old JB, Schweers BA, Boonlayangoor PW, Reich KA (2009) Developmental validation of RSID™-saliva: a lateral flow immunochromatographic strip test for the forensic detection of saliva. J Forensic Sci 54:866–873

Old JB, Schweers BA, Boonlayangoor PW, Reich KA (2012) Developmental validation of RSID-semen: a lateral flow immunochromatography strip test for the forensic identification of semen. J Forensic Sci 57:489–499

Palagummi S, Harbison SA, Fleming RI (2014) A time-course analysis of mRNA expression during injury healing in human dermal injuries. Int J Legal Med 128:403–414

Pang BCM, Cheung BKK (2007) Identification of human semenogelin in membrane strip test as an alternative method for the detection of semen. Forensic Sci Int 169:27–31

Pang BCM, Cheung BKK (2008) Applicability of two commercially available kits for forensic identification of saliva stains. J Forensic Sci 53:1117–1122

Park NJ, Zhou X, Yu T et al (2007) Characterization of salivary RNA by cDNA library analysis. Arch Oral Biol 52:30–35

Park SM, Park SY, Kim JH et al (2013a) Genome-wide mRNA profiling and multiplex quantitative RT-PCR for forensic body fluid identification. Forensic Sci Int Genet 7:143–150

Park JL, Park SM, Kim JH et al (2013b) Forensic body fluid identification by analysis of multiple RNA markers using NanoString technology. Genomics Inform 11:277–281

Paterson SK, Jenson CG, Vintiner SK, McGlashan SR (2006) Immunohistochemical staining as a potential method for the identification of vaginal epithelial cells in forensic casework. J Forensic Sci 51:1138–1143

Quinones I, Sheppard D, Harbison SA, Elliot DA (2006) Comparative analysis of luminol formulations. Can Soc Forensic Sci 40:53–63

Richard ML, Harper KA, Craig RL, Onorato AJ, Robertson JM, Donfack J (2012) Evaluation of mRNA marker specificity for the identification of five human body fluids by capillary electrophoresis. Forensic Sci Int Genet 6:452–460

Roeder AD, Haas C (2013) mRNA profiling using a minimum of five mRNA markers per body fluid and a novel scoring method for body fluid identification. Int J Legal Med 27:707–721

Sagawa K, Kimura A, Saito Y et al (2003) Production and characterisation of a monoclonal antibody for sweat-specific protein and its application for sweat identification. Int J Legal Med 117:90–95

Sakurada K, Akutsu T, Fukushima H, Watanabe K, Yoshino M (2010) Detection of dermicidin for sweat identification by real time RT PCR and ELISA. Forensic Sci Int 194:80–84

Sakurada K, Akutsu T, Watanabe K, Fujinami Y, Yoshino M (2011) Expression of statherin mRNA and proteins in nasal and vaginal secretions. Legal Med 13:309–313

Sato I, Kojima K, Yamasaki T et al (2004) Rapid detection of semenogelin by one step immunochromatographic assay for semen identification. J Immunol Methods 287:137–145

Schulz MM, Buschner MGD, Leidig R et al (2010) A new approach to the investigation of sexual offenses-cytoskeleton analysis reveals the origin of cells found on forensic swabs. J Forensic Sci 55:492–498

Schweers BA, Old J, Boonlayangoor PW, Reich KA (2008) Developmental validation of a novel lateral flow strip test for rapid identification of human blood (Rapid Stain Identification™--blood). Forensic Sci Int Genet 2:243–247

Setzer M, Juusola J, Ballantyne J (2008) Recovery and stability of RNA in vaginal swabs and blood, semen and saliva stains. J Forensic Sci 53:296–305

Simich JP, Morris SL, Klick RL et al (1999) Validation of the use of a commercially available kit for the identification of prostate specific antigen (PSA) in semen stains. J Forensic Sci 44:1229–1231

Simons JL, Vintiner SK (2012) Efficacy of several candidate protein biomarkers in the differentiation of vaginal from buccal epithelial cells. J Forensic Sci 57:1585–1590

Stombaugh PM, Kearney JJ (1987) Factors affecting the use of lactate dehydrogenase as a means of bloodstain differentiation. J Forensic Sci 23:94–105

Su CW, Li CY, Lee JCI et al (2015) A novel application of real-time RT-LAMP for body fluid identification: using HBB detection as the model. Forensic Sci Med Pathol 11(2):208–215

Thorogate R, Moreira JCS, Jickells S, Miele MMP, Daniel B (2008) A novel fluorescence-based method in forensic science for the detection of blood in situ. Forensic Sci Int Genet 2:363–371

Tobe SS, Watson N, Daeid NN (2007) Evaluation of six presumptive tests for blood, their specificity, sensitivity, and effect on high molecular-weight DNA. J Forensic Sci 52:102–109

Van den Berge M, Carracedo A, Gomes I et al (2014) A collaborative European exercise on mRNA-based body fluid/skin typing and interpretation of DNA and RNA results. Forensic Sci Int Genet 10:40–48

Vandenberg N, van Oorshot RAH (2006) The use of Polilight in the detection of seminal fluid, saliva and bloodstains and comparison with conventional chemical based screening tests. J Forensic Sci 51:361–370

Virkler K, Lednev IK (2009) Analysis of body fluids for forensic purposes: from laboratory testing to non-destructive rapid confirmatory identification at a crime scene. Forensic Sci Int 188:1–17

Visser M, Zubakov D, Ballantyne KA, Kayser M (2011) mRNA based skin identification for forensic applications. Int J Legal Med 125:253–263

Wasserstrom A, Frumkin D, Davidson A, Shpitzen M, Herman Y, Gafny R (2013) Demonstration of DSI-semen—a novel DNA methylation-based forensic semen identification assay. Forensic Sci Int Genet 7:136–142

Webb JL, Creamer TI, Quickenden TI (2006) A comparison of the presumptive luminol test for blood with four non-chemiluminescent forensic techniques. J Forensic Sci 51:361–370

Williams E, Lin MH, Harbison SA, Fleming RI (2014) The development of a method of suspension RNA-FISH for forensically relevant epithelial cells using LNA probes. Forensic Sci Int Genet 9:85–92

Yu H, Diamandis EP (1995) Prostate specific antigen in milk of lactating women. Clin Chem 41:54–58

Zubakov D, Hanekamp E, Kokshoorn M, Van Ijcken W, Kayser M (2008) Stable mRNA markers for identification of blood and saliva stains revealed from whole genome expression analysis of time-wise degraded samples. Int J Legal Med 122:135–142

Zubakov D, Kokshoorn M, Kloosterman A, Kayser M (2009) New markers for old stains: stable mRNA markers for blood and saliva identification from up to 16-year-old stains. Int J Legal Med 123:71–74

Zubakov D, Boersma AW, Choi Y, van Kuijk PF, Wiemer EA, Kayser M (2010) MicroRNA markers for forensic body fluid identification obtained from microarray screening and quantitative RT-PCR confirmation. Int J Legal Med 124:217–226

Zubakov D, Kokmeijer I, Ralf A et al (2015) Towards simultaneous individual and tissue identification: a proof-of-principle study on parallel sequencing of STRs, amelogenin, and mRNAs with the Ion Torrent PGM. Forensic Sci Int Genet 17:122–128

6 Statistical Softwares Used in Evaluation of Forensic DNA Typing

R. K. Kumawat, Aditi Mishra, and Pankaj Shrivastava

Abstract

Statistical analysis and evaluation of genetic data of an individual and interpretation of obtained results, especially in mixed DNA profiles obtained in sexual assault cases is ideally required to be strengthened by statistical evaluation, and for handling this use of many software is recommended. The present chapter is intended to provide some of the software used globally for mixed profile interpretation. The chapter also includes information about softwares being used for population data analysis which is further a very important area in forensic DNA analysis which can be evidenced with the number of publications on population-based DNA analysis in almost all the leading forensic journals.

Keywords

Forensic · Interpretation · Statistical evaluation

R. K. Kumawat (✉)
DNA Division, State Forensic Science Laboratory, Rajasthan, Jaipur, India

A. Mishra
Raksha Shakti University, Ahmedabad, India

DNA Fingerprinting Unit, State Forensic Science Laboratory, Department of Home (Police), Government of Madhya Pradesh, Sagar, Madhya Pradesh, India

P. Shrivastava
DNA Fingerprinting Unit, State Forensic Science Laboratory, Department of Home (Police), Government of Madhya Pradesh, Sagar, Madhya Pradesh, India

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_6

6.1 Introduction

6.1.1 The Software Used for Population Data Analysis and Genetic Study

In forensic DNA typing, DNA profiles obtained from crime-related exhibits are used to link or identify the accused, involve in the crime. As we know, that we do not have DNA profiles of all the individuals living on the earth. So it is very difficult to conclude that the DNA profiles obtained from crime-related exhibits are from the same suspect individuals or is there someone another who might be actual criminal. Might that suspect happen to match the DNA profiles with the crime-related evidence? It is answered with the use of population data statistics, and it is necessary to conclude from DNA testing about the matching or not matching of DNA profile with crime evidence after calculating the possibility of a random match (Butler 2001). Population statistics is a powerful tool that deals in the measurement of the uncertainty and provides the probability of random match. The population statistics for the random match is usually estimated using the frequency of a particular genotype (DNA) profile in a population. Frequency data of a population might be prepared using DNA profiles of a small sample size of randomly selected individuals residing in a particular geographic region. However, there are several software which are being used in statistical evaluation of genetic data speedily. Recent and most useful genetic data analysis software are presented as following in two categories.

1. Software being used for population genetic analysis
2. Software being used for interpretation of forensic genetic evidence

6.2 Software Being Used for Population Genetic Analysis

PowerStats Spreadsheet Program: PowerStats spreadsheet program (Tereba 1999) calculates forensic interest parameters, namely, Allele frequency (AF), Discrimination power, Paternity index, Exclusion power, Polymorphic information content, Probability of match, Homozygous and Heterozygous in percent and the total number of alleles (Fig. 6.1).

Note: This is a very user-friendly software and is based on Excel. The software was earlier made available from Promega and no longer available online now. But still the excel worksheet of this programme is available with scientists working on population-based DNA analysis and is still being used and cited.

6.2.1 GenAlEx Software

GenAlEx: Genetic Analysis in Excel software (Peakall and Smouse 2006) is useful for the analysis of co-dominant, haploid and binary genetic loci and DNA sequences. This software performs analysis based on frequency and distance.

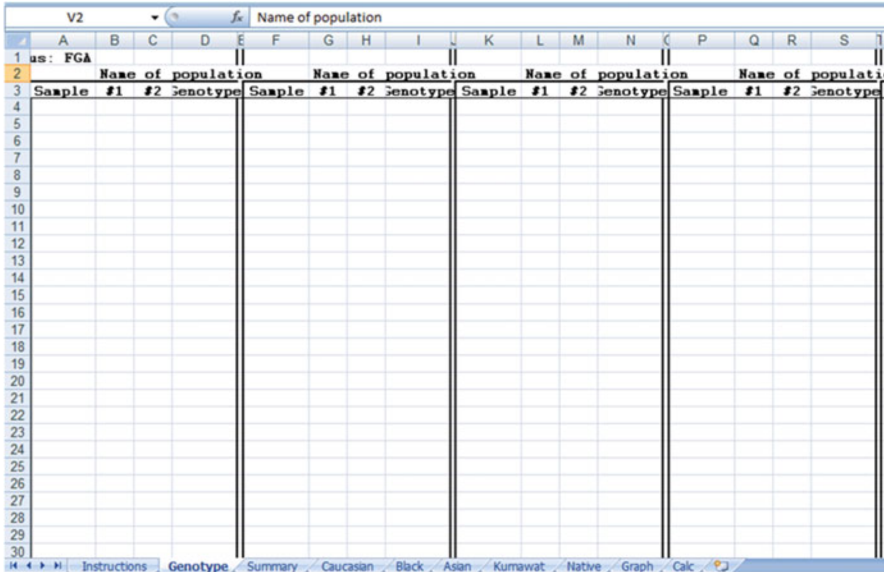


Fig. 6.1 Image of home page of PowerStats v1.2 spreadsheet program

***Frequency-based analysis*:** Allele frequency, Observed and expected heterozygosity, Hardy-Weinberg equilibrium, Relatedness, Fixation index, etc.

Distance-based analysis: Mantel test, Multivariate spatial autocorrelation, Analysis of molecular variance and principal coordinate analysis, etc. (Fig. 6.2).

6.2.2 Arlequin Software

Arlequin software (Excoffier and Lischer 2010) is a useful tool for the analysis of broad spectrum of genetic data i.e., STRs, RFLP, standard and frequency based data to analyses inter and intrapopulation statistics.

- In the intrapopulation statistics, various parameters such as gene diversity, polymorphic sites, pair wise difference between individuals, estimation of haplotype frequency, minimum spanning network among the individuals estimation of randomization, etc.
- In the interpopulation, comparison in haplotypic content of population, structure of population, *Fst*-based genetic distances, estimation of isolation by distance using Mantel test can be analyzed (Fig. 6.3).

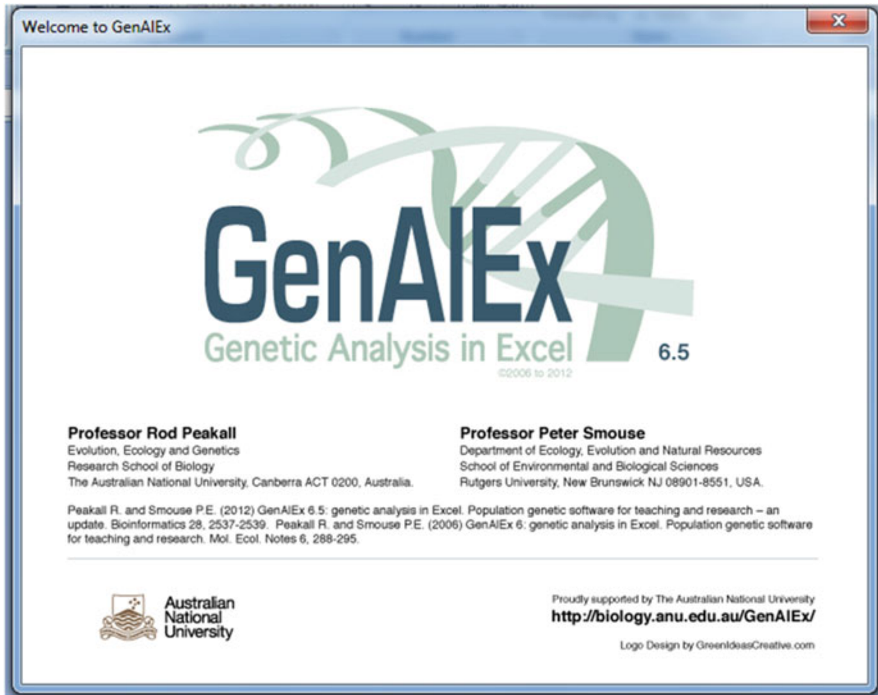


Fig. 6.2 Image of home page of GenAlEx 6.5 software

6.2.3 Forensic Statistics Analysis Toolbox (FORSTAT)

FORSTAT (Ristow and D'Amato 2017), is a online tool which can be applied for analysis of forensic importance parameters as well as population genetic analysis can be done automated and reduced the time and labor. This software overcomes on the limitations of the PowerStats software. It provides the results in numerical format as well as graphical form to improve the visualization of the results. This software needs input file in three digit formats for microvarinet alleles in STRs and two digit formats for bialleic STR markers (Weblink: https://fdl-uwc.shinyapps.io/forstat/) (Fig. 6.4).

FORSTAT online tool can calculate following allele frequency based parameters in a single excel document:

- Homozygosity
- Observed and expected heterozygosity
- Polymorphic information of the genetic data
- Exclusion power
- Matching probability
- Discrimination power
- Paternity index

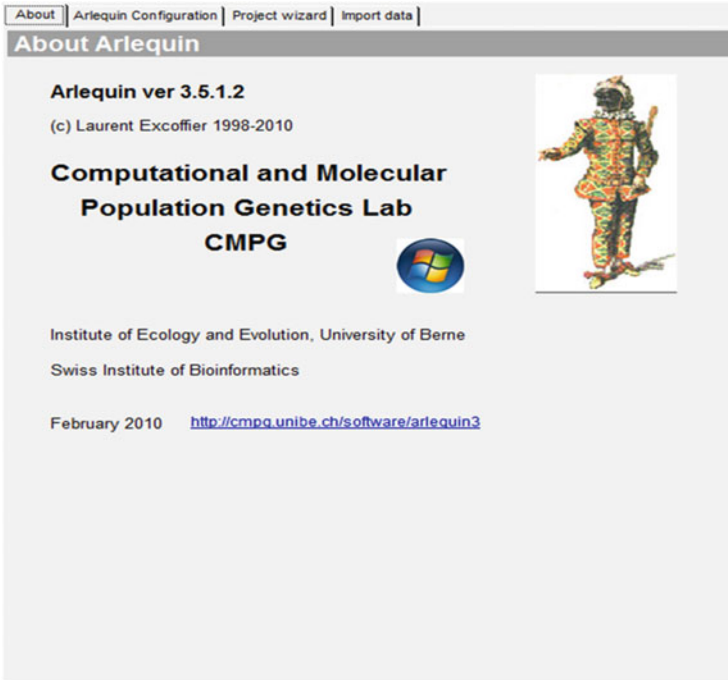


Fig. 6.3 Image of home page of Arlequin v3.5.1.2 software

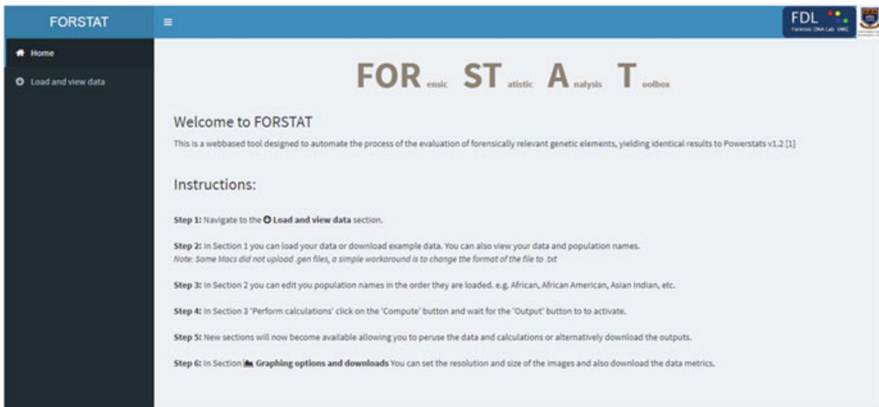


Fig. 6.4 Image of home page of FORSTAT online software

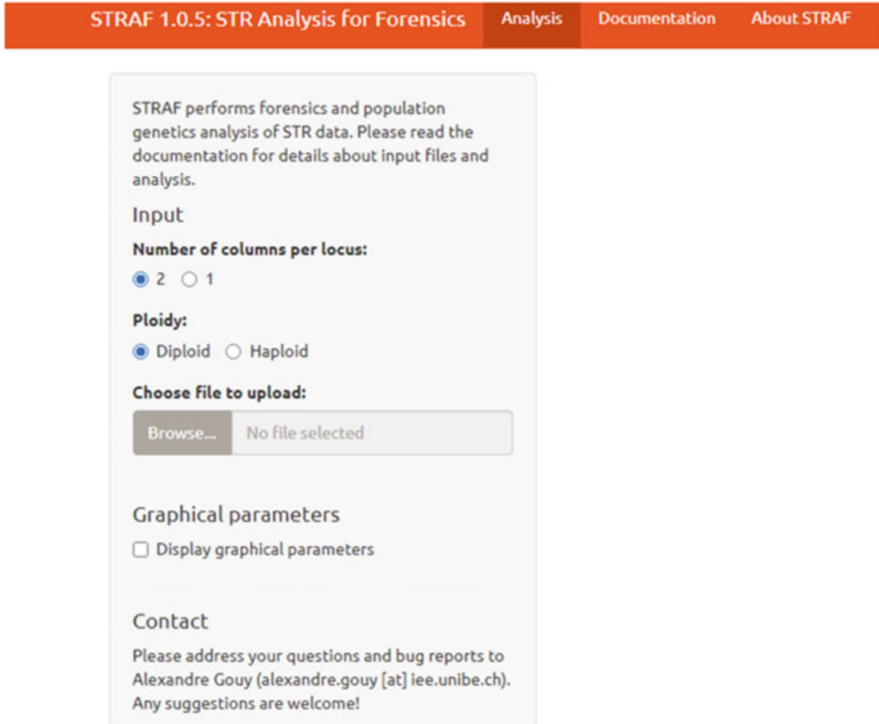


Fig. 6.5 Image of home page of STRAF online software

6.2.4 STRAF: STR Analysis for Forensics

STRAF (Gouy and Zieger 2017) is an online accessing tool to analyze variety of statistical analysis for genetic data from a single input file. For this software the input file format needs to be in .txt format and first column should contain the identification number of sample and second column needs to be the identification number of population. This software can be applied to analyze diploid as well as haploid data and it allows the microvariant alleles of STR data and 0 needs to be put for missing data (http://cmpg.unibe.ch/shiny/STRAF/) (Fig. 6.5).

This software can perform following analysis:

- Allele frequency
- Gene diversity (GD)
- Observed and expected heterozygosity
- Polymorphic information (PIC)
- Exclusion power (PE)
- Probability of a match (MP)
- Discrimination power (PD)
- Paternity index (PI)

- HWE-P value
- LDE
- Principal component analysis (PCA)

6.2.5 POPTREE Program

POPTREE2 (Takezaki et al. 2009) software can be used for the construction of phylogenetic tree based on genetic distances. The genetic distance based on allele frequency of the compared population were used in this software. The boot strap test is applied in this software to increase the acceptance level of the phylogenetic tree (Saitou and Nei 1987; Sneath and Sokal 1973; Felsenstein 1985). In POPTREE2, all the computations can be done through Windows-interface and the display of phylogenetic trees can easily be changed, copied to other applications, and printed by using icons.

Distance measures that can be used for the phylogeny construction are as follows: various genetic distances such as DA distance (Nei and Chesser 1983), Nei's genetic distance (Nei 1972), F*st* distance (Latter 1972), $(\delta\mu)^2$ distance (Goldstein et al. 1995), and DSW distance (Shriver et al. 1995) can be used for construction of the phylogenetic tree (Fig. 6.6).

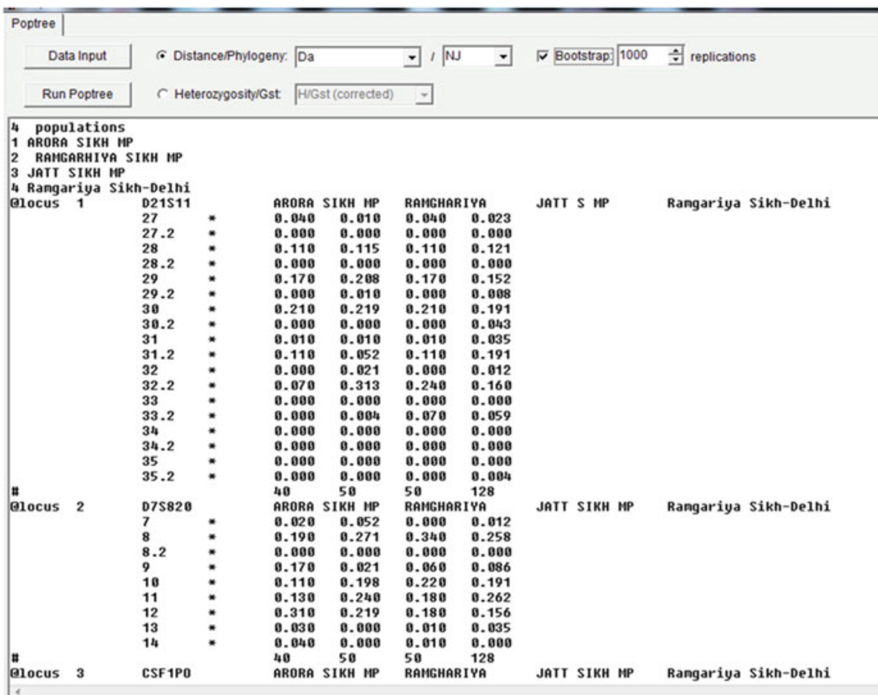


Fig. 6.6 Image of POPTREE2 program

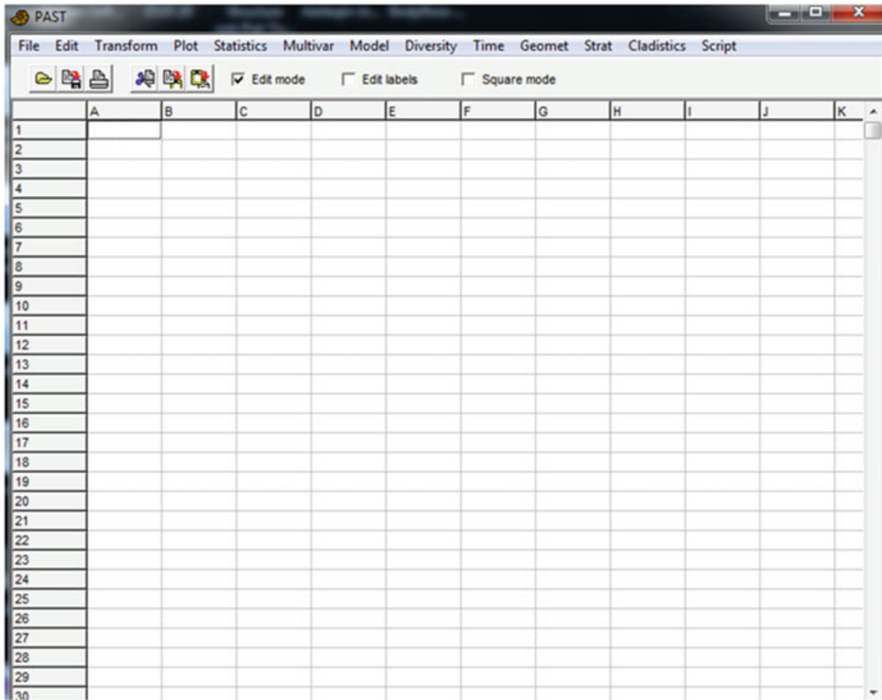


Fig. 6.7 Image of home page of PAST software

6.2.6 PAST Software

Paleontological STatistics (PAST) (Hammer et al. 2001) is a software which is applied in the vast area of research field. It allows direct copy and paste of the data from excel sheet to software, which makes it easy and user friendly. It can perform various analyses from single input file and provide graphical outcomes (Fig. 6.7).

6.2.7 STRUCTURE Software

The structure software (Evanno et al. 2005) is used for the analysis of population structure using multi locus genetic markers, Restriction Fragment Length Polymorphisms (RFLPs), and Single Nucleotide Polymorphism (SNPs). This software can be used for inferring the distance among populations, identifying of migrants, admixed, hybrid. To use this software, some parameters are to be set

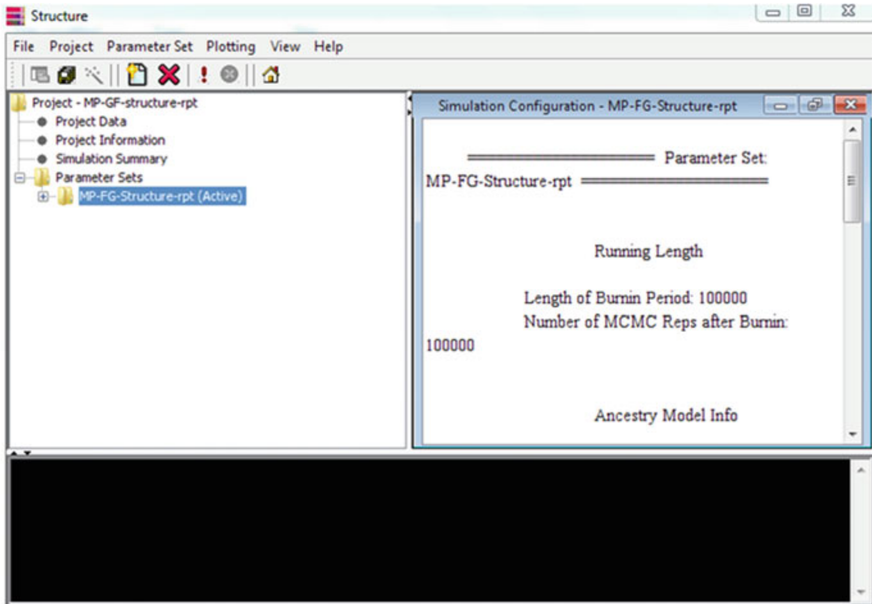


Fig. 6.8 Image of home page of STRUCTURE Software

before running it. To obtain acceptable delta *K* value, iterations should be set as 100 (Fig. 6.8).

6.2.8 Haplogroup Predictor Tools

There are several online haoplogroup predictor tools (Emmerova et al. 2017). Some of the most useful tools/softwares are as follows:

- Nevgen Y-DNA haplogroup predictor (www.nevgen.org)
  This online tool is used to predict haplogroups from Y-haplotype data as well as SNP data. In this online tool, there is pre-programmed sequence of 17 and 23 YSTR loci (Fig. 6.9).
- Whit Athey: Haplogroup predictor—HAPEST (www.hprg.com/hapest5) (Fig. 6.10) (Athey 2006)
- Jim Cullen: World Haplogroup & Haplo-'I' Subclade Predictor (members.bex.net/jtcullen515/haplotest.htm) (Fig. 6.11)
- Felix Immanuel: Y-Haplogroup Predictor (www.y-str.org/2013/06/y-haplogroup-predictor.html) (Fig. 6.12)

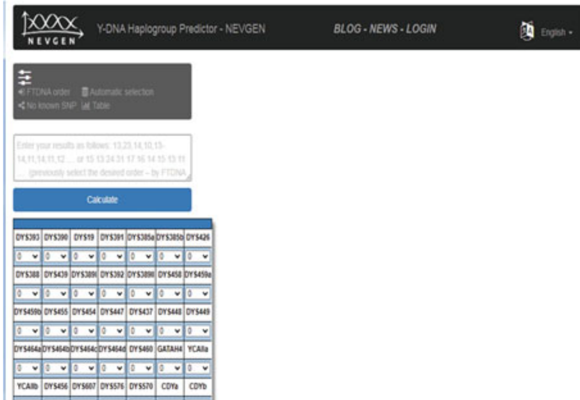


Fig. 6.9 Image of home page of NevGen Online software

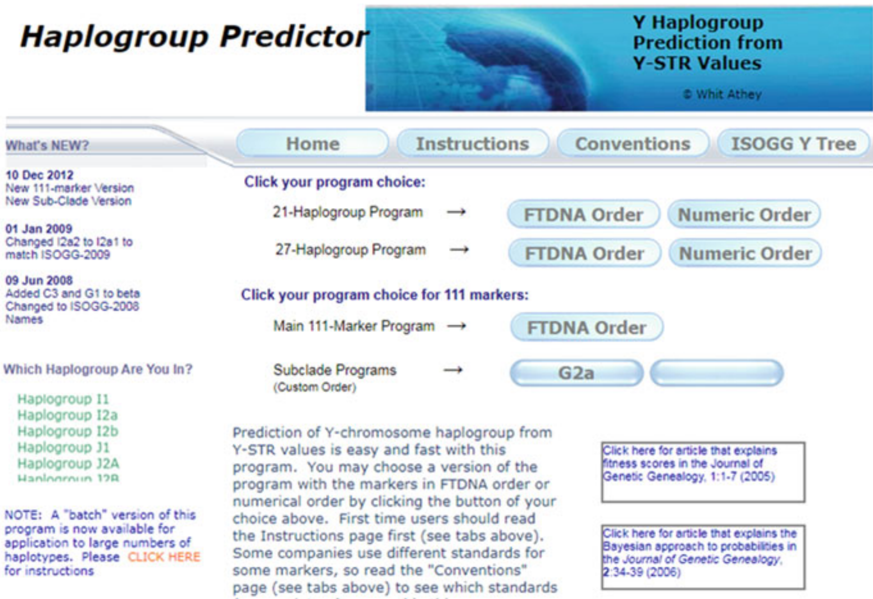


Fig. 6.10 Image of home page of Haplogroup predictor online software

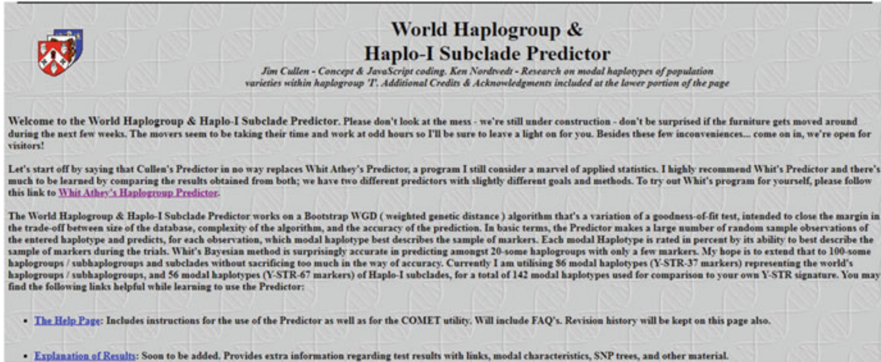


Fig. 6.11 Image of home page of World Haplogroup & Haplo-'I' Subclade Predictor online software

Fig. 6.12 Image of home page of Y-Haplogroup Predictor online software

6.3 Software Being Used for Interpretation of Forensic Genetic Evidence

6.3.1 OSIRIS (Open Source Independent Review and Interpretation System)

OSIRIS (Open Source Independent Review and Interpretation System) is the public domain application software and considered in light of suggestions of the interdisciplinary directorate (Kinship and DNA Advisory Panel, KADAP) consolidated by the

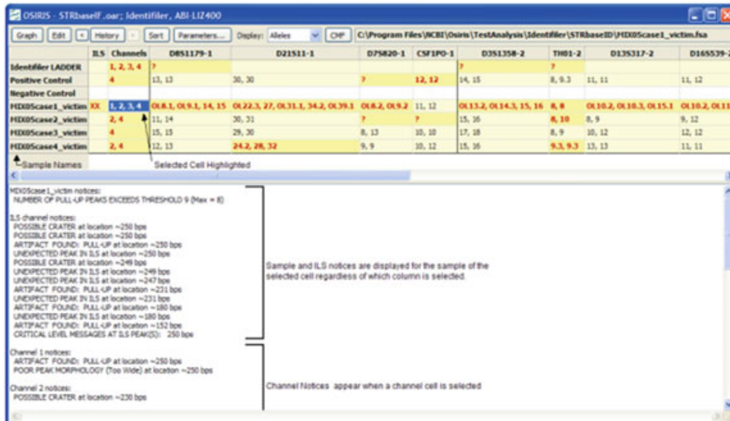


Fig. 6.13 Image of home page of OSIRIS software

US Department of Justice during the victim identification projects of the World Trade Center. It was first tried in Hurricane Katina victim identification and further created in a joint effort with government and state laboratories. It was designed by NCBI to give rapid STR analysis, fragment analysis, and identification of artifacts in forensic casework. This software runs on Windows XP, Windows 7, and Macintosh and analyzes data into .fsa and .hid file formats. It is not limited by number of STR kit dye color and also not limited to number of users on network (Fig. 6.13).

OSIRIS is a fast and powerful software based on mathematical algorithms and physical principles. This algorithm assists to coordinate the raw data generated by genetic analyzers with the mathematically characterized curves to the observed baseline and the peaks which permits OSIRIS to assess quality standards like:

- Displacement (shifting)
- Sample-to-ladder fit for alleles
- Residual displacement (displaced shifting)
- Noise and baseline
- Artifacts

OSIRIS software has different application such as:

- Clinical testing
- Forensic casework
- Identification such as paternity testing
- Biosample identification
- Kinship analysis and
- Cell line authentication
- Training purposes

Apart from these applications, it is used in DNA profiling quality check, DNA profile analysis, CODIS, automatic reanalysis, and process monitoring. OSIRIS software is also compatible with mixture interpretation software (Riley et al. n.d.).

OSIRIS software works with all commercial marker systems leading companies i.e., Applied Biosystems-Now Thermo Fisher scientific, Promega; Qiagen (Web link: http://www.ncbi.nlm.nih.gov/projects/SNP/osiris/).

6.3.2 LRmix Studio

LRmix Studio is an unpaid, open source, and nonproprietary software which is devoted in interpretations of complex DNA mixtures in forensic samples. This software measures the supplementary value of any forensic complex DNA profiles based on autosomal STR kits. It interprets data by using likelihood ratio model (Haned et al. 2012; Gill and Haned 2013). This model accommodates for the dropout/drop-in phenomena and evaluates the allelic dropout rate among the obtainable information, and adopts these calculations to create LR ratio. This tool (software) was formulated by Jeroen de Jong and Hinda Haned and supported by Netherlands Organization for Scientific Research (NWO). LRmix Studio software can be tested for propositions of at least four unknown and maximum of up to five contributors. Application of this software includes comparison of any number of replicates produced from a specific DNA sample to any reference profiles. Despite that, this has been examined and authenticated for five replicates and not more than three reference profiles. LRmix Studio can also be useful to interpret SNP mixtures (Web link: http://lrmixstudio.org/) (Fig. 6.14).

About LRmix Studio

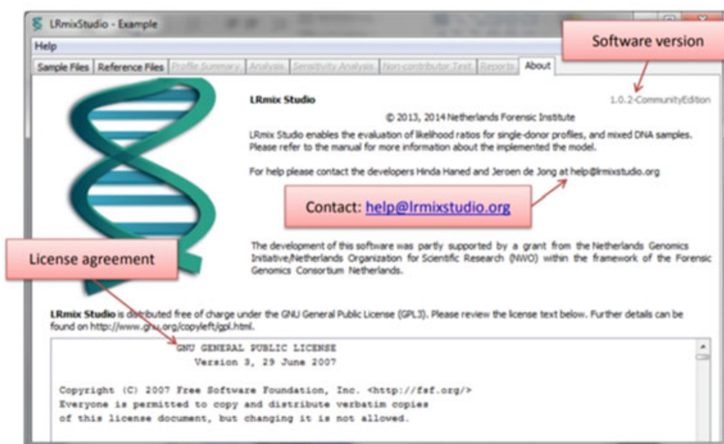


Fig. 6.14 Image of home page of LRmix Studio software

Fig. 6.15 Image of home page of forensim software

6.3.3 Forensim Package for the Open Source R Statistical Software

This is free software and available in R language. Forensim is a statistical tool being used for the elucidation of forensic samples containing DNA mixtures (Haned 2011). This was the first open source software, dedicated toward the simulation of forensic genetics data. It depends upon three object classes for the generation and the storage of forensic casework data: genotypes, mixtures of DNA samples and allele frequencies (Web link: http://forensim.r-forge.r-project.org/).

Forensim tool provides different features based on the simulation tools and statistical tools.

Simulation methods are dedicated to the simulation of genetics data obtained from forensic casework like:

- Individual genotypes
- Mixed DNA samples
- Individual genotypes (Fig. 6.15)

Statistical methods are dedicated to facilitating the evidentiary value of the DNA exhibits: It is used to interpret: number of patron in the given mixed stain and to address the statistical weight of mixed DNA stains. This package also includes some other methods for statistical analysis:

- Random match probability
- Likelihood ratios
- Random man exclusion probability
- Profile probability

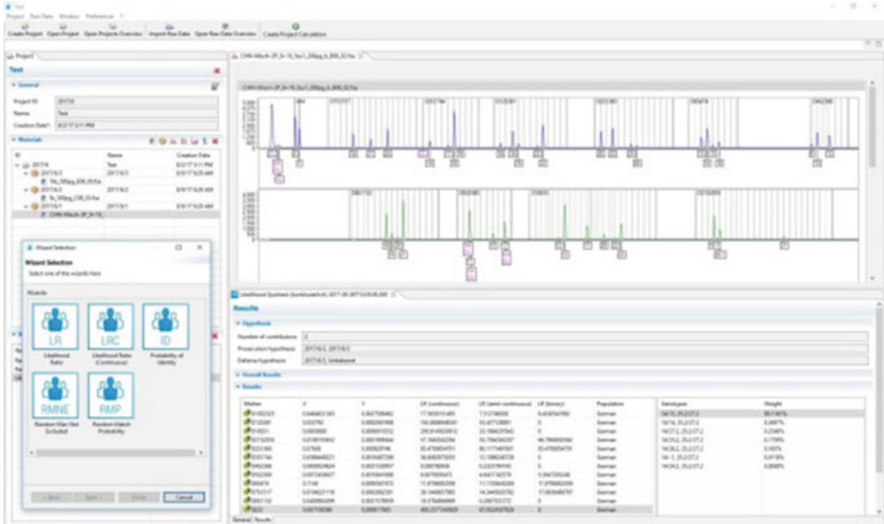


Fig. 6.16 Image of home page of GenoProof Mixture 3 User Interface software

6.3.4 Genoproof Mixture 3

Genoproof Mixture 3 software is new software used by forensics expert for the analysis of complex DNA mixtures. This software approach is a fully continuous model which allows using all the available information of profile and thus increases the mixture interpretation capabilities. Several other statistical methods are used for the interpretation of samples containing complex DNA mixtures (Fig. 6.16) (Götz et al. 2017).

The binary model only holds on presence but not on dropout/dropin events; semicontinuous model consider for dropout/dropin events however, fully continuous model not only consider the above events but also the other parameters such as: prestutter ratio, fragment size, and peakheight. Genopoof Mixture software supports all the three models for mixture interpretation. It is all in one software (weblink: https://www.qualitype.de/en/solutions/products/evaluation-software/genoproof-mixture/genoproof-mixture-3/).

The following analysis can be done with Genopoof Mixture software:

- Probability of identity
- Data management
- Quality assurance
- Raw data analysis
- Biostatistical calculations
- Mixed sample analysis

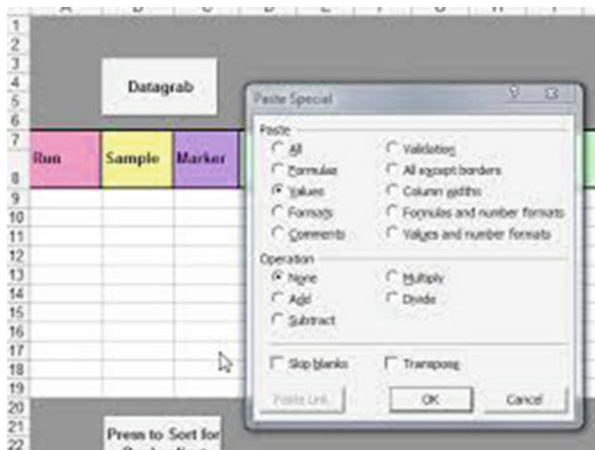


Fig. 6.17 Image of home page of MixtureCalc v1.2 software

6.3.5 MixtureCalc v1.2 Software

MixtureCalc v1.2 is an excel-based software which is used for interpretation of mixed samples and qualitative and quantitative analysis of the data from GeneMapper file (Exelsheetlink: MixtureCalc v1.2 Excel sheet).

It is freeware version of software and only supports SPSA casework. It allows user to perform several tasks such as:

- Datagrab the data from an exported GeneMapper file
- Peak adjustment or stutter adjustment
- Mixture sheets (final) calculation
- Print the sheets data and save the calculation sheets
- Provides calculation sheets for multiple contributors
- Computes likelihood ratio
- Calculate conditional match probability and match probability (Fig. 6.17)

6.3.6 Mixture Analysis Software

Mixture Analysis software is a freeware software which interprets the mixtures by utilizing peak height or peak area information. The analysis can be further carried on excel spreadsheet and used to eliminate certain combinations of mixture (Gill et al. 1998). A total of 270 known loci mixtures were successfully validated against this model. In most of these the right derivation positioned first and none were disposed of (Fig. 6.18).

Note: This software isn't approved for casework and no warrantee is accommodated such a reason.

Fig. 6.18 Image of home page of Mixture Analysis software

Fig. 6.19 Image of home page of EuroForMix software

6.3.7 EuroForMix

EuroForMix is an open source, continuous model for forensic professionals to determine the likelihood ratio (LR) where a statistical model for the peak heights is incorporated. It is also used to interpret complex SNP mixtures (Bleka et al. 2016) (Web link: http://www.euroformix.com) (Fig. 6.19).

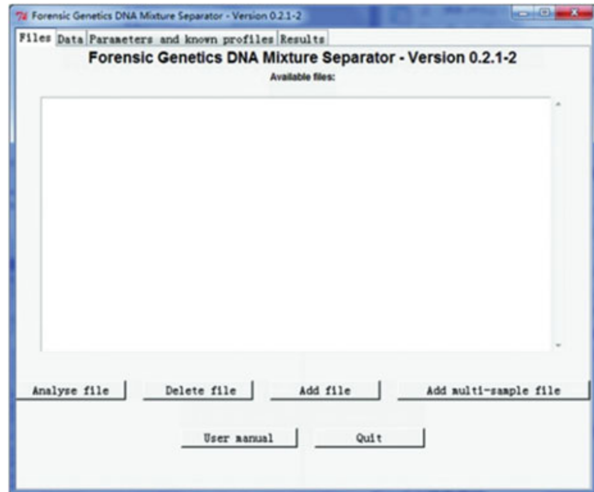


Fig. 6.20 Image of home page of mixsep software

6.3.8 Torben's Mixture Separation Software

The mixsep software (Tvedebrink 2013, 2011) is a forensic genetics statistical tool to interpret mixed DNA samples using greedy algorithm. This software is able to separate the DNA profiles of two or three persons and identification of individual DNA profile is not influenced by the allele dropin and out, and stutter (Fig. 6.20).

The mixsep software contains six variables for data analysis namely; locus, allele, height of allelic peak, area of allelic peak, base pair and dye.

However majority of the data are analyzed using first four variables and saved as CSV file (weblink: http://cran.r-project.org/web/packages/mixsep/)

6.3.9 DNAMIX Software

DNAMIX Software is the tool for forensic geneticist to calculate the likelihood ratios from forensic samples containing mixed DNA stains. This tool calculates the complex DNA mixtures and as well as single source contributors. The new version DNAMIX v.3 is based on the John Storey's DNAMIX v.1 and v.2. The formulas and methods explained by Curran et al. 1999 are based on assumption of independence of all allele's proportions in the mixture by using Hardy-Weinberg and linkage equilibrium. The assumptions are completely independence among individuals however; within the same population low-level dependency between individuals is ignored. The latest of DNAMIX v.2 is written in FORTRAN 90 and includes population structure into the calculations. If θ is equal to 0, the likelihood ratios in DNAMIX v.2 version will be equal to DNAMIX v.1 version (weblink: http://www.biostat.washington.edu/~bsweir/DNAMIX3/webpage/).

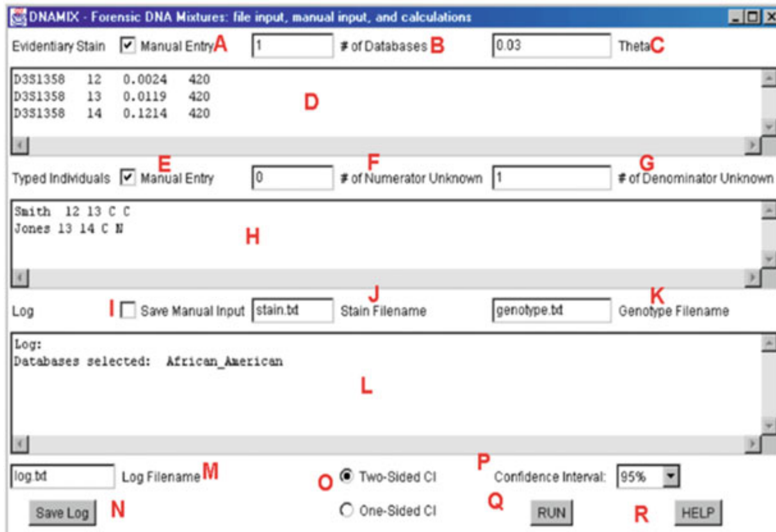


Fig. 6.21 Image of home page of DNAMIX software

Two important points should be remembered while using DNAMIX v.2 version:

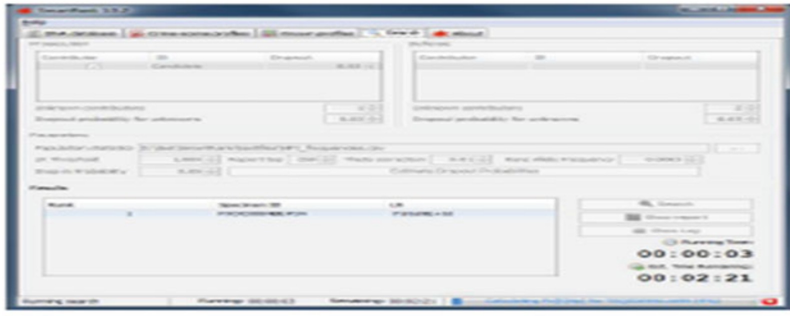
- Same Likelihood ratios probability but different numerator and denominator probability
- The individual profile is identified as known contributors/unknown contributors in numerator and as well as denominator of the likelihood ratios (vice versa) (Fig. 6.21).

6.3.10 FamLink Kinship Software

FamLink kinship software is free; user-friendly statistical software gives a convenient graphical interface for the user. This application uses pair wise linked DNA marker data to calculate the likelihood ratio for family relationships or pedigrees testing (Kling et al. 2012). The statistical calculations are dependent on ancestry structure and record for connection between sets of markers. FamLink utilizes the Merlin engine for numerical computations (web link: http://www.FamLink.se).

6.3.11 SmartRank Software

SmartRank is a free and user-friendly software in forensics which allows searching voluminous DNA databases to complex/partial DNA mixtures. While searching the



Percentage of drop-out	Number of contributors 2	3	4	5 (1 known)	5 (2 knowns)
0% [a]	Successful SmartRank search expected				
5% [b]			The loss of true donors is expected		
15%		Successfulness dependent on complexity; further considerations advised			
30%		Some or (all) donors may be missed			
50%	SmartRank searches did not yield informative results and are therefore not advised				

Fig. 6.22 Image of home page of SmartRank software

large national database, SmartRank software keeps minimum rate of false negative and false positive errors. This software is successful for searching voluminous DNA database with complex/partial mixtures containing 0% dropout, 15 or 16 alleles and maximum four unknown source however, not applicable on DNA containing 50% dropout (Fig. 6.22) (Benschop et al. 2017).

SmartRank software calculates the likelihood ratio for each tested individual in the DNA database and distributed for ranked genotypes. SmartRank software also follows the same model behind LRmix Studio (web link: http://www.LRmixStudio.org/SmartRank).

6.3.12 Bracket Method Software

This software is used to generate DNA profiles from single or multiple polymerase chain reactions using a system of soft and hard brackets. The major and minor donors separated in the DNA profile using allelic balance threshold. In addition to reproducible alleles, bracket method also permits the irreproducible alleles and keeps them segregated from confirmed alleles with systems of brackets.

The DNA profile is constructed by the use of qualitative (allele calls) and quantitative (peak heights) data. This DNA profile contains three layers: (1) a dominant profile with duplicatable alleles, (2) reproducible alleles whose peak

Fig. 6.23 Image of home page of bracket software

heights are underneath the allelic balance variable threshold, and (3) alleles which are non-reproducible. The balance of allelic threshold variable can be balanced through the Graphical User Interface and will decide at what threshold alleles will be appointed as related to a dominant profile or to a minor donor profile (Fig. 6.23) (Bekaert et al. 2012).

The script needs information from at least a single PCR replicate of one STR multiplex kit and can use a maximum of two replicates by two different STR multiplex kits.

6.3.13 STRmix™ Software

STRmix™ software is used in forensic casework to resolve the complex DNA mixtures. This software approaches fully continuous model and combines mathematical and biological modeling to interpret complex mixtures (Bright et al. 2018) (Fig. 6.24).

The characteristic of STRmix™ software is:

- Faster interpretation
- Easy accessible
- Easily understood in court
- Uses analytical threshold
- Model stutter and peak height
- Model dropout and drop-in
- Integrate DNA profiles from different kits in the same elucidation

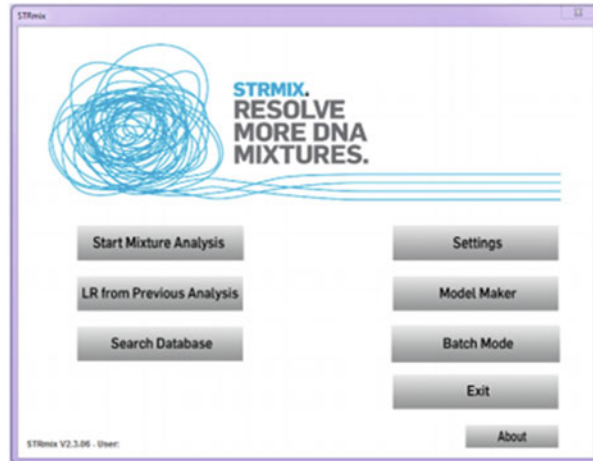


Fig. 6.24 Image of home page of STRmix™ software

Likelihood Ratio is used for the comparisons of samples to reference profiles in database. STRmix™ incorporates a capacity that permits the software to coordinate blended DNA profiles straight forwardly against a database. This is a meaningful step forward for situations where there are no suspects and there is DNA from various suspects in a single sample. STRmix™ is extensively used in forensic casework at many laboratories (web link: https://www.strmix.com/).

6.3.14 DNA Data Bases

6.3.14.1 Autosomal STR Database

STRidER (STRs for Identity ENFSI Reference Database) (Bodner et al. 2016) is the extended and improved adaptation of the ENFSI STRbASE (2004–2016). This curated online top notch STR allele frequency population database empowers scientifically reliable STR genotype likelihood assesses and gives quality control of autosomal STR information. STRidER act as recurrence database and programming stage for the advancement of novel tools for STR QC and other forensic investigations. It can oblige length-based STR alleles and sequence information.

STRidER database are used in forensic community and serves many purposes such as:

- It can calculate allele frequency.
- It can directly search the high quality STR data like autosomal allele frequency.
- It offers quality control of the STR data before it gets published.
- It gives accessibility to download allele-frequency tables of STR loci.
- Once the data gets accepted, it will be shown online with unique and traceable STRidER accession number.
- It does not give permission to access the individual STR genotypes data in order to maintain privacy regulations (Fig. 6.25).

STRidER
STRs for identity ENFSI Reference database, v2/R2

Contribute

Submitting lab — e.g. Institute of Legal Medicine, Innsbruck, Austria
Contact person
Contact email
Repeat email address
Type of population — e.g. tribe, admixed, general population
Exclusion/inclusion criteria — e.g. none, only male, only over 35 y.o.
Unrelatedness — e.g. ten generations, assessed by interviews
Geographic origin country
Region/city — city or region specifying the origin more closely

Fig. 6.25 Image of home page of STRidER online tool

6.3.14.2 Y Chromosomal STR Database

Y-STR Haplotype Reference Database (YHRD)

This is largest Y chromosomal data bank which allows the submission of validated Y-haplotype data. It is an open access database includes Y-STR data of various populations. Currently, the YHRD includes 137 national databases and these data allow statistical analyses for various forensic casework applications and interpretation purposes such as ancestry prediction, frequency estimation and kinship and mixture analysis. YHRD is also used to analyze meta population information. This database is widely useful for law-enforcement agencies and forensic personnel.

The main objectives of the Y-STR Haplotype Reference Database are:

- To evaluate the authentic frequency for Y-STR and Y-SNP haplotypes with the goal that they can be used in analysis of matches in forensic cases.
- It portrays the male ancestries to draw deduction about the causes and family of human population.
- It provides all the tools and resources related to Y-STRs and Y-SNPs.

Most of the international journal published only those Y-haplotype data that have been submitted in YHRD database and have accession number.

Y-STR Haplotype Reference Database provides various tools for haplotype analyses like:

- Validation of Data
- AMOVA (Analysis of Molecular Variance)

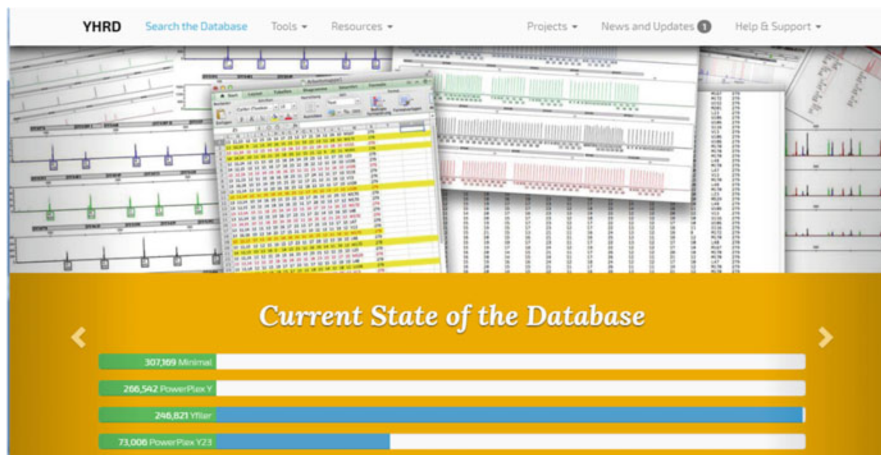


Fig. 6.26 Image of home page of YHRD online tool

Table 6.1 Current state of Y-STR Haplotype Reference Database is as follows

Dataset	Number of haplotypes
Minimal	307,169
PowerPlex Y	266,542
Yfiler	246,821
PowerPlex Y23	73,006
Yfiler Plus	73,810
Maximal	23,221

- Mixture analysis
- Kinship Analysis
- Multidimensional Scaling plot (MDS)

It was created in 1999 by Sascha Willuweit and Lutz Roewer Charité – Universitätsmedizin Berlin and inscribed by the ISFG (Fig. 6.26, Table 6.1).

6.3.14.3 X Chromosomal STR Database

ChrX-STR.org 2.0 (Web link: www.ChrX-STR.org). This web site was founded by: Reinhard Szibor, Jeanett Edelmann, and Sandra Hering. ChrX-STR.org 2.0 is the database of X chromosomal markers especially forensic interests. This database provides ChrX STRs and ChrX STR haplotypes for forensic purposes related to X-STR analysis. Various software which are useful in X-STR analysis, are directly linked with this X-STR database. This database incorporates many affairs concerning the utilization of X chromosomal markers for forensic purposes (Fig. 6.27).

This database contains four population haplotypes in which different-different number of chromosomes is investigated:

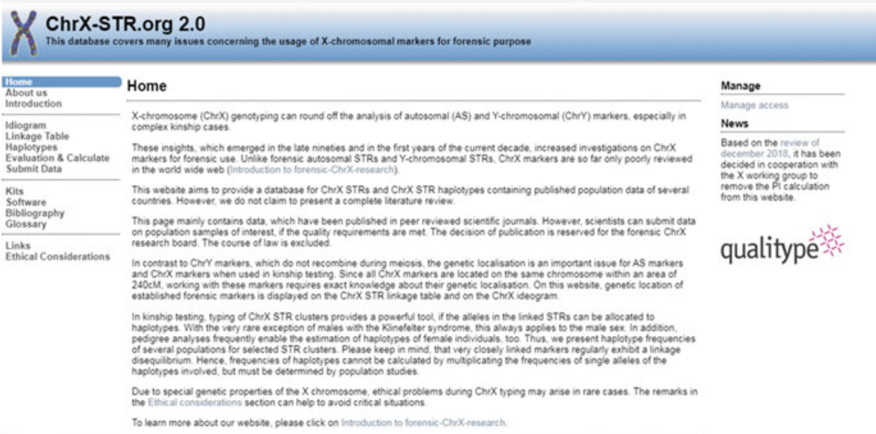


Fig. 6.27 Image of home page of ChrX-STR.org 2.0 online tool

- German: 627 chromosomes
- Ghanesen: 59 chromosomes
- Japanese: 93 chromosomes
- Chinese Hen: 144 male chromosomes

Software including X-Linked STR Markers

- GenoProof
- GenoProof Mixture
- VAT
- DNAVIEW
- Rutgers combined linkage-physical human genome map

6.3.14.4 Mt. DNA Database

Mitomap

Mitomap is the database of mitochondrial DNA, which provides comprehensive information related to mutations and polymorphisms in human mitochondrial DNA. MITOMAP attempts to bring the broad spectrum of structure and function of mtDNA, pathogenic, molecular, useful and clinical data into a single unified element brought together component which can be questioned from a wide range of perspectives. MITOMAP is right now utilizing the Sybase rational database management software. It is both a free information system for the mitochondrial researcher and moreover gives the mtDNA map and clinical information assets for the Genome and for Online Mendelian Inheritance Man.

Population dataset gives the permission to polymorphic restriction sites: clinical dataset for base changes and nucleotide positions like disease-associated mutations

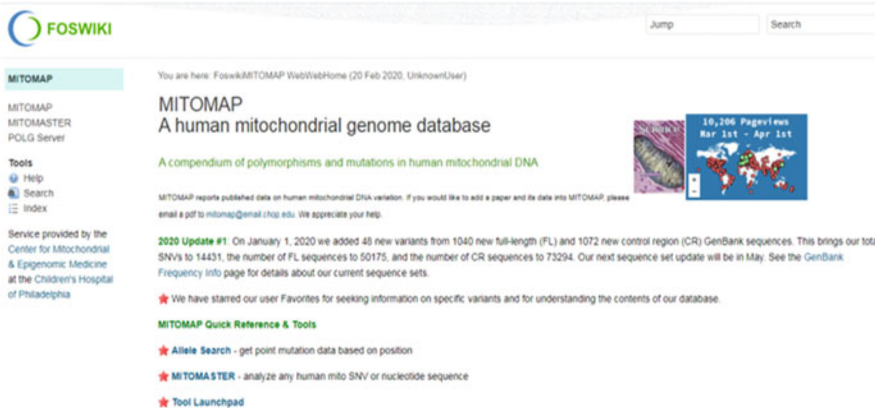


Fig. 6.28 Image of home page of Mitomap

and their clinical phenotypes; functional data for genetic element functions of mtDNA; and gene-gene interaction data for information of mitochondrial DNA. Mitomap is online database available for both browsing and querying capacity (Fig. 6.28).

EMPOP

EDNAP Mitochondrial DNA Population Database EMPOP is the largest and high-quality database which has the following characteristics:

- Free available database
- It uses SAM 2 software for the evaluation and query engine
- It conducts quality control check on mitochondrial DNA data prior to manuscript submission
- It uses quasi-median network for quality control
- It uses maximum likelihood approach (EMMA) for haplogrouping
- It conducts alignment free database searches (SAM)
- The updated version of EMPOP determines haplogoup of mitochondrial DNA sequences with the help of maximum likelihood concept (Fig. 6.29, Table 6.2).

The EMPOP uses some of the tool tools to perform the analysis and interpretation of mtDNA sequence variation such as:

- Haplogroup Browser
- EMPcheck
- NETWORK

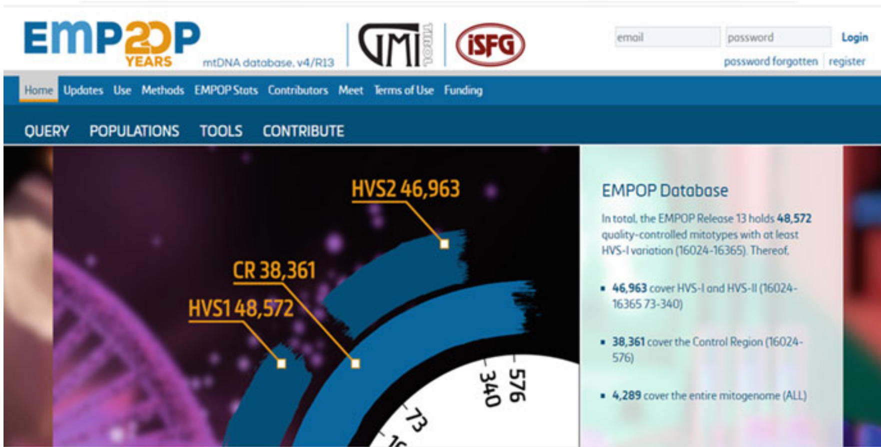


Fig. 6.29 Image of home page of EMPOP

Table 6.2 EMPOP database

Mitotypes	Region
48,572	Total with at least HVS-1 Variation
46,963	HVS-I and HVS-II
38,361	Control Region
4289	Entire mitogenome

6.3.14.5 Short Tandem Repeat Makers Database

Short Tandem Repeat DNA Internet Database which is commonly known as STRBase has been regulated by National Institute of Standards and Technology (NIST) which contains data on commonly used STR markers. This internet available datasheet is used forensic genetic community.

STRBase includes:

- Population data
- PCR primer sequences
- Commercial multiplex STR kits
- Sequence information
- Mutation Rates for Common Loci
- Technology related to STR analysis
- Comprehensive reference lists
- Y chromosome STRs
- Mixture Interpretation
- Tri-allelic pattern
- DNA Marker Information

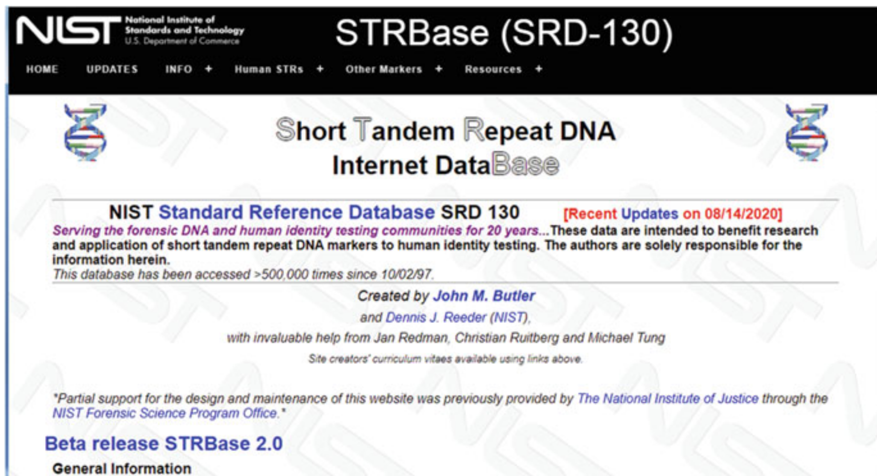


Fig. 6.30 Image of home page of STRBase

Table 6.3 List of databases used for forensic genetics

Database	DNA data types to be allow to submit	Year of launch and name of curator	Link
STRidER	Autosomal STR	2016 and Walther Parson	www.strider.online
YHRD	Y chromosomal STR data	1999 and Lutz Roewer and Sascha Willuweit	www.yhrd.org
ChrX-STR	X chromosomal STR data	Not available	www.ChrX-STR.org
Mitomap	Mitochondrial DNA data	1996 and Dr Douglas C. Wallace	www.mitomap.org
EMPOP	Mitochondrial DNA database	2006 and Walther Parson	http://www.empop.org
STRBase	Short tandem repeat DNA Database	1997 and John M. Butler	https://www.cstl.nist.gov/biotech/strbase/

Population data studies gets revealed in the writing and feed in the STRBase by posting the STR framework, target population, the quantity of irrelevant people tested, and the reference. These set of population examines gives a significant device to finding references that contain STR allele frequencies to help in ascertaining matching probabilities for DNA profiling cases (Fig. 6.30, Table 6.3).

References

Athey TW (2006) Haplogroup prediction from Y-STR values using a Bayesian-allele-frequency approach. J Genet Geneal 2(2):34–39

Bekaert B, Van Geystelen A, Vanderheyden N, Larmuseau MHD, Decorte R (2012) Automating a combined composite–consensus method to generate DNA profiles from low and high template mixture samples. Forensic Sci Int Genet 6(5):588–593

Benschop CCG et al (2017) Validation of SmartRank: a likelihood ratio software for searching national DNA databases with complex DNA profiles. Forensic Sci Int Genet 29:145–153

Bleka Ø, Storvik G, Gill P (2016) EuroForMix: an open source software based on a continuous model to evaluate STR DNA profiles from a mixture of contributors with artefacts. Forensic Sci Int Genet 21:35–44

Bodner M et al (2016) Recommendations of the DNA Commission of the International Society for Forensic Genetics (ISFG) on quality control of autosomal Short Tandem Repeat allele frequency databasing (STRidER). Forensic Sci Int Genet 24:97–102

Bright J-A et al (2018) Internal validation of STRmix™–a multi laboratory response to PCAST. Forensic Sci Int Genet 34:11–24

Butler JM (2001) Forensic DNA typing: biology & technology behind STR markers. Academic Press, New York, NY

Curran JM, Triggs CM, Buckleton J, Weir BS (1999) Interpreting DNA mixtures in structured populations. J Forensic Sci 44(5):987–995

Emmerova B, Ehler E, Comas D, Votrubova J, Vanek D (2017) Comparison of Y-chromosomal haplogroup predictors. Forensic Sci Int Genet Suppl Ser 6:e145–e147

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14(8):2611–2620

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10(3):564–567

Felsenstein J (1985) Confidence-limits on phylogenies–an approach using the bootstrap. Evolution 39:783–791. https://doi.org/10.2307/2408678

Gill P, Haned H (2013) A new methodological framework to interpret complex DNA profiles using likelihood ratios. Forensic Sci Int Genet 7(2):251–263

Gill P, Sparkes R, Pinchin R, Clayton T, Whitaker J, Buckleton J (1998) Interpreting simple STR mixtures using allele peak areas. Forensic Sci Int 91(1):41–53

Goldstein DB, Linares AR, Cavalli-Sforza LL, Feldman MW (1995) Genetic absolute dating based on microsatellites and the origin of modern humans. Proc Natl Acad Sci 92(15):6723–6727

Götz FM, Schönborn H, Borsdorf V, Pflugbeil A-M, Labudde D (2017) GenoProof Mixture 3—new software and process to resolve complex DNA mixtures. Forensic Sci Int Genet Suppl Ser 6:e549–e551

Gouy A, Zieger M (2017) STRAF—a convenient online tool for STR data evaluation in forensic genetics. Forensic Sci Int Genet 30:148–151

Hammer Ø, Harper DAT, Ryan PD (2001) PAST: paleontological statistics software package for education and data analysis palaeontol. Electronica 4:1–9

Haned H (2011) Forensim: an open-source initiative for the evaluation of statistical methods in forensic genetics. Forensic Sci Int Genet 5(4):265–268

Haned H, Slooten K, Gill P (2012) Exploratory data analysis for the interpretation of low template DNA mixtures. Forensic Sci Int Genet 6(6):762–774

Kling D, Egeland T, Tillmar AO (2012) FamLink–a user friendly software for linkage calculations in family genetics. Forensic Sci Int Genet 6(5):616–620

Latter BDH (1972) Selection in finite populations with multiple alleles. III. Genetic divergence with centripetal selection and mutation. Genetics 70(3):475–490

Nei M (1972) Genetic distance between populations. Am Nat 106(949):283–292

Nei M, Chesser RK (1983) Estimation of fixation indices and gene diversities. Ann Hum Genet 47 (3):253–259

Peakall ROD, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Resour 6(1):288–295

Riley SSG, Goor R, Hoffman D, Kalafut T, Noreuil C, Scott T, Sutton J Osiris version 2.2: intelligent casework and codis str analysis

Ristow PG, D'Amato ME (2017) Forensic statistics analysis toolbox (FORSTAT): a streamlined workflow for forensic statistics. Forensic Sci Int Genet Suppl Ser 6:e52–e54

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4(4):406–425

Shriver MD, Jin L, Boerwinkle E, Deka R, Ferrell RE, Chakraborty R (1995) A novel measure of genetic distance for highly polymorphic tandem repeat loci. Mol Biol Evol 12(5):914–920

Sneath PHA, Sokal RR (1973) Numerical taxonomy. The principles and practice of numerical classification. W. H. Freeman and Co., San Francisco, CA

Takezaki N, Nei M, Tamura K (2009) POPTREE2: software for constructing population trees from allele frequency data and computing other population statistics with Windows interface. Mol Biol Evol 27(4):747–752

Tereba A (1999) Tools for analysis of population statistics. Profiles DNA 2:14–16

Tvedebrink T (2011) mixsep: an R-package for DNA mixture separation. Forensic Sci Int Genet Suppl Ser 3(1):e486–e488

Tvedebrink T (2013) mixsep: DNA mixture separation. R package version 0.2. Computer Software, pp 1–2

Part II

Challenges in Forensic DNA Typing

7 Ancient DNA Analysis and Its Relevance in Forensic DNA Fingerprinting

Mian Sahib Zar and M. Aslamkhan

Abstract

The chapter presents the timeline, developments, nature, challenges, applications and prospects of ancient deoxyribonucleic acid (aDNA) technology. It also highlights the vision of future perspective that makes the field of aDNA technology so important and applicable to the people today. DNA is a blueprint of book of life that encodes genetic information. Human DNA consists of 3.2 billion nucleotides, comparable to letters in a book. DNA, the language of life, is present in all biological material—animals, plants. It degrades very slowly even after the so called death of living beings. This degrading DNA remains preserved in dead and fossilized living material is extractable and analyzable, called as ancient DNA (aDNA). Mostly aDNA is recovered from archaeological, anthropological, paleontological, museum specimen and clinical material. The age of aDNA ranges from 100 years, or even less, to millions of years and it has undergone natural environmental impact. Ancient DNA technology can be used in evolutionary biology, anthropology, medicine, agriculture, ancient diseases, living descendants by establishing a direct genetic link between a dead and a living individual, sex identification, animal or plant identification, kinship analysis, law enforcement and even determining the migration patterns of human populations. Due to recent ongoing development, the aDNA technology has become a factual, truthful, reliable and reproducible research field that recuperate DNA information

M. S. Zar
Department of Forensic Sciences, University of Health Sciences, Lahore, Pakistan

Center for Synthetic Biology Engineering Research (CSynBER), Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, P. R. China

M. Aslamkhan (✉)
Department of Human Genetics and Molecular Biology, University of Health Sciences, Lahore, Pakistan
e-mail: genetics@uhs.edu.pk

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_7

from paleontological, anthropological and archaeological finds allowing us to study the genetic associations between past and the existent organisms.

Keywords

aDNA · Paleontology · Archaeology · Genetics · Inter disciplinary research

7.1 Introduction

DNA, abbreviation for deoxyribonucleic acid, is the basic life giving material in all species of animal and plants. It is generally present in the nucleus of cell in the form of a double helix and called as nuclear DNA. Besides nuclei, DNA is also present in mitochondria and hence referred as mitochondrial DNA. Genome is the name given to the total amount of DNA in a species. In human genome, mitochondrial DNA (mtDNA) consists of 16.5 kbp (kilo base pair) and nuclear DNA is 3 billion bp (Lander et al. 2001). DNA in humans is present in 46 chromosomes, which exist in the nucleus of all body cells except the red blood cells (RBCs). The human nuclear genome contains 5% coding region referred as exons and 95% noncoding region called as introns. The exons are highly conserved and the introns are highly polymorphic (Daniel and Walsh 2006). In each human body cell, DNA is present as a 6–9 ft long coiled thread. On unwinding of all DNA in human body cells it would reach to the sun and back one thousand times. It is interesting to note that sequencing reveals 99% similar DNA in all humans. Of the remaining 1% only, it's one-tenth, is responsible to make us unique (Feuk et al. 2006). Remember that human DNA is not unique as more or less of it is shared with many other past and present species of animals and plants. Thus 98% of human DNA is similar with that of chimpanzee and 50% is similar to that of banana. If we write the information in a single cell on paper it needs 46 volumes of Webster dictionary. DNA is a blueprint or book of life. Human DNA consists of 3.2 billion letters. Human genome sequence was completed in 2000, by a consortium of 18 laboratories in five countries. There are three chemical substances in DNA: (a) "deoxyribose sugar", (b) "phosphate group" and (c) "nitrogenous base". Of these three, deoxyribose, sugar and phosphate group, are similar in DNA of all persons, but the nitrogenous base differs as it characterizes individual constituent of the polymer. This variation gives differentiating characters in each person, which helps in individualistic identification. Each nitrogenous base consists of, as a general rule, one of the four letters of nucleotide; A, T, C and G. These four nucleotides—adenine, thymine, cytosine and guanine are complement to each other in a permanent order or sequence: A pairs with T and C pairs with G (Fig. 7.1). This complementation of four nucleotides makes a DNA molecule (Luftig and Richey 2001).

Nucleic acid molecule, as we know, is very long and called as macromolecules, which means that it is a molecule made up of smaller molecular units, which are called nucleotides. The nucleotide linked chemically to each other to form a chain. In DNA, the four nucleotides named as adenine, thymine, cytosine and guanine, as

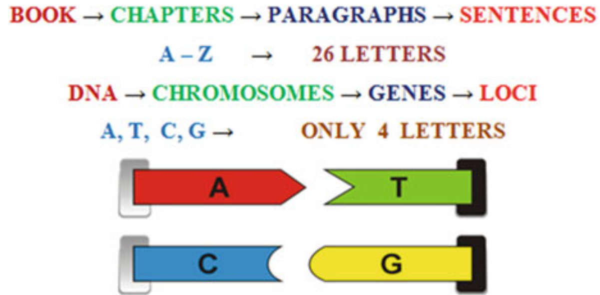


Fig. 7.1 Fundamentals of DNA and its constituents

mentioned earlier and A, T, C, G, respectively is their abbreviation. Another major nucleic acid, present in the cells, called as ribonucleic acid, abbreviated as RNA, which differs from DNA by replacing thymine with uracil (U). The sequence of the RNA nucleotides allows nucleic acid to encode genetic blueprint of an organism.

7.2 Ancient Deoxyribonucleic Acid (aDNA)

Ancient Deoxyribonucleic Acid, abbreviate as "aDNA", is a nuclear material, which is extracted from long dead biological organisms—plants, animals and humans, that were not preserved for the purpose of DNA isolation. Examples are all archaeological finds of biological origin, fossil remains, museum specimen and any other biological material not preserved for genetic or genomic studies but remained lying under natural environmental conditions over substantial period of time. DNA is very complex molecular structure, which starts degrading just after the death of organism as bacteria initiate the process of decomposition. Ancient DNA material sources are herbarium, museum specimens, bones, teeth, soft tissue and ancient plant seeds or compression fossils (Sutton et al. 1996; Poinar et al. 2001). Mostly ancient DNA samples possess less amount and highly degraded quality of DNA because of negative impact of various environmental factors such as microorganisms, humidity, light and heat, which may be both quantitatively and qualitatively inadequate (Bender et al. 2004). In archaeological, paleontological and anthropological finds, teeth and skeletal material are the best for sampling ancient DNA analysis as well as for forensics identification of missing persons and mass disasters. Biological material buried under debris for a long time has the ability of preservation for long time (Loreille et al. 2007; Seo et al. 2010).

7.3 Timeline of DNA Research

Let us go to the timeline of DNA research, which led us to successfully isolate and interpret the result of hereditary material. About one and a half century ago, in 1869, a young Swiss biologist, Johannes Friedrich Miescher, accidently found various phosphate-rich chemicals from the white blood cell nuclei. Friedrich Miescher got

his MD in 1868 from the University of Basal Switzerland. During his studies in 1865, Miescher worked in the organic chemistry laboratory of Adolf Stecker at Gottingen University, Germany, but due to his illness he had to interrupt his studies. Unfortunately his illness of typhoid fever impaired his hearing ability and he could not pursue medical profession like his father and uncle who were chair of Anatomy at University of Basal. Because of partial deafness, Miescher said good bye to medical profession and decided to pursue his studies towards physiological chemistry and joined University of Tübingen, Germany, at newly established faculty of natural science, under Felix Hoppe-Seyer, a chemist and physiologist, who later became the founder of biochemistry. Miescher was interested to study the chemistry of the nucleus of the white blood cell, but Hoppe-Seyer encouraged him to study neutrophil instead for the simple reason that it is difficult to get sufficient number of lymphocytes while it was known to get sufficient number of neutrophils in the pus, which could be obtained from the bandage of patients. The problem was how to wash cells off the bandage without damaging them. Miescher solved the problem, isolated the nuclei and on analyses of the contents he found a new compound from the nuclei of lymphocytes in 1868.

Young 24-year-old Miescher must have been amazed to find a compound that could not be characterized as a protein or lipid or a carbohydrate. He simply named it as "Nuclein" as it was isolated from the nucleus of blood cell. Interestingly Felix Hoppe-Seyer did not allow Miescher to publish it before his personal verification that took some time and finally Miescher published his findings in 1871. Nuclein remained in obscurity for 20 years till 1889, when Richard Altman, a German Pathologist, found the acidic properties of nuclein and renamed it as "nucleic acid". Today, this nucleic acid is known as deoxyribonucleic acid (DNA).

As is known now, nucleic acid research is the hub of current biomedical research. The reason of anonymity of nucleic acid in the nineteenth century was the publication of Charles Darwin's book in 1859, entitled "*On the Origin of Species*". The book generated hot debates on the theory of evolution, nevertheless by the 1870s, not only the scientific community but also people from the general public had started believing in evolution as a scientific fact. Another scientific investigation which was ignored at that time is the work of Gregor Mendel, "*Versuche über Pflanzenhybriden*" (Experiments on Plant Hybridization), who read it at two meetings (8 February and 8 March 1865) at the "Natural History Society of Brno", and published it in "Vehandlungen des naturforschenden Vereins Brünn" in 1966. However, Mendel work remained ignored till its rediscovery in 1900, by De Vries, Correns and Tsehermak.

7.4 Timeline of Ancient DNA Research

In the early 1980s, researchers started extracting DNA from different ancient sources such as fossil remains, archaeological finds, museum specimens, permafrost ice cores, fossilized faeces, and other uncommon antique biological sources. Such extractions were named "Ancient DNA". In 1984, the field of ancient DNA research

began with genetic analysis of Quaga (*Equas quagga*) DNA, an 140-year-old museum specimen (Higuchi et al. 1984). In 1985, the 2400 years old Egyptian mummies (*Homo sapiens*) were analyzed to understand ancient DNA technology (Pääbo 1985). In 1989, ancient DNA study divulged closest living relatives to extinct Tasmanian tiger (*Thylacinus cynocephalus*). In 1990, ancient DNA technology was introduced to readers by the best-selling novel Jurassic Park. In 1992, the analysis of aDNA from 3550 years old museum species revealed the relationships between moas (*Moa* sp.) and living flightless birds, e.g. ostriches (*Struthio camelus*) and kiwis (*Actinidia deliciosa*). In 1994, DNA of *Mycobacterium tuberculosis was detected in* 1000 years old archaeological specimens from Peruvian mummy (*Homo sapiens*). In 1994, DNA was extracted from 9000 to 50,000 years old Siberian woolly mammoth (*Mammuthus primigenius*) (Hänni et al. 1994). In 1997, mitochon drial DNA of Cheddar Man (*Homo sapiens*) from Britain's 9000-year-old complete skeleton was matched to a local Cheddar school teacher. In 1997, ancient malarial parasite (*Plasmodium falciparum) DNA was detected in a 60-year-old human* (*Homo sapiens*) *skeletal remains*. *In 1998, DNA extracted from* fossilized dung revealed the ecology (diet) of long extinct ground sloths (*Megatherium* sp.).

In 2003, archaeologists studied domestication of corn (*Zea mays tortillas*) from 6300-year-old Mexican wild grass teosinte. In 2003, DNA of ancient plants and animals was detected in 10,000–400,000-year-old Ice cores of Siberia. In 2005, researchers analyzed ancient DNA from 40,000-year-old cave bear (*Ursus spelaeus*) samples and their environment using a novel metagenomic approach. In 2005, DNA of ancient bottle gourds (*Lagenaria siceraria*) was compared with modern Asian and African varieties, and hence the scientists concluded that about 10,000 years ago, Paleoindians (*Homo sapiens*) brought gourds from Asia to Americas (Nicholls 2005). In 2007, researchers isolated oldest DNA (at least 450,000 years old) from southern Greenland cores that were compared with DNA of existing animals and plants. The comparison revealed DNA of conifers (Briggs et al. 2007). In 2008, for the first time, the woolly mammoth's (*Mammulus primigenius*) nuclear genome reported from an extinct animal. In 2010, both nuclear and mitochondrial DNA extracted from fossilized eggs of extinct elephant-birds (*Aepyonus* sp.), owl, ducks, emu (*Dromalius novaehollandiae*) and extinct moas. In 2010, scientists conducted genetic study on five 100-year-old royal Egyptian mummies to find out the relationship of these mummies to King Tutankhamen family and to indicate inherited disorders and infectious diseases in these mummies. The parents of king Tutankhamun were identified from these mummies and genes specific to *Plasmodium falciparum* (malaria pathogen) were found in four mummies. In 2010, scientists published the draft of genome sequence of 40,000-year-old material of Neanderthal man (*Homo neanderthalensis*), which play a role in the understanding of human evolution.

In 2011, DNA of the Black Death pathogen, *Yersinia pestis*, recovered from the 660-year-old skeletal remains of the people who died due to the "Black Death", and buried in East Smithfield Cemetery London. In 2011, fossilized fingertips and two teeth found in Siberia revealed the missing link of human species (*Homo*). In 2012, ancient syphilis pathogen (*Treponema pallidum*) DNA was isolated from

sixteenth- to seventeenth-century-old two newborns deceased affected by syphilis from birth. In 2012, genetic analysis of the 5300 year old frozen mummy of the Tyrolean Iceman, "Otzi", (*Homo sapiens*) carried out. In 2013, comparison of complete genomes of 1000-year-old *Mycobacterium leprae* (leprosy causing bacteria) with modern strains exposed little genomic changes, and indicated that leprosy in the Americas originated in Europe. In 2013, the genome of the oldest known European fungus, *Phytophthora infestans*, was compared with modern strains. The study revealed that *Phytophthora infestans* is an aggressive pathogen, which mutates rapidly in response to fungicides and use of new resistant potato crops. In 2013, the analysis of the 560,000–780,000 years old ancient horse (*Equas* sp.) genome exposed that the common lineage of present day zebra (*Equus zebra*), donkeys (*Equus asunus*) and horses (*Equus Caballus*) arose 4–4.5 million years ago. In 2013, analysis of DNA obtained from ancient dental plaque traced human diet from Neolithic European skeletons through modern times. In 2013, a unique family reunion was celebrated by the Australian and British descendants of Cheddar man in the cave where his 9000-year-old skeleton remains were discovered. In 2014, the analysis of DNA obtained from old bones of Clovis boy linked to American Indians and other native peoples in the Americas. In 2014, the genetic study revealed that Kiwis (*Actinidia deliciosa*) are the closest living relatives of the massive elephant-birds (*Aepyonus* sp.) (Gansauge and Meyer 2014).

In 2016, researchers were able to find out that all human beings are descendant of one common female ancestor called "The Mitochondrial Eve", who lived circa 150,000 years ago in Kenya (Jones 2016). Ancient DNA research has exposed the evolutionary history of earlier man, *Homo denisova*, *Homo neanderthalensis,* etc. and in future aDNA studies will further increase our knowledge about the history of human (Slatkin and Racimo 2016). The field of ancient DNA is an emerging field and currently passing through a phase of exponential growth (Woods et al. 2017). We, human beings, share our DNA genome with other living beings, e.g. chimpanzee (*Pan troglodytes*) [98%], slug (*Arion distictus*) [75%] and banana (*Musa* sp.) [50%] (Haviland et al. 2013).

7.4.1 Development of Ancient DNA Research

To extract and analyze DNA and/or genomes of ancient biological specimens, last decade saw a substantial development of approaches, that opened a new window to clearly understand the genotypes of past organisms, to know the place and time of their extinction and biological variation of species that took place as well as to gain clear insights that how they were evolved from the previous position/stage towards the existing (Orlando et al. 2013).

A massive, methodological transformation has undergone in the study area of ancient DNA technology since 1984 Higuchi and his colleagues initiated its research. The analysis of DNA from ancient biological samples was a laborious process at that time. However, the discovery of polymerase chain reaction (PCR) initiated a rapid progress in the field of ancient DNA research (Mullis and Faloona

1987). The organeller DNA, such as chloroplast DNA in plants, and mitochondrial DNA in animals, is protected by extra membranes and its multiple copies are found in the cells, therefore, the early analyses of the past organisms were focused on organeller DNA. However, the analysis of the nuclear DNA/genomes of ancient material made possible with the discovery of NGS, i.e. next-generation sequencing techniques. Insights into the geographical area and its extents, the types of organisms and taphonomy of ancient organisms were clearly understood with advancement in bioinformatics tools. The advancement in all these technologies provides new information about the identification of fossils and the precise association between past and present organisms, communities and populations, and gives insights into past fossilized life as well as present living organisms from the evolutionary point within a tree of life with the passage of time and environmental changes.

The new era of ancient DNA technology is very important as it provides new avenues in understanding the genomes of fossils and past organisms and even explore about their associations with their existing living organisms. To get new knowledge and breakouts in the field of aDNA technology, the collaborations and combination of multiple fields such as genomics, taphonomy, archaeology, anthropology, biochemistry, bioinformatics and chronology are very critical for making accurate predictions about the future status of existing organisms under future climatic change scenarios.

The increasing collaborations in multiple research fields, improvement of sophisticated sequencing, coupled with use of tools of bioinformatics, development of “New Generation Sequencing” based aDNA data sets, improvement of phylogenetic, denser ancient DNA material sampling around times of extinction and changes in the environment make it possible to use a potpourri of phenotypic and genotypic characters in both fossils and existing organisms. This approach allows us to know as to how populations and ancient species diverge, how diversification patterns develop and how they interact, which pave the way to evolve independently through periods of environmental change (Sjögren et al. 2017).

7.5 Techniques Used for DNA Extraction from Past Organisms

Extraction of DNA from past organisms has obvious impact in archaeology, anthropology and forensic DNA studies, however, the existing protocols used for DNA extraction are far from satisfaction (Kalmár et al. 2000). The extraction process of ancient DNA from the old specimens is very difficult because antique archaeological material is strongly influenced not only by environmental factors but also by microbial contamination, sample age, substrate properties, minimum amount of material to start, as well as presence of PCR inhibitors. From the plant fossils extraction of DNA is more difficult compared to animal fossils, because generally animal old skeletal fossils material possesses well-preserved DNA (Brown and Barnes 2015). In plant fossils, pollen grains are not only well preserved but found in abundance for ancient DNA extraction. The old skeletal remains contain inorganic minerals, which create hurdles during the extraction of DNA, therefore, the buffer

containing EDTA is used for incubation to demineralize these specimens and inactivates DNAses by chelating bivalent Ca^{2+} or Mg^{2+} cations (Hochmeister et al. 1991; Loreille et al. 2007; Barbaro et al. 2011). The extraction of DNA from ancient specimens and its amplification are issues that are important in current DNA technology. Occasionally, the history and condition of the old specimens is not known, however, to get good, reliable and reproducible results a standard and perfect protocol for ancient DNA extraction is needed. A wide range of DNA extraction protocols is available, e.g. alcohol precipitation, spin columns or silica columns. Davoren and his colleagues in 2007, made a comparison of the silica column-based and the phenol/chloroform-based DNA extraction technology for old skeletal materials. Significant results were obtained with silica column-based DNA extraction technique. Loreille et al. (2007) compared the phenol/chloroform and total demineralization DNA extraction methods. Maximum quantity of DNA was obtained with total demineralization extraction protocol using old bone specimens. In 2012, phenol/chloroform, total demineralization and crystal aggregates DNA extraction methods were compared for both fresh and old bone specimens. The phenol/chloroform extraction method was found excellent for fresh bones while the total demineralization extraction method was significantly better for old specimens (Jakubowska et al. 2012). Based on the previous history, silica column and total demineralization extraction methods were merged to develop a new modified silica column-based total demineralization extraction method for DNA extraction from ancient skeletal remains, which were highly degraded (Zar et al. 2013, 2015).

7.6 Applications of Ancient DNA (aDNA) Technology

Ancient DNA technology is an emerging field, which plays a key role in anthropology, archaeology, palaeontology, medicine, agriculture and law enforcement (Papiha et al. 1999; Burger et al. 2000; Pääbo et al. 2004). Some major applications of aDNA are mentioned as follow:

- **Historical Mysteries**: aDNA is used to identify an unknown person from his/her old skeletal material by matching his/her DNA to a known relative.
- **Human Evolution**: aDNA is analyzed to study human evolution.
- **Ancient Diseases**: The old skeletal remains and tissues of dead peoples may still contain pathogens. Therefore, aDNA is used to identify ancient diseases such as Bubonic Plague, Tuberculosis, etc.
- **Living Descendants**: aDNA is analyzed to establish a direct genetic link between a dead and a living individual.
- **Sex Identification**: aDNA is used to distinguish between male and female from the analysis of old skeletal remains or aDNA samples.
- **Animal or Plant Identification**: aDNA can be used to gain insight into prehistoric subsistence practices and the local economy through identification of archaeological/palaeontological remains of plant and animal species.

- **Migration Patterns**: aDNA can be used to ascertain the migration patterns of human populations by comparing the DNA of ancestral remains to that of people living there today.

7.7 Requirements and Challenges of Ancient DNA Research

Genetic markers are required for ancient DNA analysis. These markers must be polymorphic, i.e. marker level differences must be identifiable in the populations that are compared or among the individuals being studied. If differences are not perceived, then it will not be possible to differentiate individuals and about their movements, relationships and evolution. Several calculations are made to know genetic distance, which depends on genetic discrepancies and their number between individuals and/or populations. These genetic discrepancies are used to measure resemblance or dissimilarity between individuals or populations. Often, DNA data interpretation is carried out through comparison of the new data with previous data obtained from previous studies on the same genomic region. For this purpose, access to the previous data stored in the comparative database is crucial. Mitochondrial genome has the most widespread comparative database, because mitochondrial DNA (mtDNA) has been studied more than any other genomic DNA.

Mitochondrial DNA molecule is the preferred organelle for the analysis of ancient DNA because mtDNA persist in multiple copies per cell, having large number of DNA than nuclear DNA, which has single copy and diploid in any given specimen. Nuclear DNA is inherited, through haploid number of chromosomes, from both parents while mitochondrial DNA is inherited only through mitochondria of the maternal genome. Single-copy-chromosome markers (Y-Chromosomal DNA) are inherited through the paternal-lineage. If the analysis of mtDNA is successful, subsequent analysis of nuclear DNA (paternal-lineage or autosomal nuclear loci) may be attempted for ancient DNA studies.

Based on the results, ancient DNA studies are grouped as (a) DNA yield authentic and endogenous, (b) DNA yield contaminated or (c) DNA yield nil. This classification is actually carried out for authenticity of aDNA analysis. As clarified earlier that precise interpretation of ancient DNA data needs a well-defined comparative database for identification and classification of the new data. If no resemblance persists between the new ancient DNA (aDNA) sequences and other earlier reported aDNA sequences, then the interpretation of significance of newly sequenced aDNA becomes extremely difficult. In that case, removal of contamination also becomes very difficult. The uniqueness/variation of a sequence is considered as authentic ancient DNA data, provided all sources of contamination are recognized and eradicated. Some time, the novelty of sequences is so thrilling that contamination looks extremely doubtful, for example, the DNA sequences of *Homo neanderthalensis* were clustered outside the range of the current *Homo sapiens* DNA sequences (Krings et al. 1997, 1999, 2000; Ovchinnikov et al. 2000; Schmitz et al. 2002; Serre et al. 2004).

The analysis of ancient DNA, isolated from past specimens, is an emerging technology which is growing very rapidly. Regardless of pervasive interest in analysis of ancient DNA, the problems occur in ancient DNA analysis. The main problems in ancient DNA analysis are firstly, the contamination of samples with exogenous DNA and secondly, the inability to successfully produce good results from many past specimens. In most cases, majority of samples failed to produce results due to low copy number and highly degraded DNA, and fruitful analysis of past specimens is not achieved. Contamination of DNA samples is another big problem of DNA analysis, therefore, research strategies must be carefully designed to remove and identify all foreign DNA contaminants. Generally, ancient DNA study is extremely specialized and technical. It is fast emerging and quite expensive and needs widespread training. In ancient DNA technology, each step should be carefully planned, DNA samples are needed to be well preserved, and should focus on well-trained physical anthropologists, palaeontologists, archaeologists, geneticists, linguists, related researchers (Kelman and Kelman 1999; Kolman and Tuross 2000; Pääbo et al. 2004; Mulligan 2006).

The two other important challenges in ancient DNA analysis are Informed Consent and Intellectual Property. It is important to decide as to who has the right to give informed consent for ancient DNA research and who owns the information obtained from ancient DNA research are very difficult questions. Since the past specimens cannot provide consent, therefore, we must go to their descendants, but how can we come to know about their descendants, who controls the information about their descendants and past specimens? Will the information created or source of information are authentic, and the way of information sharing is appropriate or according to the descendant communities?

7.8 DNA Evidence Overview

The role of DNA in discrimination or elimination of a suspect in a criminal offense is becoming an important forensic tool with new advancement in its technology. In 1987, DNA was allowed as legal evidence in the United States of America judicial system. DNA as evidence on crime scene can be isolated from saliva, tissue, blood, hair or semen. DNA is also an investigative tool to solve problem of identification of individual human being. Some of the sources which contain DNA as evidence are mentioned below: carpets (analyzed for biological DNA evidence), automobiles and automobile seats (kidnapping, carjacking, rapes, etc.), Guns (for blood and associations made in assaults, battery, homicide case, etc.), Shoes (almost all types of cases), cigarette butts, hairs (battery, armed robbery, sexual assault and homicides), bones (identification of missing people, missing children, war casualties, reconstruction of bodies and mass disaster such as airplane crashes), teeth (identification of missing people, missing children, war casualties, reconstruction of bodies and mass disaster such as airplane crashes), fingernail (sexual assaults, rapes or murders), blood stains (sexual assaults, rapes or murders), semen stains (sexual assaults, rapes or murders), vaginal secretions (sexual assaults, rapes or

murders), chewing gum (saliva), razor blades (missing kids, missing females, etc.), ski masks (saliva and hair, bank robberies, carjacking, rapes or whenever the suspect wants to hide his identity), caps (hair), tooth brush (saliva), stamps and envelops (saliva), ear rings, nasal secretions, ear wax, watches, coffee cups, soft drink cans, beer bottles, etc., shirts (neck area for sweat), tooth picks, eye glasses (nose piece and the back of arm), hearing aids, condoms, saliva stains (threatening letters, or claiming the responsibility of certain crime in a letter), finger rings, bracelets, telephones, cell phones, computer keyboards, steering wheels and dashboards, forks and spoons, date pits and mango stones, gun triggers and magazines. *Locard's principle* states that "Wherever there is contact between two objects, there is a transfer of material between them".

7.9 DNA Database

A DNA database/databank contains profiles of DNA to be used in the analysis of DNA, fingerprinting, genetic genealogy and genetic diseases. DNA databases may be private or public, however, many developed countries have developed their large national database. In 1997, forensic DNA scientists prepared a standardized set of STR markers for identification of *Homo sapiens*. This is called "Combined DNA Index System" abbreviated as "CODIS". For CODIS project as many as 17 STR loci were identified and studied as candidates but only 13 were finally selected as part of the CODIS system. The 13 CODIS STR loci include: "TH01, TPOX, FGA, D21S11, CSF1PO, D7S820, vWA, D18S51, D8S1179, D13S317, D5S818, D16S539, and D3S1358". Among them, D21S11, D18S51 and FGA are the most polymorphic STR markers while TPOX displays the least variation between persons (Chakraborty et al. 1999; Butler 2001).

The benefits of DNA database are as follows:

- In case without any lead, DNA can efficiently and effectively identify suspects by eliminating possible suspects
- Identification of serial offenders by determining the links with different crime scene
- DNA analysis would serve as an investigative tool in tandem with finger prints
- Hard evidence in addition to soft evidence to be presented in the court of law
- DNA database profiles may act as deterrence and create awareness to some habitual offenders from committing crime in future
- Solution of past unsolved cases where DNA evidence was found and available but could not be used because of non-availability of DNA technology but linked with a present suspect offender and matched with the present crime scene
- DNA analysis could not be used in pre-DNA-database era and hence offenders from such cases could not be identified

Jerry Miller Case (DNA That Cleared One Man Implicate Another) The DNA profile obtained from Item(s) M16 and M17 was searched in the Indiana DNA

AMERICAN MORNING | SITUATION ROOM | LOU DOBBS TONIGHT | PAULA ZAHN NOW | LARRY KING LIVE | ANDE

CNN.com

Member Center: Sign In | Register

SEARCH THE WEB CNN.COM

Home World U.S. Weather Business Sports Analysis Politics Law Tech Science Health Entertainment Offbeat Trave

U.S.

Tools: Save | Print | E-mail | Most Popular

DNA exonerates man who spent 25 years in prison

POSTED: 7:40 p.m. EDT, April 23, 2007

Jerry Miller was exonerated of rape charges Monday, after he spent 25 years in prison.

STORY HIGHLIGHTS
- DNA exonerates man who spent 25 years in prison for rape
- Jerry Miller was paroled in March 2006 and lives with family member
- The case is 200th in which a person was convicted then exonerated due to DNA

Adjust font size:

CHICAGO, Illinois (AP) -- A man who spent 25 years in prison for rape was exonerated Monday after a judge threw out his convictions because DNA evidence showed he couldn't have committed the attack.

Jerry Miller smiled and the courtroom erupted into cheers after Cook County Circuit Court Judge Diane G. Cannon read the ruling that cleared him of all charges.

Miller, 48, had been found guilty of rape, robbery, aggravated kidnapping and aggravated battery even though he testified he was at home watching television at the time of the 1981 attack.

He was paroled in March 2006 and now works two jobs and lives with a family member in a Chicago suburb.

"I want to get on with my life, start a life, have a life," Miller said after the hearing. "I'm just thankful for this day."

The Innocence Project, a New York-based group, had persuaded prosecutors last year to conduct DNA tests on a semen sample taken from the rape victim's clothes. Those results excluded Miller as the attacker.

The case is the 200th in the United States in which a person was convicted, then exonerated based on DNA evidence, the group says.

The first exonerations based on DNA testing were in 1989, and in all, the 200

ADVERTISER LINKS
- Fresh Flowers
- Get a New Car Cheap
- Notebook Computers
- Online Book Stores

breast cancer DNA test
Find Answers to Health Questions Surgery, Medicine & Prevention.
RevolutionHealth.com

Identical or Fraternal Twins?
Affordable DNA based zygosity testing for twins.
www.proactivegenetics.com

DNA Testing Centre - Accurate and Fast
By representing many labs, we offer more order options than any DNA paternity...
www.dnatestingcentre.com

Fig. 7.2 Newspaper cutting highlighting Jerry Miller case

Database and was found to be consistent with the convicted offender sample of Jeffery Jones (DOC inmate #974912) (Fig. 7.2).

7.9.1 Importance of Forensic DNA Database Legislation

There are over 50 countries throughout the world that have already passed DNA Database legislation such as Europe (every country in Europe required by EU legislation for data sharing), the United States (every state has passed separate legislation), Australia (every state has passed separate legislation), Canada, Russia, Kenya, Malaysia, etc. No country has ever reduced the scope of DNA database testing done (Doleac 2017). Countries have only expanded database utilization based on:

- The effectiveness of the databases
- The security of the database systems
- NO example of database misuse anywhere in the world

Each country and state has had to address similar issues that arise in the context of Forensic DNA data basing in the context of its own legal, historical and cultural frameworks, for example—destruction of samples in Germany. Each DNA database legislation is different; there are many issues which, after extensive scientific and legal research, have been addressed in very similar ways (Levitt 2007; Wallace et al. 2014). Some of these issues include:

- The protection of Human/Constitutional and other legal rights.
- Sample and profile retention
- The type of sample to be collected and who may collect it

7.9.2 Common Features of DNA Database Authorizing Legislation

- Entry criteria
- Sample collection
- Removal criteria
- Sample retention
- Database access

Legislation is required for:

- Effective utilization of DNA analysis facility of all forensic labs
- Eradicate the confusion of admissibility of DNA evidence
- Ensuring security of DNA profile
- Linking victim and suspect
- Linking suspect/victim with crime scene

References

Barbaro A, Cormaci P, Falcone G (2011) Validation of BTA™ lysis buffer for DNA extraction from challenged forensic samples. Forensic Sci Int Genet Suppl Ser 3:e61–e62

Bender K, Farfán MXAJ, Schneider PM (2004) Preparation of degraded human DNA under controlled conditions. Forensic Sci Int 139:135–140

Briggs AW, Stenzel U, Johnson PL, Green RE, Kelso J, Prüfer K, Meyer M, Krause J, Ronan MT, Lachmann M (2007) Patterns of damage in genomic DNA sequences from a Neandertal. Proc Natl Acad Sci 104:14616–14621

Brown TA, Barnes IM (2015) The current and future applications of ancient DNA in Quaternary science. J Quat Sci 30:144–153

Burger J, Hummel S, Pfeiffer I, Herrmann B (2000) Palaeogenetic analysis of (pre) historic artifacts and its significance for anthropology. Anthropol Anz 58:69–76

Butler JM (2001) Forensic DNA typing: biology & technology behind STR markers. Academic Press, New York, NY

Chakraborty R, Stivers DN, Su B, Zhong Y, Budowle B (1999) The utility of short tandem repeat loci beyond human identification: implications for development of new DNA typing systems. Electrophoresis 20(8):1682–1696

Daniel R, Walsh SJ (2006) The continuing evolution of forensic DNA profiling-from STRS to SNPS. Aust J Forensic Sci 38:59–74
Doleac JL (2017) The effects of DNA databases on crime. Am Econ J Appl Econ 9(1):165–201
Feuk L, Carson AR, Scherer SW (2006) Structural variation in the human genome. Nat Rev Genet 7:85–97
Gansauge M-T, Meyer M (2014) Selective enrichment of damaged DNA molecules for ancient genome sequencing. Genome Res 24:1543–1549
Hänni C, Laudet V, Coll J, Stehelin D (1994) Sequence variant from an Egyptian mummy. Genomics 22:487–489
Haviland WA, Prins HE, Mcbride B, Walrath D (2013) Cultural anthropology: the human challenge. Cengage Learning, Boston, MA
Higuchi R, Bowman B, Freiberger M, Ryder OA, Wilson AC (1984) DNA sequences from the quagga, an extinct member of the horse family. Nature 312:282
Hochmeister MN, Budowle B, Borer UV, Eggmann U, Comey CT, Dirnhofer R (1991) Typing of deoxyribonucleic acid (DNA) extracted from compact bone from human remains. J Forensic Sci 36:1649–1661
Jakubowska J, Maciejewska A, Pawłowski R (2012) Comparison of three methods of DNA extraction from human bones with different degrees of degradation. Int J Legal Med 126 (1):173–178
Jones M (2016) Unlocking the past: how archaeologists are rewriting human history with ancient DNA. Arcade Publishing, New York, NY
Kalmár T, Bachrati CZ, Marcsik A, Raskó I (2000) A simple and efficient method for PCR amplifiable DNA extraction from ancient bones. Nucleic Acids Res 28:e67–e67
Kelman LM, Kelman Z (1999) The use of ancient DNA in paleontological studies. J Vertebr Paleontol 19:8–20
Kolman CJ, Tuross N (2000) Ancient DNA analysis of human populations. Am J Phys Anthropol 111:5–23
Krings M, Stone A, Schmitz RW, Krainitzki H, Stoneking M, Pääbo S (1997) Neandertal DNA sequences and the origin of modern humans. Cell 90:19–30
Krings M, Geisert H, Schmitz RW, Krainitzki H, Pääbo S (1999) DNA sequence of the mitochondrial hypervariable region II from the Neandertal type specimen. Proc Natl Acad Sci 96:5581–5585
Krings M, Capelli C, Tschentscher F, Geisert H, Meyer S, Von Haeseler A, Grossschmidt K, Possnert G, Paunovic M, Pääbo S (2000) A view of Neandertal genetic diversity. Nat Genet 26:144–146
Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, Fitzhugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, Levine R, Mcewan P, Mckernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, Mcmurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, Mcpherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M et al (2001) Initial sequencing and analysis of the human genome. Nature 409:860–921
Levitt M (2007) Forensic databases: benefits and ethical and social costs. Br Med Bull 83:235–248
Loreille OM, Diegoli TM, Irwin JA, Coble MD, Parsons TJ (2007) High efficiency DNA extraction from bone by total demineralization. Forensic Sci Int Genet 1:191–195
Luftig MA, Richey S (2001) DNA and forensic science. N Engl Law Rev 35(3):609–613

Mulligan CJ (2006) Anthropological applications of ancient DNA: problems and prospects. Am Antiq 71:365–380

Mullis KB, Faloona FA (1987) Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. Methods Enzymol 155:335–350

Nicholls H (2005) Ancient DNA comes of age. PLoS Biol 3:e56

Orlando L, Ginolhac A, Zhang G, Froese D, Albrechtsen A, Stiller M, Schubert M, Cappellini E, Petersen B, Moltke I (2013) Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. Nature 499:74–78

Ovchinnikov IV, Götherström A, Romanova GP, Kharitonov VM, Lidén K, Goodwin W (2000) Molecular analysis of Neanderthal DNA from the northern Caucasus. Nature 404:490–493

Pääbo S (1985) Molecular cloning of ancient Egyptian mummy DNA. Nature 314:644–645

Pääbo S, Poinar H, Serre D, Jaenicke-Després V, Hebler J, Rohland N, Kuch M, Krause J, Vigilant L, Hofreiter M (2004) Genetic analyses from ancient DNA. Annu Rev Genet 38:645–679

Papiha SS, Deka R, Chakraborty R (1999) Genomic diversity: applications in human population genetics. In: Proceedings of a Symposium on Molecular Anthropology in the Twenty-First Century, Held During the 14th International Congress of the Association of Anthropological and Ethnological Sciences, Held July 26-August 1, 1998, in Williamsburg, Virginia. Springer Science & Business Media, Berlin

Poinar HN, Kuch M, Sobolik KD, Barnes I, Stankiewicz AB, Kuder T, Spaulding WG, Bryant VM, Cooper A, Pääbo S (2001) A molecular analysis of dietary diversity for three archaic Native Americans. Proc Natl Acad Sci 98:4317–4322

Schmitz RW, Serre D, Bonani G, Feine S, Hillgruber F, Krainitzki H, Pääbo S, Smith FH (2002) The Neandertal type site revisited: interdisciplinary investigations of skeletal remains from the Neander Valley, Germany. Proc Natl Acad Sci 99:13342–13347

Seo SB, Zhang A, Kim HY, Yi JA, Lee HY, Shin DH, Lee SD (2010) Efficiency of total demineralization and ion-exchange column for DNA extraction from bone. Am J Phys Anthropol 141:158–162

Serre D, Langaney A, Chech M, Teschler-Nicola M, Paunovic M, Mennecier P, Hofreiter M, Possnert G, Pääbo S (2004) No evidence of Neandertal mtDNA contribution to early modern humans. PLoS Biol 2:e57

Sjögren P, Edwards ME, Gielly L, Langdon CT, Croudace IW, Merkel MKF, Fonville T, Alsos IG (2017) Lake sedimentary DNA accurately records 20th Century introductions of exotic conifers in Scotland. New Phytol 213:929–941

Slatkin M, Racimo F (2016) Ancient DNA and human history. Proc Natl Acad Sci 113:6380–6387

Sutton M, Malik M, Ogram A (1996) Experiments on the determination of gender from coprolites by DNA analysis. J Archaeol Sci 23:263–267

Wallace HM, Jackson AR, Gruber J, Thibedeau AD (2014) Forensic DNA databases–ethical and legal standards: a global review. Egyptian J Forensic Sci 4(3):57–63

Woods R, Marr MM, Brace S, Barnes (2017) The small and the dead: a review of ancient DNA studies analysing micromammal species. Genes (Basel) 8:8(11)

Zar MS, Shahid AA, Shahzad MS, Shin KJ, Lee HY, Israr M, Kim EH, Rahman ZU, Husnain T (2013) Forensic DNA typing of old skeletal remains using AmpFlSTR®Identifiler® PCR amplification kit. J Forensic Res 5(1):211–216

Zar MS, Shahid AA, Shahzad MS, Shin KJ, Lee HY, Israr M, Husnain T (2015) Comparative study of STR loci for typing old skeletal remains with modified protocols of AmpFlSTR® Identifiler® and AmpFlSTR® MiniFiler™ STR Kits. Aust J Forensic Sci 47(2):200–223

8 Analyses of Second World War Skeletal Remains Using a Forensic Approach

Irena Zupanič Pajnič

Abstract

This chapter presents analyses of skeletal remains from World War II using a forensic approach. Aged bones are challenging samples of biological material for DNA typing because their DNA content is very low and greatly degraded. The exceptional risk of contamination and the presence of inhibitors further limit the success of DNA typing. Because the DNA content is so limited, aged skeletons are exposed to contamination by people involved in excavation, anthropological analyses, and genetic testing. To prevent and track potential contamination by contemporary DNA, a number of standard precautions are used and they are described. The composition of bones and teeth and their degradation process is discussed. In addition to morphological structure, special attention is paid to factors affecting the preservation of DNA in old bones and teeth. Based on the literature reviewed and some analyses performed, the chapter summarizes which skeletal elements are most suitable for investigating World War II skeletal remains. It discusses how to clean and grind bone and tooth samples, how DNA can be extracted from the powder obtained, and how DNA quality and quantity can be determined for extracts using real-time quantification. The genetic markers most frequently examined in aged DNA and the advantages of new, high-performing sequencing techniques for the development and study of aged DNA are described. Using innovative methods that may help in retrieving higher-quality and increased data makes it possible to investigate more degraded DNA. Storing samples is especially important in laboratories engaged in forensic genetics. Efficient long-term bone storage is necessary to guarantee sample stability across time so that new markers as well as new technologies can be used for future retesting. The results of some aged bone sample storage studies are

I. Zupanič Pajnič (✉)
Institute of Forensic Medicine, Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia
e-mail: irena.zupanic@mf.uni-lj.si

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_8

presented. The chapter concludes with a presentation of World War II victim identifications performed in Slovenia.

Keywords

Disaster victim identification · World War II skeletal remains · Degraded DNA · DNA typing

8.1 Introduction

In forensic investigations, unknown remains of human bones must often be identified. If the identity of the remains cannot be confirmed by standard forensic methods, it is possible to use the teeth and bones to perform molecular genetic identification. This has proven to be a very successful method for identifying missing persons, victims of mass disasters, and victims of wartime conflicts (Biesecker et al. 2005; Prinz et al. 2007; Zupanič Pajnič 2013b; Hartman et al. 2011; Montellius and Lindblom 2012; Ossowski et al. 2013; Morild et al. 2015). Among biological samples, skeletal remains are some of the most challenging for DNA typing because the condition of the skeletal remains is often not conducive for recovering DNA (Burger et al. 1999). Aged bones have a very low content of DNA that is badly degraded, and the success of DNA typing is further limited by the presence of inhibitors of PCR (Zietkiewicz et al. 2012; Irwin et al. 2012). A high contamination risk also constrains successful DNA typing (Anderung et al. 2008).

8.2 Composition of Bones and Teeth and Their Degradation

The amount of DNA in old skeletal remains and its quality depends on the preservation of tissues that are subjected to various environmental factors after death (Iwamura et al. 2004). In a living cell, special mechanisms take care of repairing damaged DNA. DNA repair mechanisms do not function after the death of the organism because the cellular balance is broken. Due to exposure to bacteria, fungi, and nucleases, DNA damage continues to occur, resulting in the survival of very fragmented DNA (Baum et al. 2013). Understanding the composition of bones and teeth and their degradation process is a key factor in determining DNA content and plays an important role in selecting suitable samples for obtaining endogenous DNA.

Teeth are anatomically divided into the crown, neck, and root. The outer layer of the crown is covered with enamel, which is almost entirely of mineral origin and does not contain cells. The tooth root is covered with cementum, which is a mineralized tissue consisting of hydroxyapatite, collagen, and other non-collagen proteins (Higgins and Austin 2013). Cell-free cementum is located in the neck of the tooth root. Cementum containing cells is located in the apical part of the tooth root and is a good source of DNA (Mansour et al. 2018). In the dental neck, enamel and

cementum come into contact. Below both of them, dentine protects the tooth pulp. The dentine and pulp comprise the majority of the tooth and, unlike the enamel, are rich in cells (mostly odontoblasts and fibroblasts). The dentine consists of hydroxyapatite, collagen type I, and water (Higgins and Austin 2013). The quantity of DNA recoverable from a particular tooth and from one group of teeth to another greatly varies; the amount of DNA isolated and its quality are also affected by dental pathologies, dental procedures that were carried out, how much time has gone by since the tooth was extracted until the DNA was isolated, and the age of the donor (Schwartz et al. 1991). Dental pulp is a good source of DNA. The quantity of DNA is related to the volume of the tooth pulp and the kind of tooth. Teeth with larger pulp and multi-root teeth are the best source of DNA because they contain many pulp cells and have more tooth cementum compared to single-root teeth. According to the recommendations, molars contain the most DNA and are the best for DNA analyses, followed by premolars (Higgins and Austin 2013). The best teeth for isolating DNA are as follows: first of all, molars, premolars, canines, and incisors that have not undergone endodontic treatment, and then molars, premolars, canines, and incisors that have been treated.

Bones are macroscopically divided into compact and spongy bone. Compact bone can be found in the outer part of the ends of the long bones (epiphysis), the outer part of the flat bones, and the middle part of the long bones (diaphysis). Spongy bone can be found inside the flat bones and inside of the ends of the long bones. Because of the fragility of spongy bone, it needs an outer protective layer of compact bone (Camposa et al. 2012). Microscopically, bone consists of cells and an extracellular matrix made up of an organic part and an inorganic part. The organic part mostly consists of type I collagen plus some other proteins and glycoproteins, and the inorganic part mostly consists of hydroxyapatite composed of calcium and phosphate ions (Camposa et al. 2012). Following death, the DNA is maintained much better in teeth and bones than it is in soft tissue (Iwamura et al. 2004; Camposa et al. 2012; Kendall et al. 2018). The slow degradation of DNA and the protection against enzymatic processes in the skeletal remains is influenced by the binding of negatively charged phosphate groups to hydroxyapatite hydroxyl groups, and the DNA molecule is bound to collagen as well (Camposa et al. 2012; Coulson-Thomas et al. 2015). Collagen and hydroxyapatite together form a tight structure that, due to small pores, prevents collagenases of microorganisms from breaking into the structure (Turner-Walker 2008). The size of the pores and their interconnection determine how water, microorganisms, and other particles pass through and out of the structure (Kendall et al. 2018). The environment that skeletal remains are located in from death until they are found greatly impacts the degradation of bones and teeth. The decay of bone cells and the soft tissue in the blood vessels make the bone more porous. Bone porosity determines the speed and manner in which environmental factors affect bone changes. Fungi and bacteria present in the soil and cyanobacteria present in the water can easily break into the bone tissues due to increased porosity with time, making the bone tissue even more porous (Camposa et al. 2012). In contrast, the environmental impact can also reduce porosity due to permineralization, which gradually leads to fossilization, making the bone tissue

better preserved (Kendall et al. 2018). The pores in the teeth are smaller than in bone, and therefore the DNA in teeth is preserved better than in bones. Thus, it is possible for DNA to persist for centuries or even millennia (Turner-Walker 2008).

8.3 Factors Affecting the Preservation of DNA in Skeletal Remains

Analysis of aged skeletal remains found at various locations proves that the preservation of DNA is affected not only by the postmortem interval, but primarily by the environment the remains are subjected to. Factors affecting DNA preservation in skeletal remains include the soil's humidity, temperature, geochemical properties, and pH as well as microorganisms in the soil (Höss et al. 1996; Poinar et al. 1996; Putkonen et al. 2010; Higgins and Austin 2013). The most important factors for preserving DNA are the soil temperature and its humidity, and environments that are very stable with minimal yearly temperature or humidity fluctuation are favorable for preserving DNA. Wet warm environments dramatically reduce the amount of DNA because of extensive damage and fragmentation (Smith et al. 2003). The presence of water, the pH of the environment, and mineralized tissue's porosity are factors that influence the solubility of hydroxyapatite. Water helps dissolve mineral ions, and a decrease in pH increases solubility. Alternating dry and humid environments and environments with continual water flow are more harmful to DNA preservation than a continuously wet environment (Turner-Walker 2008; Higgins and Austin 2013; Kendall et al. 2018). There also exist factors particular to individuals affecting the preservation of DNA; these include race, age, sex, age, and the type of skeletal elements (Camposa et al. 2012; Romanini et al. 2012; Mundorff and Davoren 2014).

8.4 Sampling Strategy and DNA Typing of Skeletal Remains

The assessment of bone and tooth suitability for genetic analysis is difficult, and investigations are costly and often unsuccessful. Consequently, much recent research focuses on finding the most appropriate skeletal elements that would allow the acquisition of well-preserved DNA for genetic typing (Hansen et al. 2017; Pilli et al. 2018). Several studies of using skeletal remains for nuclear DNA typing determined that teeth yield the best typing results, followed by femurs and tibias (Edson et al. 2004; Miloš et al. 2007; Misner et al. 2009). These bones are not always available for use in sampling and genetic testing. In the study to identify victims of the World Trade Center attack, DNA was extracted from small bones, including patellae, foot phalanges, and metatarsals; these produced rates that were comparable to both femurs and tibias (Mundorff et al. 2009). Despite the wide spread conviction that dense (i.e., cortical) bone contains DNA that is better preserved than in spongy (i.e., cancellous) bone, a few researchers have reported that commonly overlooked small skeletal types of the feet and hands have been even more useful than weight-bearing long bones. Among them, Mundorff et al. (2009) and Mundorff

and Davoren (2014) found that metatarsals, metacarpals, and phalanges as well as patellae were very similar to or even better than the femurs and tibias in the DNA that they yield. It is easy to sample such bones with a disposable scalpel, thereby reducing potential contamination of DNA. Recent studies of teeth show that, for genetic typing of skeletal remains, it is most appropriate to sample tooth cementum in the region of the tooth root and not whole teeth (Mansour et al. 2018). A study performed by Pilli et al. (2018) compares various skeletal elements with the middle section of the petrous bone (the temporal bone), which is among the toughest human bones (Hansen et al. 2017). Pinhasi et al. (2015) determined that the petrous bone is the most suitable skeletal part for sampling ancient human skeletons. For identifying which skeletal elements yield superior results in DNA preservation in 70-year-old skeletons, different skeletal element types were sampled from three World War II skeletons. The goal was to optimize the sampling process for identifying victims of World War II (Obal et al. 2019). With the help of measurements such as the quantity and quality of DNA (measured using qPCR, in which short and long fragments were targeted and degradation rates determined), and autosomal STR typing success, an effort was made to determine the best kinds of skeletal elements for identifying the identity of the victims. The best parameters for assessing skeletal types that show superior preservation of DNA were the DNA quantity and number of STR loci that were typed successfully. Among 56 different elements, 15 elements in all three skeletons yielded a full profile. Among the skeletal types containing the most DNA on average in all three skeletons studied were mostly bones of the hand and the feet and metatarsal and metacarpal bones exhibited the highest yield. Among 15 bones yielding full genetic profiles in the three skeletons, 14 also contained the greatest average DNA amounts except for tibias, for which the average yield of DNA did not correspond to the success of profiling because the amount was lower (Obal et al. 2019). Our second study was performed on ancient bones. The aim was to establish if acquiring sufficient DNA quantities and then profiling the autosomal STR markers would succeed for small bones of the feet (i.e., metatarsal bones) and of the hands (i.e., phalanges and metacarpal bones) from ancient skeletons in comparison to skeletal types suggested by current recommendations (i.e., petrous portion of temporal bones, femurs, and teeth). Therefore, 48 samples of bone from eight skeletons (six eighteenth-century skeletons and two third-century skeletons) were acquired from five Slovenian archeological locations. The study sampled six skeletal types from each skeleton (molar, temporal bone, femur, metacarpal bone, proximal phalanx from the hand, and metatarsal bone). Careful precautions were adhered to for preventing any contamination. The greatest yields were identified in temporal bones, and the lowest were in femurs. The degree of success for STR typing was evaluated based on the number of loci that were successfully typed, and the study confirmed a robust correlation between the quantity of DNA extracted and the degree of success for typing STRs. The results indicate that it would be appropriate to involve metacarpal and metatarsal bones in the sampling strategies for identification even in the study of ancient bones (Geršak et al. 2019). Based on recent studies performed by various research groups (including this author's research on World War II and ancient skeletons), the present recommendations for preferentially testing long bones

taken from the legs must be re-evaluated, and in the future there should be changes to the sampling strategy used by laboratories that type bone samples.

Recent forensic studies have focused on identifying which skeletal types offer superior DNA preservation, but there has been a lack of focus on measuring variation within particular bone type (Barta et al. 2014; Higgins et al. 2015; Alberti et al. 2018). As early as 2004, Pääbo noticed that yields of aged DNA might vary among extracts of different samples taken from the same bone, even when the same extraction method is used (Pääbo et al. 2004). It is necessary to focus sampling efforts on sections of bone that will result in maximum recovery of genetic material, and so it is of great interest to explore the variation in DNA content within a bone in the future. DNA content variation within a bone was explored by measuring nuclear DNA quantity and quality (using qPCR in which short and long fragments were targeted and degradation rates determined) in World War II metatarsal and metacarpal bones (data not published yet). To exclude the influence of taphonomic issues (postmortem interval and the environment that the remains were subjected to), 213 bones from a single World War II mass grave were examined to ensure that the skeletons had decomposed for the same duration and under the same conditions. From each bone, DNA was extracted from the compact diaphysis and from spongy epiphyses. A statistical analysis indicated that DNA yields differed significantly in extracts from the diaphysis and extracts obtained from the epiphyses of the same bones in metacarpals and metatarsals. A higher difference in DNA yield between the diaphysis and epiphyses was found in metacarpals in comparison to metatarsals. On average, 12 times more DNA was obtained from metatarsal epiphyses than the diaphysis, and 26 times more from metacarpal epiphyses than the diaphysis. It was determined that the best place to sample from within metatarsal and metacarpal bones from World War II is spongy epiphyses (data not published yet).

8.4.1 Contamination Monitoring and Its Prevention

Preventing and detecting contamination by contemporary or modern DNA is essential in forensic DNA examinations of aged remains. Modern DNA contamination can occur during excavation, as well as anthropological and molecular genetic investigations. Because of the small DNA quantities extracted from old teeth and bones, it is difficult to distinguish between endogenous DNA and much more prevalent modern DNA. Studies of old skeletal remains in forensic laboratories should be guided by experience in analysis of ancient DNA, which usually deals with significantly older samples, for which problems caused by inhibition, degradation, and contamination consequently usually increase (Pääbo et al. 2004; Gilbert et al. 2005; Rohland and Hofreiter 2007). Recommendations for preventing contamination and criteria that confirm the authenticity of the genetic profiles are therefore used. In order to ascertain the authentic character of bone and tooth genetic profiles, an elimination database is always created that contains the genetic profiles of every person that came in contact with the skeletal remains. To prevent contamination in the laboratory as much as possible, it is necessary to handle samples and carry out

DNA extraction in dedicated laboratory premises where post-PCR work has never been carried out (Pääbo 1990). The laboratory may further be equipped with Hepa-filtered hoods to ensure that no DNA has entered the facility. All of the work for extraction should be carried out using protective garments, the workspace should be cleaned using bleach, and UV lights should be used to irradiate it (Pääbo 1989). Testing of an extraction-negative control and PCR-negative controls must be carried out in parallel with bone DNA typing to detect any contamination that has been introduced from reagents at the time of the extraction and amplification procedure. Excavators, anthropologists, and laboratory researchers (included in the elimination database) should be typed for the same genetic markers whenever possible, and their genetic profiles should not match the profiles obtained from bones. Each specimen should be used to prepare multiple extracts, and these should produce identical DNA profiles (Pääbo et al. 2004). It is necessary to quantify the number of amplifiable DNA molecules. An inverse correlation must exist between the size of the amplification product and amplification efficiency, which reflects damage and degradation in the aged DNA template (Pääbo et al. 2004).

To reduce the contamination possibility to a minimum at the author's laboratory, the extraction procedure is performed in a room exclusively designed to process World War II remains. Mechanical cleaning is carried out in a closed cytostatic C-(MaxPro)^3-130 (IskraPio) safety cabinet. To reduce cross-contamination to a minimum, a maximum of eight bone or tooth samples are processed during each extraction batch, and a blank control is included with each batch. DNA quantification is performed to test whether an inverse correlation exists between the size of the amplification product and the amplification efficiency. Cleaning in detergent, ethanol, and bi-distilled water, plus UV-light irradiation are carried out to minimize contamination of surfaces from prior handling. The extraction process includes extraction-negative controls to check their cleanliness. Cleanliness of amplification-negative controls is checked as well. Analyses of tooth and bone sample are separated physically and time-wise from sample analyses for the elimination database and family reference samples. Duplicate analysis is performed for all bone and tooth samples, extraction-negative controls, and negative template controls. DNA is extracted from teeth and bones a minimum of two times or from different skeletal types from the same person to verify the genotyping results, and results that are reproducible are utilized for interpretation. No contamination of old bone and tooth DNA samples was detected in the laboratory because strict precautions to prevent contamination are followed.

8.4.2 Bone and Tooth Sample Preparation

When directly handled and washed, the sample surface seems to be contaminated, as well as (to various degrees, depending on porosity and preservation), bone and tooth interiors (Gilbert et al. 2005; Salamon et al. 2005; Sampierto et al. 2006). Because of this, bone and tooth samples are purified and decontaminated in the laboratory. Samples of bone are cleaned chemically and mechanically, and they are irradiated

with UV light, and the samples from teeth are chemically cleaned and then subjected to UV-light irradiation prior to being ground into a powder. No procedures for removing contamination occurring during excavation or storage of remains, or during collection are 100% efficient, but cleaning may enhance the ratio between endogenous DNA and contaminating DNA, and it may help reduce the quantity of inhibitors that are introduced into the extraction (Rohland and Hofreiter 2007). The samples of bone are cleaned by physically removing the surface with a rotary sanding device, and they washed successively in detergent, sterile bi-distilled water, and ethanol. An approximately 1–3 mm layer is stripped from the bone surface through sanding to reduce any contamination caused by prior handling. Skeletal remains are cleaned in a MC 3 (IskraPio) closed microbiological safety cabinet located in premises exclusively designed for handling old skeletal remains. Between processing each sample, all of the tools used to drill, cut, and grind the bones are cleaned. They are washed with water, bleach (6% sodium hypochlorite) or DNA Away™ (Molecular BioProducts) and sterile bi-distilled water (Sartorius-Stedim Biotech or Millipore), followed by 80% ethanol. After this, all of the material is sterilized and undergoes UV-irradiation a minimum of overnight, or as long as 72 h. Warming bones may cause endogenous DNA to degrade (Alonso et al. 2001). To keep the bones from warming while drilling and cutting take place, abrasion and cutting are performed at a lower speed, and the bone is frequently cooled in liquid nitrogen. A factor that affects DNA extract quality is the quality of bone powder used for the extraction procedure. Very fine bone powder must be obtained for extraction of a sufficient amount of DNA from skeletal remains. Very fine-grained powder yields improved and more rapid demineralization, and the generation of fine powder maximizes the sample surface area for eventual contact with the chelating solution (Rohland and Hofreiter 2007). Rohland and Hofreiter (2007) compared samples ground into a fine powder with a freezer mill versus samples that were ground into pieces measuring a few millimeters in diameter with a mortar and pestle; they determined that higher yields were obtained from the fine powder. The performance of a Bead Beater MillMix 20 tissue homogenizer (Tehtnica–Domel, Železniki, Slovenia) was tested for grinding the human bone samples on 60 World War II femurs and tibias, and it was found to be very successful in the recovery of high-quality powder from skeletal remains from World War II (Zupanič Pajnič 2017). Grinding is performed in grinding vials made of metal using metal balls that are cooled with liquid nitrogen before use. Samples of teeth and bones are cooled with liquid nitrogen before grinding as well to avoid overheating during powdering. Grinding is performed at 30 Hz for 1 min. Pulverization of skeletal remains is performed in premises exclusively designed for handling old skeletal remains. The grinding vials must be fully cleaned before they are used again. Between samples, they are washed with water, and then bleach (6% sodium hypochlorite) or DNA Away™ (Molecular BioProducts), followed by sterile bi-distilled water (Sartorius-Stedim Biotech or Millipore), and finally 80% ethanol. In the end, the vials undergo sterilization and UV irradiation a minimum of overnight or as long as 72 h, plus another 30 min before they are used. There must be an adequate number of grinding vials for preparing more than one sample each day.

8.4.3 Extraction of DNA

Harmful effects due to different environmental factors impact the ability to obtain uncontaminated intact DNA from aged bones, and so it is very important to take proper steps to guarantee that the highest possible DNA yields are obtained when analysing poor-condition skeletal remains. The DNA extraction method that is used as key for the quantity and quality of the DNA acquired, and it has a great effect on STR typing success (Putkonen et al. 2010; Irwin et al. 2012). Optimization of extraction methods is necessary for obtaining DNA from old skeletons. Highly efficient extraction methods are the basis for studying and obtaining any genetic data from old bones in forensic investigations. The most recent studies indicate that total demineralization is the optimum method for extracting DNA from old bones (Jakubowska et al. 2012; Amory et al. 2012) because total demineralization greatly increases the share of full profiles, and this correlates with better quality of DNA. The DNA extraction method that was developed at the author's laboratory for acquiring DNA of high quality from remains of World War II skeletons was developed based on a total demineralization process, and it has proved effective from as little as 0.5 g of powder from tooth or bone (Zupanič Pajnič et al. 2016a; Zupanič Pajnič 2016). Analysis was performed on 111 teeth and bones taken from World War II mass graves for evaluation of this method (Zupanič Pajnič 2011), as well as on 54 samples of skeletal remains from World War II (Zupanc et al. 2013). No undigested tooth of bone powder remained after the demineralization and lysis stages because complete decalcification was applied with ethylenediaminetetraacetic acid (Na_2EDTA). Na_2EDTA is a powerful chelator that can bind metallic ions like calcium in the tooth or bone powder, making it possible to remove it. High quantities of Na_2EDTA are needed to dissolve some of the hydroxyapatite matrix that is specific to tooth and bone samples (Rohland and Hofreiter 2007). To achieve complete demineralization, 15 mL of 0.5 M Na_2EDTA per gram of tooth or bone powder is required. In theory, this amount of Na_2EDTA can only bind the quantity of calcium in 1 g of tooth or bone powder (Loreille et al. 2007). To totally demineralize 0.5 g of tooth of bone powder, 10 mL of Na_2EDTA was used in the extraction protocol. After decalcification, an extraction buffer, proteinase K, and DTT are mixed in to the precipitate, and this is incubated for 2–3 h at 56 °C. Alongside calcium chloride, humic acids are general PCR inhibitors that are often encountered in samples bone taken from soil (Krenke et al. 2008). When PCR inhibitors are present, this can create imbalances or the loss of peaks at particular STR loci, and this may produce a genetic profile that is incomplete. The purification of DNA using an organic extraction method (commonly used in forensic laboratories) is performed using chloroform and phenol. They are both dangerous, and processing using them should always be carried out with a vented fume hood. Because they are toxic, using other efficient methods to purify DNA is much safer. The purification of DNA using magnetic beads and robotic devices was included in the extraction protocol (Zupanič Pajnič et al. 2016a; Zupanič Pajnič 2016) and no inhibition was detected. High efficiency of magnetic particles in DNA purification and PCR inhibitor removal was also confirmed in some other studies that were not conducted on skeletal remains but

on casework samples (Nagy et al. 2005; Montpetit et al. 2005; Kishore et al. 2006; Valgren et al. 2008). The purification procedure using automated devices does not use aggressive organic solvents such as phenol or chloroform, devices are easy to use, and it is not necessary to prepare reagents (all of the reagents except DTT are available in the kits). The purification procedure is automated and it requires only 20–30 min to complete. The entire process of extraction is performed in a large filter tip for single use, and the remainder of the extraction reagents is also safely put in a cartridge for one-time use. As a result, manual pipetting is not necessary and the extraction process is less prone to contamination. Purification based on magnetic particles can be modified for robotic machines from different suppliers or can be carried out by hand. DNA from skeletal material from World War II was successfully purified using Biorobot EZ1 (Qiagen) and the AutoMate Express Instrument (TFS). It is believed that the established extraction methods (Zupanič Pajnič et al. 2016a, Zupanič Pajnič 2016) will make a contribution to the options for automated DNA extraction from skeletal remains, thus streamlining procedures for extracting high-quality DNA from skeletal material in forensic laboratories. In addition, commercial kits for automated DNA extraction from bone samples will make possible test and result standardization in forensic laboratories, will increase tooth and bone sample throughput, and will minimize the possibility of human error resulting in sample mixing.

8.4.4 Real-Time Quantification

In compromised bone samples, it is vital to examine the quality and quantity of human DNA that is available for genetic analysis. To obtain this information, the DNA extracts from the bone samples should be quantified by multiplex real-time or quantitative PCR (qPCR). New, highly informative multiplex qPCR kits that quantify total human nuclear DNA, male DNA, the level of DNA degradation, and PCR inhibitor presence all at the same time are used by the author's laboratory to determine the DNA quality and quantity from old tooth and bone samples. The multicopy targets that new kits use offer a much lower detection level than the single-copy assays used in the past. Samples of World War II skeletal remains from mass graves were subjected to extreme environmental conditions, which resulted in extensive degradation of DNA. For assessing the extent of degradation of DNA, two regions with different lengths from the same autosomal multicopy locus are amplified using new qPCR kits. The ratio of the relative amounts of short autosomal (Auto) and long degradation (Deg) amplicons provides an evaluation of DNA quality through a quantitative estimate of the extent of DNA degradation (i.e., the [Auto]/[Deg] ratio) for each sample. It was determined that the DNA extracted from the skeletons recovered from World War II Slovenian mass graves is highly degraded with an [Auto]/[Deg] ratio or degradation index much higher than 2 (Zupanič Pajnič et al. 2017). In addition to determining the concentration of DNA, new kits also detect the Internal Positive Control (IPC), which is part of each amplification reaction, and its amplification is utilized for detecting inhibitors

found in the sample. A quantification screening method that is both effective and able to identify DNA samples that do not result in useful short tandem repeat (STR) profiles is greatly desired in forensic genetic laboratories because STR typing has a strong impact on the costs and time connected with DNA profile development. Preservation of the sample to allow single nucleotide polymorphism (SNP) or mitochondrial DNA (mtDNA) analyses is also essential. The great sensitivity of the new qPCR assays could be applied to identify DNA samples with quantities below the STR reaction sensitivity and thus for avoiding analysing samples that are less likely to be amplified, resulting in less expenditure of money and time. When testing the qPCR PowerQuant assay (Promega) on 60 World War II bone samples, it was wondered whether the kit is capable of predicting downstream STR typing success and whether it could be used to screen for autosomal STR typing success. STR profiles that were full or partial but useful were only produced from bone extracts in which short autosomal (Auto) PowerQuant targets and long degradation (Deg) ones were detected. It was determined that, after PowerQuant quantification, STR typing of old bones should be carried out only when both the Auto and Deg targets are simultaneously detected. The PowerQuant kit can identify bone DNA samples that will not produce useful STR profiles and that would be analysed with greater success using other genetic markers like SNPs and mtDNA (Zupanič Pajnič et al. 2017).

8.4.5 Genetic Markers

DNA obtained from aged skeletons has deteriorated to a small size on average, usually between 100 and 500 bp, and bases become altered due to molecular damage (Pääbo 1989; Hofreiter et al. 2001). Hydrolytic and oxidative changes are most probable to damage DNA as time passes. Oxidative changes result in modifications of bases, whereas damage due to hydrolytic changes causes the deamination of bases as well as depyrimidination and depurination. Both of these mechanisms diminish the size and number of fragments that PCR can amplify (Keyser-Tracqui and Ludes 2005). Knowledge of these processes and the effects they have on DNA aids researchers in using appropriate assays (targeting of short amplicons) in order to retrieve DNA from aged skeletons. Sex determination was the first study performed on skeletal remains (Fattorini et al. 1988). Formerly, mitochondrial DNA tests were regularly used to forensically identify old skeletal remains (Anslinger et al. 2001; Palo et al. 2007). Recently, some research groups, including the author's own, have described successful nuclear DNA typing from old skeletal remains (Lee et al. 2010; Zupanič Pajnič et al. 2010, 2016b, 2017; Zupanič Pajnič 2013a). Autosomal nuclear DNA is usually typed for missing person identification because it is specific to individuals due to recombination and offers kinship information about both parents. In autosomal DNA, STRs or SNPs are investigated because their high discriminatory power allows reliable identification of the individual (Zupanič Pajnič et al. 2012; Zupanič Pajnič 2013a; Quatrehomme et al. 2019; Pilli et al. 2018). Close relatives are the most suitable for analysing autosomal STRs or SNPs. When close relatives

are not available for identification, it is possible to use distant relatives on the paternal and maternal sides and lineage markers from mtDNA or from the Y chromosome analysed. The Y chromosome passes in unchanged form from fathers to sons, and mtDNA from mothers to all offspring, regardless of sex, and they make it possible to trace the paternal and maternal lines. When even distant relatives are not available for identification, externally visible characteristics of skeletal remains—biogeographical ancestry, or BGA (Phillips et al. 2009; Hollard et al. 2017), eye color (Walsh et al. 2011), and color of the hair (Walsh et al. 2013) are possible to predict through the recent establishment of DNA phenotyping in forensics (Kayser and de Knjiff 2011; Kayser 2015).

The HIrisPlex system for predicting parallel color of eyes and hair from DNA recently showed successful retrieval of eye and hair color information from World War II skeletal remains recovered from Slovenian mass graves using capillary electrophoresis technology for separation of phenotypic SNPs and its possible use for identifying missing persons and disaster victims (Chaitanya et al. 2017). The HIrisPlex system was applied to 49 samples of DNA acquired from the teeth and bones of World War II victims, and all 49 of the samples resulted in complete HIrisPlex profiles except for one MC1R DNA marker (N29insA) that was absent in 83.7% of the samples (Chaitanya et al. 2017).

Until recently, the chief methods for ancestry inference were Y-chromosome STR/SNP and mtDNA control region and mtDNA SNP analyses because they are geographically highly differentiated (Karafet et al. 2008). By increasing the phylogenetic resolution of Y chromosome, researchers recently proved additional informativeness in western European populations (Larmuseau et al. 2014, 2017; Ralf et al. 2015). Analysis of Y-chromosomal and mtDNA markers offers useful information for studying paternal and maternal lineages. Nonetheless, they are not subjected to recombination, and they show only an incomplete picture of an individual ancestry, particularly in the case of complex admixture within a particular individual's genealogy. In contrast, autosomal ancestry-informative markers (AIMs) or ancestry SNPs that have a broad distribution through the autosomes offer a fuller image of overall ancestry and they have utility for identifying an individual's most likely BGA or population of origin, and they have become the main markers for investigating individual ancestry (Kosoy et al. 2009; Tasker et al. 2017). AIMs are classified as genetic polymorphisms; they are mainly SNPs with significant allele frequency divergence between ethnic groups to better characterize genetic differences between them (Espregueira et al. 2016). It is possible to infer genetic ancestry by comparing the genetic diversity of a sample with variation patterns in contemporary populations. In addition to tracing genealogies of individuals, AIMs can play an important role in identifying missing persons and the victims of large-scale disasters (Phillips et al. 2007; Kidd et al. 2014; Puente et al. 2016) and in BGA prediction in crime case work (Pereira et al. 2017; Tasker et al. 2017; Hollard et al. 2017). In comparison to STRs, they are very short and can be used for genotyping difficult forensic samples from badly degraded skeletal remains with minimal amounts of DNA (Romanini et al. 2015; Silvia et al. 2017; Al-Asfi et al. 2018). By genetic typing of STRs and SNPs found on Y chromosome and

sequencing of mtDNA control region and mtDNA coding region SNPs to determine Y-chromosomal and mtDNA haplogroups, and investigations of AIMs, the BGA of World War II victims' skeletons could be determined in the future, and thus their possible ethnic background. In addition, various databases, such as YHRD (the Y-Chromosome STR Haplotype Reference Database) and also the EMPOP (the European Mitochondrial DNA Population Database) can be used to search for matching haplotypes and determination of BGA in the case of a match. There are over 150,000 haplotypes of different populations from across the world in the YHRD database (Willuwiet and Roewer 2007). Because the Y chromosome is inherited on the paternal line, the matching of haplotypes indicates a common ancestral paternal line. By searching for ancestry information using the YHRD database, it is possible to obtain information about matching Y haplotypes and which population they belong to. The same can be applied to analyses of the mtDNA control region. MtDNA is inherited on the maternal line, and the matching of haplotypes indicates a common ancestral maternal line. By using the EMPOP database (Parson and Dür 2007), which contains more than 34,000 haplotypes of different populations from all over the world, it is possible to obtain information about matching mtDNA haplotypes and which population they belong to. Recently, Polish researchers performed genetic analyses of early World War II Sobibór death camp victims. Even though they could not identify them through phylogenetic analyses, they determined that the skeletal remains, originally believed to originate from victims of Russian–Polish conflicts, actually seem to belong to Jews. This is the first solid evidence of Holocaust crimes (Diepenbroek et al. 2018), and it shows the importance of forensic genetics for aiding historical research.

In highly degraded skeletal remains, particularly old ones, STR typing is often not possible. Advances made recently in massively parallel sequencing (MPS), also known as next-generation sequencing (NGS), offers strong advantages over technologies used previously for analysing DNA recovered from human remains (Knapp et al. 2015). With NGS technology, identification of highly degraded skeletal remains could be achieved using identity SNPs (Mehta et al. 2017), which can be used because they are much shorter than STRs. Because of the high copy number per individual cell, mitochondrial DNA analysis is a key DNA detection method used by forensic scientists when old skeletal remains have to be identified. Traditional Sanger sequencing using capillary electrophoresis is time-consuming and cost-prohibitive, and NGS technology provides many advantages. To predict eye and hair color, phenotypic SNPs can be analysed with NGS technology and AIMs to provide BGA information.

8.5 Sample Storage

Storage of samples is of great importance at laboratories engaged in forensic genetics because only high-quality storage makes it possible to successfully recover DNA from a small amount of badly degraded DNA. The quality and quantity of DNA in specimens of aged skeletal remains is influenced by post-excavation treatment, and

samples that have been stored for years in collections at museums, generally at room temperature, exhibit a poorer amplification success rate in comparison to samples that have been freshly excavated (Pruvost et al. 2007). According to Malmstrom (2007), freezing skeletal remains is a preferred method to ensure optimum preservation of DNA. The main reason for freezing samples is to keep DNA degradation to a minimum, preventing the loss of cell integrity and maintaining the quantity and quality of genomic DNA (Lee et al. 2010). Previous studies indicate that there is a reduction in DNA yield with refrigerated liquid DNA (Lee et al. 2010), with temperature changes during freezing and thawing of the samples (Lee et al. 2010; McElderry et al. 2011), and in samples kept in microcentrifuge tubes (Gaillard and Strauss 2000). In addition, spontaneous decay has been reported with progressive molecular damage to DNA (Lindahl 1993). When identification of missing persons based on skeletal remains has been completed, the bone samples typically undergo long-term storage at −20 °C for possible future retesting (Ballou and Stolorow 2013).

Bone samples from World War II can yield low-quantity and low-quality DNA, and it is necessary to perform duplicate analyses of various genetic markers to identify World War II victims. No DNA extract generally remains after analyses, and efficient bone storage is required to guarantee sample stability over time for possible retesting making use of new markers and technologies. Following molecular genetic analysis of World War II victims from Slovenia, small femur fragments are routinely kept at −20 °C. Some authors have observed lower DNA recovery from frozen liquid DNA extracts (Anchordoquyn and Molina 2007; Lee et al. 2010; Hubel et al. 2014), and a study was performed on World War II femurs frozen for 10 years to investigate how freezing bone samples impacts DNA preservation (Friš et al. 2019). To achieve this goal, the quantity of DNA acquired from 57 femurs of World War II victims in a 2009 study (data in Zupanič Pajnič et al. 2010) was compared with the DNA quantities acquired from the identical bones after 10 years in a freezer using the same extraction method and quantification kit. The sample of bone used in this study was adjacent to the one that was used in 2009. The bones were all stored in their original form for 10 years at a temperature of −20 °C. Up to 100 ng DNA/g of powdered bone was acquired from femurs in 2009 and up to 31 ng DNA/g of powdered bone from the same femurs studied after lengthy storage. Statistical analysis demonstrated a marked difference in DNA yield from extracts acquired from World War II bones in 2009 and from extracts acquired from the same bones kept at −20 °C after 10 years, with a greater amount of DNA extracted from bones in 2009 compared to 2019. As described for frozen liquid DNA extracts (Lee et al. 2010; Pokines 2016), reduced DNA recovery was also confirmed for frozen samples of bone, with a marked decrease in yield of DNA after being frozen for 10 years (Friš et al. 2019). The findings showed a reduced DNA recovery for frozen samples of bone.

The second study on storing World War II bone samples focused on comparing the yield of DNA obtained from fragments of bone and powdered bone stored in a freezer (Grdina et al. 2019). Generally not all of the bone powder prepared with a grinder is used for extraction, and because the powdering procedure is

time-consuming it was asked whether storing powdered bone would be useful for future analysis. Hummel (2003) suggests that DNA extraction should take place as soon as possible after homogenizing of bone samples, and long-term bone powder storage could lead to degradation of DNA (even at a low temperature). Researchers concur that storing samples of bone is crucial because preservation of DNA primarily depends on the environmental conditions (Lindahl 1993). Hummel (2003) states that powdering the bones increases the surface area that is exposed and as a result oxidative damage to DNA. Considering this, the aim of the study was to investigate DNA yield in World War II fragments of bone and compare this to the DNA yield from powdered bone stored under identical conditions for 10 years at a temperature of −20 °C. The samples that were analysed had remained following molecular genetic identification carried out in 2009 for 88 Slovenian victims buried in the World War II Konfin Shaft 1 Mass Grave (Zupanič Pajnič et al. 2010). The amount of DNA acquired from 57 femur fragments was compared to the amount of DNA acquired from powdered bone from the same femurs that had been frozen for the same length of time. DNA was extracted from a bone fragment using a piece sampled next to the one used in 2009. This involved working with old DNA, and so every precaution was considered to avoid possible contamination. Statistical analyses showed a difference significant at the 0.05 level in the DNA yield when comparing bones fragments and powdered bone stored at −20 °C for 10 years, and the findings showed there is a greater amount of DNA in powdered bone than in fragments of bone. Because the powdering procedure is time-consuming, it is recommended that, in addition to storing the bone fragment, the powdered bone also be stored for future analyses (Grdina et al. 2019). The results show that lengthy storage of bone powder does not necessarily correspond to a lower quantity of DNA that is extracted from frozen powdered bone when compared to a bone fragment that is frozen, and it is suggested that powdered bone left over from the initial extraction of DNA should be stored long term for investigations in the future along with fragments of bone.

Efforts in forensic research with DNA have focused on the development of innovative methods for long-term storage of samples because today's standard is suboptimal (Lee et al. 2010). Multiple factors have to be considered. Humidity is important because dry storage with a solid matrix inhibits molecular-level movement, and chemical reactions are thus less likely (Pokines 2016). However, when moisture is introduced into the sample again or temperature variations occur, it is again possible for chemical reactions to occur (Pokines 2016). The temperature that the sample is maintained at also impacts the preservation of DNA, and changes in the crystallinity values of the bone mineral are visible after a single freezing, and the decrease in DNA yield is greater after each cycle of freezing and thawing (McElderry et al. 2011). DNA is damaged by repeated cycles of freezing and thawing, and it should therefore be defrosted as rarely as possible (Lindahl 1993). Initial contamination, the effects of light, seal efficiency, and intrinsic bone factors, such as its type and size, also have to be considered (Lee et al. 2010). Fulton (2012) states that the best protocol for lengthy storage of aged specimens varies and that it depends on the way the specimens were collected. Samples that were frozen when

collected should be maintained at that temperature. Samples that were collected at room temperature ought to be stored in a cool, dry environment with a stable temperature, but they might not benefit from freezing, especially if several cycles of freezing and thawing are expected (Fulton 2012). More studies are needed to improve the understanding of forensic DNA sample methodologies for long-term storage and how they affect the quantity of DNA.

8.6 Identification of World War II Victims in Slovenia

Even though seven decades have elapsed since the end of World War II, identification of World War II victims remains relevant because many individuals killed at that time are still missing and have not been identified. Discovering remains dating from World War II is a common occurrence in Slovenia. The Slovenian Government Commission on Concealed Mass Graves has recorded more than 600 hidden mass graves from World War II over the last quarter century (Ferenc 2008). These killings that took place in World War II, amounting to almost 100,000 victims, represent one of the largest-scale losses of life in the modern history of Slovenia. Most victims of killings during and after the war remain buried and are still unidentified. The victims did not have court trials and they were not convicted of crimes. There exist no documents for most mass graves in Slovenia that victim identification could be based on. For some of them, a list of the victims can be made based on archives. Even for those, it is difficult to find living relatives because such a long time has elapsed since the massacres during World War II. The author's team has been conducting genetic investigations of victims of World War II from only a few mass graves, where a list of the Slovenian victims was found in archives. Some of the victims have been identified because putative relatives (close and/or distant) were available for identification (Zupanič Pajnič 2008; Zupanič Pajnič et al. 2010, 2018a, b). A 99.9% recommended posterior probability was applied, with the aim of high confidence in correctly identifying the individuals (Brenner and Weir 2003; Biesecker et al. 2005; Prinz et al. 2007). By combining various genetic markers and involving both close and distant relatives, a sufficiently high statistical probability of positively identifying the victims from mass graves was achieved, even if it was possible to trace only distant relatives of the victims. This proved that by combining many genetic markers it is possible to identify victims of World War II despite an absence of close relatives. Relatives of identified victims have attained certainty about the disappearance of their loved ones after over seven decades, and they have buried the identified skeletal remains in family graves. Genetic identifications have helped reveal information about historical events in Slovenia immediately after World War II. Recent phylogenetic analysis of Polish World War II death camp victims and the establishment of their Jewish origin shows the great importance of forensic genetics for aiding historical research even when it is not possible to identify victims by name (Diepenbroek et al. 2018).

For traceability of possible contamination, for each mass grave an elimination database was created. The database included genetic profiles for the nuclear DNA

(autosomal and Y-chromosomal STRs) and mtDNA of every person that had had contact with the skeletal remains. When identifying victims of the World War II mass graves, the chance of contamination occurring during genetic investigation was minimized. The authenticity of bones' genetic profiles was checked through clean isolation and amplification-negative controls used for nuclear DNA, identical genetic profiles acquired using different PCR amplification kits, and mismatch of genetic profiles for bones with people included in the elimination database.

8.6.1 Identification of the Konfin Shaft 1 Mass Grave Victims

In 2008/2009, one-third of the remains of 88 Slovenian victims from the Konfin Shaft 1 Mass Grave (*Grobišče Brezno pri Konfinu 1*) were identified (Zupanič Pajnič et al. 2010). Victims were taken from prison on the night of 24 June 1945 and transported to the place of execution. The victims included 40 men that were wounded (patients that had been taken from the hospital in Ljubljana and moved to the central prison of the OZNA (secret police) 2 weeks before their execution), and 48 that had been selected from the prisoners. These men did not have court trials and they were not convicted of any crimes. Their corpses were dumped into a 45-m-deep cave and the opening was dynamited. The remains were not covered with soil, which would have held the skeletons in their original positions. Precipitation runoff flowed unimpeded into the cave, and its bottom, measuring 20 m^2, was covered completely with a layer of bone sand mud 2 m thick. Living relatives were traced for 36 of the victims. Eighty-four right femurs were analysed and the genetic profiles were compared to genetic material from living relatives. Y-chromosome haplotypes and autosomal genetic profiles were acquired from 98% of the skeletal remains, and mtDNA haplotypes from 95% of the remains for the HVI region and from 97% of the remains for the HVII region. The genetic profiles for nuclear DNA and mtDNA were established for the reference persons. When comparing genetic profiles, 28 out of the 84 bones analysed were matched with relatives still living (siblings, children, nephews, or cousins) and the statistical analysis showed a high degree of confidence for correct identification of all of the 28 victims (PP ranged from 99.9% to more than 99.999999%). Nuclear DNA was obtained from skeletal remains over 70 years old, and STR typing was successful. This shows that combining a large number of genetic markers offers very high likelihood ratios, supporting the hypothesis that the bones of individuals are related to the family references instead of unrelated individuals. If the analysis had involved only autosomal STR loci, only 12 victims could have been identified with high confidence that the identification was correct. It is therefore also deemed necessary to include Y-STRs and mtDNA analyses, and also to include both close and distant relatives on both the maternal and paternal lines when identifying World War II victims. Identification of the remains in the Konfin Shaft 1 Mass Grave established a basis for future molecular genetic investigations of Slovenian post-war mass graves when the opportunity arose. It should be stressed that mass graves are rarely associated with lists of their victims. The Konfin

Shaft 1 Mass Grave is such a rare case because it was possible to compile a list of its victims based on archival documents. The methods for extracting and amplifying DNA that were used for identifying the Konfin Shaft 1 Mass Grave victims proved very efficient because full genetic profiles were generated for autosomal DNA, mtDNA haplotypes, and Y-STR haplotypes (Zupanič Pajnič et al. 2010).

8.6.2 Identification of the Storžič Victims

In 1944, below Mount Storžič, at an elevation of 1200 m, five men were taken to a killing site. One of them fell into a shaft during an attempt to escape, and the other four were killed and their bodies were buried in a hidden mass grave. This grave was found in 2007 and four skeletons were excavated. One victim was identified based on an anthropological study, and the other three were identified by genetic methods. Identification was performed by analysing excavated femurs and teeth and comparing the genetic profiles obtained with the living relatives' genetic material (Zupanič Pajnič 2008). Autosomal genetic profiles, Y-chromosome haplotypes, and mtDNA haplotypes, were acquired from all of the tooth and bone and from reference persons. When comparing the genetic profiles, all the teeth and bones matched with living relatives (for the first victim the reference person was a daughter, for the second victim the reference person was a niece on the maternal line, and for the third victim the reference persons were a son and a niece on the maternal line). The statistical analysis exhibited a high confidence that identification was correct with posterior probabilities (PP) that ranged from 99.9999% to 99.999999%, speaking in favor of the positive identification of the victims.

8.6.3 Identifying a Cranium from Bohinj

During World War II, a young woman was executed in Bohinj and she left behind a 1-year-old daughter. Four decades ago, a skeleton was excavated from an individual grave in the Bohinj area and only the cranium from the grave was kept at the Little War Museum (*Mali vojni muzej*). The burial site conformed to accounts from witnesses, but initial anthropological and morphological screening was unable to identify the sex of the remains. In 2015, genetic analysis was performed on the cranium to determine whether it was from the mother of the presumed daughter serving as a family reference. The genetic analysis was conducted alongside a more careful anthropological analysis of the cranium. The genetic examination was carried out on two molars and a petrous portion of the temporal bone, and the results acquired from the teeth corroborated the results from the temporal bone. Nuclear DNA for STR typing was obtained from the temporal bone and from the left second molar. A complete autosomal genetic profile, which included amelogenin locus, indicated that the cranium was from a male, and analyses of Y-STRs provided further confirmation of this. The same conclusions were reached following an anthropological analysis, which indicated that the cranium had come from a very

young Caucasoid male. The male sex of the cranium excluded the possibility that it had belonged to the mother of the female reference used for comparison (Zupanič Pajnič et al. 2016b).

8.6.4 Identification of the Mače Victims

In 2015, genetic identification of a Slovenian banker and industrialist was performed using autosomal and Y-STR markers (Zupanič Pajnič et al. 2018a). He was killed along with his wife (she was the first Yugoslav female pilot) in January of 1944 close to his residence (Strmol Castle). The couple came from well-known and wealthy Slovenian families, who were part of the pre-war elite in Slovenia. The concealed grave with the skeletal remains of the husband and wife was discovered in 2015 in Mače, and only the incomplete remains of a male skeleton and a female skeleton were recovered. A piece of the skull, ribs, and vertebrae were absent from the male's skeleton, and of the female's skeleton the only parts found were the skull lacking the lower jaw and a foot. It is presumed that animals carried the bones away from the shallow grave. Because of the lengthy timespan since the World War II massacres, finding living relatives for identifying victims in such mass graves is difficult. Living relatives could be found only for the husband (two nephews and a niece on the paternal side), and, because his wife had no children, it was not possible to genetically identify her. From the poorly preserved and incomplete remains of the skeletons, genetic typing was performed on the female skeleton's right third molar (skeleton B) and on the male skeleton's left second molar, right tibia, and right femur (skeleton A). The sexes of the two skeletons were confirmed using amelogenin and Y-STR typing. Skeleton A's left second molar, right tibia, and right femur yielded identical profiles. Full autosomal and Y-STR profiles allowed identification of the male skeleton through a comparison to family references. Y-STRs confirmed the relationships between selected males (an uncle and nephews). The prior probability was established on the basis of the number of victims that were reported, and the study used a recommended PP (for kinship) of 99.9% with the aim of high confidence for correctly identifying the mass grave victim. The product rule was applied in order to estimate a combined LR for autosomal and Y-STRs, and the results of statistical analysis indicated a high confidence for correct identification with a PP of 99.997%. After over seven decades, the couple's skeletal remains were returned to their living relatives, and the relatives buried them in a family grave (Zupanič Pajnič et al. 2018a).

8.6.5 Identification of the Mačkovec Victims

In 1943, according to historical data, 43 prisoners were executed at Mačkovec Hill; two others were able to escape the killings. Bones lying in the forest next to the road were buried in subsequent years; although they could have had been exposed to wild animals prior to their burial. The number of excavated skeletons thus differed from

the number of victims presented in historical documents. Excavations in recent years have proven the existence of two hidden mass graves at the location in question. The first one is designated "Mačkovec, large grave," where the remains of 14 different individuals buried in a shallow grave were found, and the second one "Mačkovec, small grave," where two individuals' remains were found. The human remains were originally covered only with fallen leaves and branches. Consequently, the skeletal remains discovered were poorly preserved, being exposed to harmful environmental factors. In 2013, the author's team was asked to perform genetic identification of victims whose skeletal remains had been excavated from the Mačkovec Mass Grave. Anthropological study determined how many victims there were based on the number of ulnas. However, all the tibias and femurs were included in the genetic analyses because genetic profiling from these delivers far greater success than those from ulnas. All the tibias and femurs were analysed to obtain as many unique genetic profiles as possible. Moreover, the poorly preserved nature of remains dating back to World War II presents an obstacle in STR typing, resulting in allele dropouts and partial profiles, and therefore more skeletal elements should be genotyped to determine a genetic profile. DNA analysis was performed on nine right femurs, eight left femurs, ten right tibias, 11 left tibias, and one tibia with undetermined laterality, or a total of 39 bones. Buccal swab samples were collected for 13 family references for ten different victims according to the victims list; for three of the victims, two family references were obtained. Presumed family relations with the deceased were sons (for four victims), daughters (for three victims), a sister (for one victim), nephews on the father's side (for three victims), and grandchildren on the father's side (for two victims). An anthropological report for the larger of the two graves described eight right and seven left femurs, and genetic profile matches were found for five pairs of femurs. Furthermore, out of ten right and ten left tibias and one tibia without determined laterality found in the larger grave, as described by anthropology analysis, it was possible to match nine pairs according to genetic profiles. Left and right femurs, as determined by an anthropologist, and a tibia, all of which were found in the smaller grave, had matching genetic profiles and thus belong to the same victim. Four different genetic profile matches were ascertained, among which two cases were a match between a victim and a family reference, and two cases a match between two respective victims, the latter highlighting the fact that some of the victims were related. The product rule was applied in order to estimate a combined LR for autosomal and Y-STRs, and a high confidence of correct identification was demonstrated through statistical analyses with PP greater than 99.9% for all of the victims identified (data not published yet).

8.6.6 Identification of the Babna Gora Victims

In 2016, the victims of the single largest massacre of a family in Slovenia were genetically identified (Zupanič Pajnič et al. 2018b, 2020). Ten people from the same family were killed in World War II: nine of them in 1942 and one immediately after the war. During the massacre in 1942, seven members were killed in the forest and

their bodies were hidden in a mass grave. Soon after these seven people were killed, the mother and grandmother of the children that had survived were murdered at their home, and the two of them were buried in the local cemetery. Two children, a 4-year-old girl and her 3-year-old brother, were the only ones who survived the massacre. Besides these two children, an uncle also survived the killings in 1942, but he was murdered immediately after World War II. A total of ten people from the same family were killed during World War II. Exhumation of the remains started in March 2015, but archeologists exhumed only three female skeletons, which were incomplete. Twenty meters from this site, relatives later found bones, and the burial location of at least three males was excavated in August 2016. The grave excavated in 2015 was labeled a female grave, and the one excavated in 2016 was labeled a male grave. The family members were buried in shallow graves in the woods, and the skeletons that were exhumed were incomplete at both sites. Aside from teeth in one mandible, no teeth were found. The reason why the skeletons were incomplete could be because the graves were very shallow, making it very probably that wild animals scattered the victims' remains in the woods. The remains underwent taphonomic processes after deposition, which resulted in a poor state of preservation. The bones that were exhumed were abraded, otherwise damaged, and porous, without features that would permit anthropologists to macroscopically determine their laterality. The graves were positioned along a creek in the woods, and so water exposure caused diagenetic and degradation processes. For the first skeleton recovered from the female grave, the only parts found were a piece of the skull without teeth, part of a femur, and part of a tibia. For the second, the pieces recovered were part of the skull, a segment of spine, a humerus, and the right leg bones. The skull preserved three mandibular teeth: the left and right second molars and also the left first premolar. For the third skeleton in the female grave, only part of the skull lacking teeth, a piece of the spine, and a tibia were discovered. In the male grave, only individual bones (including the diaphysis from six femurs) were recovered by archeologists, and the number of individuals buried there was at least three. Following the removal of the skeletal remains from the two graves, the author's team was requested to carry out genetic identification by comparing findings with two living relatives (the aforementioned daughter and son). Altogether, analysis was made of 12 bones and teeth, and these were compared with the two living relatives. Seven bones and one molar yielded nuclear DNA for successful STR typing. The female grave yielded autosomal profiles from only one of the skeletons, and autosomal profiles were obtained from five of the six femurs in the male grave. The relationships among the males were further confirmed through analyses of Y-STRs and STR profiles made it possible to identify four members of the family; an aunt from the female grave, and from the male grave two uncles plus the father of the living children, who served as family references. The product rule was applied to calculate a combined likelihood ratio for autosomal and Y-STRs, and high confidence of having made a correct identification was demonstrated through statistical analyses, with PP greater than 99.9% for three of the four victims identified. For identification of the aunt, the PP determined following capillary electrophoresis STR typing was not high enough. To achieve a better PP, we used the next-generation

sequencing Precision ID GlobalFiler NGS STR Panel (TFS) and, following an analysis of additional autosomal STR loci, our statistical analysis indicated a PP exceeding 99.9%, showing that a sufficient number of genetic markers had been studied to identify the aunt's skeletal remains (Zupanič Pajnič et al. 2019). After over seven decades, the remains of the victims were returned to the two children that had survived, and they laid their relatives to rest in the family grave (Zupanič Pajnič et al. 2020).

Acknowledgments The author thanks the Slovenian Government Commission on Concealed Mass Graves for its support for excavations involving World War II victims. Some of the studies described in this chapter received financial support from the Slovenian Research Agency (the project "Determination of the most appropriate skeletal elements for molecular genetic identification of aged human remains," J3-8214).

References

Al-Asfi M, McNevin D, Mehta B, Power D, Gahan ME, Daniel R (2018) Assessment of the precision ID ancestry panel. Int J Legal Med 132:1581–1594

Alberti F, Gonzalez J, Paijmans JLA (2018) Optimized DNA sampling of ancient bones using computed tomography scans. Mol Ecol Resour 18:1196–1208

Alonso A, Anđelinović Š, Martin P (2001) DNA typing from skeletal remains: evaluation of multiplex and megaplex STR systems on DNA isolated from bone and teeth samples. Croat Med J 42:260–266

Amory S, Huel R, Bilić A, Loreille O, Parsons TJ (2012) Automatable full demineralization DNA extraction procedure from degraded skeletal remains. Forensic Sci Int Genet 6:398–406

Anchordoquyn TJ, Molina MC (2007) Frontiers in clinical research preservation of DNA. Cell Preserv Tech 5:180–188

Anderung C, Persson P, Bouwman A, Elburg R, Götherström A (2008) Fishing for ancient DNA. Forensic Sci Int Genet 2:104–107

Anslinger K, Weichhold G, Keil W, Bayer B, Eisenmenger W (2001) Identification of the skeletal remains of Martin Bormann by mtDNA analysis. Int J Legal Med 114:194–196

Ballou S, Stolorow M (2013) The biological evidence preservation handbook: best practices for evidence handlers. U.S. Department of Commerce, National Institute of Standards and Technology, Washington, DC, pp 9–25

Barta JL, Monroe C, Kemp BM (2014) Mitochondrial DNA preservation across 3000-years-old northern fur seal ribs is not related to bone density: implications foe forensic investigations. Forensic Sci Int 239:11–18

Baum DA, Futuyma DJ, Hoekstra HE, Lenski RE, Moore JA, Peichel CL et al (2013) The Princeton guide to evolution. Princeton University Press, Princeton, NJ, pp 477–479

Biesecker LG, Bailey-Wilson JE, Ballantyne J, Baum H, Bieber FR, Brenner C et al (2005) DNA identification after the 9/11 World Trade Center attack. Science 310:1122–1123

Brenner CH, Weir BS (2003) Issues and strategies in the DNA identification of World Trade Center victims. Theoret Popul Biol 63:173–178

Burger J, Hummel S, Hermann B, Henke W (1999) DNA preservation: a microsatellite-DNA study on ancient skeletal remains. Electrophoresis 20:1722–1728

Camposa PF, Craig OE, Turner-Walker G, Peacock E, Willerslev E, Gilbert MTP (2012) DNA in ancient bone - where is it located and how should we extract it? Ann Anat 194:7–16

Chaitanya L, Zupanič Pajnič I, Walsh S, Balažic J, Zupanc T, Kayser M (2017) Bringing colour back after 70 years: predicting eye and hair colour from skeletal remains of World War II victims using the HIrisPlex system. Forensic Sci Int Genet 26:48–57

Coulson-Thomas YM, Norton AL, Coulson-Thomas VJ, Florencio-Silva R, Ali N, Elmrghni S et al (2015) DNA and bone structure preservation in medieval human skeletons. Forensic Sci Int 251:186–194

Diepenbroek M, Strobl C, Niederstätter H, Zimmermann B, Szargut M, Zielinska G et al (2018) The phylogenetic analyses of the human remains found in the Nazi German death camp as a proof of the Holocaust. In: The 11th Haploid Markers Conference: Inferring Ancestry from DNA, 17–19 May, 2018, Bydgoszcz, Poland

Edson SM, Ross JP, Coble MD, Parsons TJ, Barritt SM (2004) Naming the dead - confronting the realities of rapid identification of degraded skeletal remains. Forensic Sci Rev 16:64–89

Espregueira T, Smidt Mogensen H, Børsting C, Morling N (2016) Frequencies of HID-ion ampliseq ancestry panel markers among Greenlanders. Forensic Sci Int Genet 24:60–64

Fattorini P, Caccio S, Gustincich S, Altamura B, Graziosi G (1988) Sex determination from skeleton: a new method using a DNA probe. Acta Med Leg Soc (Leige) 39(2):201–205

Ferenc M (2008) Topografija evidentiranih grobišč (Topography of documented mass graves). In: Dežman J (ed) Poročilo Komisije vlade Republike Slovenije za reševanje vprašanj prikritih grobišč 2005-2008. Družina, Ljubljana, pp 7–27

Friš EL, Grdina S, Podovšovnik E, Zupanc T, Zupanič Pajnič I (2019) Comparison of DNA yield after long-term storage of Second World War bone samples. Forensic Sci Int Genet Suppl Ser 7:117

Fulton LT (2012) Ancient DNA - methods and protocols: setting up an ancient DNA laboratory. Humana Press Inc, Totowa, NJ, pp 1–11

Gaillard C, Strauss F (2000) Eliminating DNA loss and denaturation during storage in plastic microtubes. Int Biotechnol Lab 6:24

Geršak ŽM, Zupanič Pajnič I, Črešnar M, Zupanc T (2019) Determination of DNA yield rates in six different skeletal elements in ancient bones. Forensic Sci Int Genet Suppl Ser 7:120

Gilbert MTP, Rudbeck L, Willerslev E (2005) Biochemical and physical correlates of DNA contamination in archaeological bone and teeth excavated at Matera, Italy. J Archaeol Sci 32:785–793

Grdina S, Friš EL, Podovšovnik E, Zupanc T, Zupanič Pajnič I (2019) Storage of Second World War bone samples: bone fragments versus bone powder. Forensic Sci Int Genet Suppl Ser 7:175

Hansen HB, Damgaard PB, Margaryan A, Stenderup J, Lynnerup N, Willerslev E et al (2017) Comparing ancient DNA preservation in petrous bone and tooth cementum. PLoS One 12(1): e0170940

Hartman D, Drummer O, Eckhoff C, Scheffer JW, Stringer P (2011) The contribution of DNA to the disaster victim identification (DVI) effort. Forensic Sci Int 205:52–58

Higgins D, Austin JJ (2013) Teeth as a source of DNA for forensic identification of human remains: a review. Sci Justice 53:433–441

Higgins D, Rohrlach AB, Kaidonis J (2015) Differential nuclear and mitochondrial DNA preservation in post-mortem teeth with implications for forensic and ancient DNA studies. PLoS One 10: e0126935

Hofreiter M, Serre D, Poinar HN, Kuch M, Pääbo S (2001) Ancient DNA. Nat Rev Genet 2:353–359

Hollard C, Keyser C, Delabarde T, Gonzalez A, Vilela Lamego C, Zvénigorosky V, Ludes B (2017) Case report: on the use of the HID-Ion AmpliSeq Ancestry Panel in a real forensic case. Int J Legal Med 131:351–358

Höss M, Jaruga P, Zastawny TH, Dizdaroglu M, Pääbo S (1996) DNA damage and DNA sequence retrieval from ancient tissues. Nucleic Acids Res 24:1304–1307

Hubel A, Spindler R, Skubitz AP (2014) Storage of human hiospecimens: selection of the optimal storage temperature. Biopreserv Biobank 12:165–175

Hummel S (2003) Ancient DNA typing - methods, strategies and applications. Springer, Berlin, pp 225–227

Irwin JA, Just RS, Loreille OM, Parsons TJ (2012) Characterization of a modified amplification approach for improved STR recovery from severely degraded skeletal elements. Forensic Sci Int Genet 6:578–587

Iwamura ESM, Soares-Vieira JA, Muñoz DR (2004) Human identification and analysis of DNA in bones. Rev Hosp Clin Fac Med S Paulo 59:383–388

Jakubowska J, Maciejewska A, Pawłowski R (2012) Comparison of three methods of DNA extraction from human bones with different degrees of degradation. Int J Legal Med 126:173–178

Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF (2008) New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. Genome Res 18:830–838

Kayser M (2015) Forensic DNA phenotyping: predicting human appearance from crime scene material for investigative purposes. Forensic Sci Int Genet 18:33–48

Kayser M, de Knjiff P (2011) Improving human forensics through advances in genetics, genomics and molecular biology. Nat Rev Genet 12:179–192

Kendall C, Høier Eriksen AM, Kontopoulos I, Collins MJ, Turner-Walker G (2018) Diagenesis of archaeological bone and tooth. Paleogeogr Palaeocl 491:21–37

Keyser-Tracqui C, Ludes B (2005) Methods for the study of ancient DNA. In: Carracedo A (ed) Forensic DNA typing protocols. Humana Press Inc, New York, NY, pp 253–264

Kidd KK, Speed WC, Pakstis AJ, Furtado MR, Fang R, Madbouly A et al (2014) Progress toward an efficient panel of SNPs for ancestry inference. Forensic Sci Int Genet 10:23–32

Kishore R, Hardy WR, Anderson VJ, Sanchez NA, Buoncristiani MR (2006) Optimization of DNA extraction from low-yield and degraded samples using the biorobot EZ1 and biorobot M48. J Forensic Sci 51:1055–1061

Knapp M, Lalueza-Fox C, Hofreiter M (2015) Re-inventing ancient human DNA. Investig Genet 6:4

Kosoy R, Nassir R, Tian C, White PA, Butler LM, Silva G et al (2009) Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America. Hum Mutat 30:69–78

Krenke BE, Nassif N, Sprecher CJ, Knox C, Schwandt M, Storts DR (2008) Developmental validation of a real-time PCR assay for the simultaneous quantification of total human and male DNA. Forensic Sci Int Genet 3:14–21

Larmuseau MHD, Vanderheyden N, Van Geystelen A, van Oven M, Kayser M, Decorte R (2014) Increasing phylogenetic resolution still informative for Y chromosomal studies on West - European populations. Forensic Sci Int Genet 9:179–185

Larmuseau MHD, Otten GPPL, Decorte R, Van Damme P, Moisse M (2017) Defining Y-SNP variation among the Flemish population (Western Europe) by full genome sequencing. Forensic Sci Int Genet 31:e12–e16

Lee SB, Crouse CA, Kline MC (2010) Optimizing storage and handling of DNA extracts. Forensic Sci Rev 22:131–144

Lindahl T (1993) Instability and decay of the primary structure of DNA. Nature 362:709–715

Loreille OM, Diegoli TM, Irwin JA, Coble MD, Parsons TJ (2007) High efficiency DNA extraction from bone by total demineralization. Forensic Sci Int Genet 1:91–195

Malmstrom H (2007) Ancient DNA as a means to investigate the European Neolithic. Doctoral dissertation, Uppsala University, Uppsala, Sweden

Mansour H, Krebs O, Sperhake JP, Augustin C, Koehne T, Amling M et al (2018) Cementum as a source of DNA in challenging forensic cases. J Forensic Legal Med 54:76–81

McElderry JP, Kole MR, Morris MD (2011) Repeated freeze-thawing of bone tissue affects Raman bone quality measurements. J Biomed Opt 16(071407):1–4

Mehta B, Daniel R, Phillips C, McNevin D (2017) Forensically relevant SNaPshot® assays for human DNASNP analysis: a review. Int J Legal Med 131:21–37

Miloš A, Selmanović A, Smajlović L, Huel RLM, Katzmarzyk C, Rizvić A, Parsons JP (2007) Success rates of nuclear short tandem repeat typing from different skeletal elements. Croat Med J 48:486–493

Misner LM, Halvorson AC, Dreier JL, Ubelaker DH, Foran DR (2009) The correlation between skeletal weathering and DNA quality and quantity. J Forensic Sci 54:822–828

Montellius K, Lindblom B (2012) DNA analysis in disaster victim identification. Forensic Sci Med Pathol 8:140–147

Montpetit SA, Fitch IT, O'Donnell PT (2005) A simple automated instrument for DNA extraction in forensic casework. J Forensic Sci 50:555–563

Morild I, Hamre SS, Huel R, Parsons TJ (2015) Identification of missing Norwegian World War II soldiers, in Karelia Russia. J Forensic Sci 60:1104–1110

Mundorff A, Davoren JM (2014) Examination of DNA yield rates for different skeletal elemenets at increasing post mortem intervals. Forensic Sci Int Genet 8:55–63

Mundorff AZ, Bartelink EJ, Mar-Cash E (2009) DNA preservation in skeletal elements from the world trade center disaster: recommendations for mass fatality management. J Forensic Sci 54:739–745

Nagy M, Otremba P, Krüger C, Bergner-Greiner S, Anders P, Henske B, Prinz M, Roewer L (2005) Optimization and validation of a fully automated silica-coated magnetic beads technology in forensics. Forensic Sci Int 152:13–22

Obal M, Zupanič Pajnič I, Gornjak Pogorelc B, Zupanc T (2019) Different skeletal elements as a source of DNA for genetic identification of Second World War victims. Forensic Sci Int Genet Suppl Ser 7:27

Ossowski A, Kuś M, Brzeziński P, Prüffer J, Piątek J, Zielińska G (2013) Example of human individual identification from World War II gravesite. Forensic Sci Int 233:179–192

Pääbo S (1989) Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. Proc Natl Acad Sci U S A 86:1939–1943

Pääbo S (1990) Amplifying ancient DNA. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds) PCR-protocols and amplifications-a laboratory manual. Academic, San Diego, CA, pp 159–166

Pääbo S, Poinar H, Serre D (2004) Genetic analyses from ancient DNA. Annu Rev Genet 38:645–679

Palo JU, Hedman M, Soderholm N, Sajantila A (2007) Repatriation and identification of Finnish World War II soldiers. Croat Med J 48:528–535

Parson W, Dür A (2007) EMPOM - a forensic mtDNA database. Forensic Sci Int Genet 1:88–92

Pereira V, Mogensen HS, Børsting C, Morling N (2017) Evaluation of the precision ID ancestry panel for crime case work: a SNP typing assay developed for typing of 165 ancestral informative markers. Forensic Sci Int Genet 28:138–145

Phillips C, Salas A, Sánchez JJ, Fondevila M, Gómez-Tato A, Alvarez-Dios J et al (2007) Inferring ancestral origin using a single multiplex assay of ancestry-informative marker SNPs. Forensic Sci Int Genet 4:273–280

Phillips C, Prieto L, Fondevila M, Salas A, Gómez-Tato A, Alvarez-Dios J et al (2009) Ancestry analysis in the 11-M Madrid bomb attack investigation. PLoS One 4:e6583

Pilli E, Boccone S, Agostino A, Virgili A, D'Errico G, Lari M et al (2018) From unknown to known: identification of the remains at the mausoleum of fosse Ardeatine. Sci Justice 58:469–478

Pinhasi R, Fernandes D, Sirak K, Novak M, Connell S, Alpaslan-Roodenberg S et al (2015) Optimal ancient DNA yields from the inner ear part of the human petrous bone. PLoS One 10 (6):e0129102

Poinar HN, Hoss M, Bada JL, Pääbo S (1996) Amino acid racemisation and the preservation of ancient DNA. Science 272:864–866

Pokines JT (2016) The effects of experimental freeze-thaw cycles to bone as a component of subaerial weathering. J Archaeol Sci Rep 6:594–602

Prinz M, Carracedo A, Mayr WR, Morling N, Parsons TJ, Sajantila A et al (2007) DNA Commission of the International Society for Forensic Genetics (ISFG): recommendations regarding the

role of forensic genetics for disaster victims identification (DVI). Forensic Sci Int Genet 1 (1):3–12

Pruvost M, Schwarz R, Correia VB (2007) Freshly excavated fossil bones are best for amplification of ancient DNA. Proc Natl Acad Sci U S A 104:739–744

Puente M, Santos C, Fondevila M, Manzo L, EUROFORGEN-NoE Consortium, Carracedo Á, Lareu MV et al (2016) The global AIMs Nano set: a 31-plex SnaPshot assay of ancestry-informative SNPs. Forensic Sci Int Genet 22:81–88

Putkonen MT, Palo JU, Cano JM, Hedman M, Sajantila A (2010) Factors affecting the STR amplification success in poorly preserved bone samples. Investig Genet 1:9

Quatrehomme G, Unité Police d'Identification de Victimes de Catastrophes, Toupenay S, Delabarde T, Padovani B, Alunni V (2019) Forensic answers to the 14th of July 2016 terrorist attack in Nice. Int J Legal Med 133:277–287

Ralf A, van Oven M, Zhong K, Kayser M (2015) Simultaneous analysis of hundreds of Y-chromosomal SNPs for high-resolution paternal lineage classification using targeted semiconductor sequencing. Hum Mutat 36:151–159

Rohland N, Hofreiter M (2007) Ancient DNA extraction from bones and teeth. Nat Protoc 2:1756–1762

Romanini C, Catelli ML, Borosky A, Pereira R, Romero M, Salado Puerto M et al (2012) Typing short amplicon binary polymorphisms: supplementary SNP and Indel genetic information in the analysis of highly degraded skeletal remains. Forensic Sci Int Genet 6:469–476

Romanini C, Romero M, Salado Puerto M, Catelli L, Phillips C, Pereira R et al (2015) Ancestry informative markers: inference of ancestry in aged bone samples using an autosomal AIM-Indel multiplex. Forensic Sci Int Genet 16:58–63

Salamon M, Tuross N, Arensburg B, Weiner S (2005) Relatively well preserved DNA is present in the crystal aggregates of fossil bones. Proc Natl Acad Sci U S A 102:13783–13788

Sampierto ML, Gilbert MT, Lao O, Caramelli D, Lari M, Berranpetit J, Lalueza-Fox C (2006) Tracking down human contamination in ancient human teeth. Mol Biol Evol 23:1801–1807

Schwartz TR, Schwartz EA, Mieszerski L, McNally L, Kobilinsky L (1991) Characterization of deoxyribonucleic acid (DNA) obtained from teeth subjected to various environmental conditions. J Forensic Sci 36:979–990

Silvia AL, Shugarts N, Smith J (2017) A preliminary assessment of the ForenSeq™ FGx System: next generation sequencing of an STR and SNP multiplex. Int J Legal Med 131:73–86

Smith CI, Chamberlain AT, Riley MS, Stringer C, Collins MJ (2003) The thermal history of human fossils and the likelihood of successful DNA amplification. J Hum Evol 45:203–217

Tasker E, LaRue B, Beherec C, Gangitano D, Hughes-Stamm S (2017) Analysis of DNA from post-blast pipe bomb fragments for identification and determination of ancestry. Forensic Sci Int Genet 28:195–202

Turner-Walker G (2008) The chemical and microbial degradation of bones and teeth. In: Pinhasi R, Mays S (eds) Advances in human palaeopathology. John Wiley and Sons Ltd, New York, NY, pp 1–29

Valgren C, Wester S, Hansson O (2008) A comparison of three automated DNA purification methods in forensic casework. Forensic Sci Int Genet Suppl Ser 1:76–77

Walsh S, Liu F, Ballantyne KN, van Oven M, Lao O, Kayser M (2011) IrisPlex: a sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. Forensic Sci Int Genet 5:170–180

Walsh S, Liu F, Wollstein A, Kovatsi L, Ralf A, Kosiniak-Kamysz A et al (2013) The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA. Forensic Sci Int Genet 7:98–115

Willuwiet S, Roewer L (2007) Y chromosome haplotype reference database (YHRD): update. Forensic Sci Int Genet 1:83–87

Zietkiewicz E, Witt M, Daca P, Zebracka-Gala J, Goniewicz M, Jarząb B et al (2012) Current genetic methodologies in the identification of disaster victims and in forensic analysis. J Appl Genet 53:41–60

Zupanc T, Balažic J, Štefanič B, Zupanič Pajnič I (2013) Performance of the Human Quantifiler, the Investigator Quantiplex, and the Investigator ESSplex Plus kit for quantification and nuclear DNA typing of old skeletal remains. Rom J Legal Med 21:119–124
Zupanič Pajnič I (2008) Molekularno genetska identifikacija domobranskih žrtev (Molecular genetic identification of Slovenian home guard victims). Zdrav Vestn 77:745–750
Zupanič Pajnič I (2011) Visoko učinkovita metoda ekstrakcije DNA iz skeletnih ostankov (Highly efficient DNA extraction method from skeletal remains). Zdrav Vestn 80:171–181
Zupanič Pajnič I (2013a) A comparative analysis of the AmpFlSTR identifiler and PowerPlex 16 autosomal short tandem repeat (STR) amplification kits on the skeletal remains excavated from Second World War mass graves in Slovenia. Revista de medicinæalegalæa 21:73–78
Zupanič Pajnič I (2013b) Genetic identification of Second World War victim's skeletal remains. Lap Lambert Academic Publishing, Saarbrucken
Zupanič Pajnič I (2016) Extraction of DNA from human skeletal material. In: Goodwin W (ed) Forensic DNA typing protocols, methods in molecular biology, vol 1420. Springer Science & Business Media, LLC, New York, NY, pp 89–108
Zupanič Pajnič I (2017) Evaluation of the quality of bone powder for successful STR typing of human skeletal remains. Revista de medicinæ alegalæa 25:92–98
Zupanič Pajnič I, Gornjak-Pogorelc B, Balažic J (2010) Molecular genetic identification of skeletal remains from the Second world war Konfin I mass grave in Slovenia. Int J Legal Med 124:307–317
Zupanič Pajnič I, Gornjak Pogorelc B, Balažic J, Zupanc T, Štefanič B (2012) Highly efficient nuclear DNA typing of the World War II skeletal remains using three new autosomal short tandem repeat amplification kits with the extended European Standard Set of loci. Croat Med J 53:17–23
Zupanič Pajnič I, Debska M, Gornjak Pogorelc B, Vodopivec Mohorčič K, Balažic J, Zupanc T et al (2016a) Highly efficient automated extraction of DNA from old and contemporary skeletal remains. J Forensic Legal Med 37:78–86
Zupanič Pajnič I, Petaros A, Balažic J, Geršak K (2016b) Searching for the mother missed since the Second World War. J Forensic Legal Med 44:138–142
Zupanič Pajnič I, Zupanc T, Balažic J, Geršak ŽM, Stojković O, Skadrić I et al (2017) Prediction of autosomal STR typing success in ancient and Second World War bone samples. Forensic Sci Int Genet 27:17–26
Zupanič Pajnič I, Bajželj M, Gornjak-Pogorelc B, Inkret J, Balažic J (2018a) Identification of the Second World War victim using autosomal and Y - STR markers. In: The 11th Haploid Markers Conference: Inferring ancestry from DNA, 17–19 May, 2018, Bydgoszcz, Poland
Zupanič Pajnič I, Obal M, Bajželj M, Gornjak-Pogorelc B, Inkret J, Balažic J (2018b) Victims of the biggest Second World War family killing in Slovenia identified via autosomal and haploid markers. In: The 11th Haploid Markers Conference: Inferring ancestry from DNA, 17–19 May, 2018, Bydgoszcz, Poland
Zupanič Pajnič I, Gornjak Pogorelc B, Zupanc T (2019) Next generation sequencing technology in Second World victim identification. Forensic Sci Int Genet Suppl Ser 7:123
Zupanič Pajnič I, Obal M, Zupanc T (2020) Identifying victims of the largest Second World War family massacre in Slovenia. Forensic Sci Int 306:110056

9 Molecular Tools for Analysis of Archaeological and Prehistoric Human Bones: A Perspective of Anthropological and Forensic Relevance

Muktikanta Panda, Awdhesh Narayan Sharma, Pankaj Shrivastava, and Hirak Ranjan Dash

Abstract

Bones and the teeth have been considered as the most prominent preserved remains of the primates. The chapter describes conventional tools for the purpose of identification of these materials as well as the present-day DNA typing method as the advanced tool for typing these evidences. The chapter also reviews different steps in molecular typing of these evidences along with problems, challenges and upcoming advancements in the field.

Keywords

Archaeological · Prehistoric · PCR · Human bones · Anthropology · Forensic relevance

M. Panda
Department of Anthropology, Dr. Harisingh Gour Vishwavidyalaya (A Central University), Sagar, M. P., India

DNA Fingerprinting Unit, State Forensic Science Laboratory, Department of Home (Police), Government of Madhya Pradesh, Sagar, M. P., India

A. N. Sharma
Department of Anthropology, Dr. Harisingh Gour Vishwavidyalaya (A Central University), Sagar, M. P., India

P. Shrivastava
DNA Fingerprinting Unit, State Forensic Science Laboratory, Department of Home (Police), Government of Madhya Pradesh, Sagar, Madhya Pradesh, India

H. R. Dash (✉)
DNA Fingerprinting Unit, Integrated High-Tech Laboratory, Forensic Science Laboratory, Bhopal, M. P., India

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_9

9.1 Introduction

Commonly bones and the teeth are the more survived evidences of vertebrates in the archaeological findings (Lee-Thorp 2008). Hence considered as the most prominent preserved remains of the primates during the period of evolution, helps in breaking the mystery of evolutionary pathway of different primates and *Homos*. The archaeological as well as prehistoric bones were mostly found in burial conditions either in decomposed or in fossilized forms. Wide variety of sampled bones, that is complete skeletal remains, partial, burnt or damaged as well as scarped and fragmented samples are recovered from different sites like archaeological sites, mass disaster areas, war zones and fire scenes and so on. Many curiosity like biological gender, age since death, race as well as height of the deceased skeletal remains can accurately determine by the forensic anthropologists, distinguished the human skeletons from the other mammal remains at various sites using different methods; well-advanced DNA techniques add twist within those inquiries (Lundy 1998). Wide range of species recognition as well as ancestry evaluation become possible due to molecular approaches in bone analysis which increases the applicability of the DNA analysis (Ubelaker 2018). Besides, information regarding the past related to the civilization can be known from the bones excavated as the only evidence at different archaeological sites. Information about the past population like their lifestyle and living conditions can be known from morbidity-related changes due to infections like leprosy and syphilis and so on, osteoarthritis and trauma and so on; informs about origin, evolution of many diseases found at present (Ortner 2003). As an archaeological specimen, bones are usually numerous in the form of biological materials than the other soft specimens and those conveniently for vertical diversity study in the point of genetic, considered as highly informative with their ancient DNA (aDNA) analysis (Beraud-Colomb et al. 1995). Chemical analysis helps to detect the dietary habit or nutrition-related information of the past from the skeletal remains (Allmäe et al. 2012). Both organic and inorganic parts in bones and teeth contain different geochemical prints as a result of variations in diet of animal at of isotopic carbon levels (Clementz et al. 2009). Discrimination between human skeletal remains from the nonhumans should be the prime concern in archaeology and forensic biology to save time, resources and money before proceeding for any analysis (Johnson et al. 2017). From anthropological and forensic investigations point of view, proper age estimation of the old bones is quite necessary with minimum biasness. During the death investigation, a forensic anthropologist uses knowledge regarding human body and calculates age by analysis of skeletons with a first priority (Stanojevich 2012). Different methods are employed by the investigators to estimate the accurate age of death. Age-related useful indicators are usually missing or completely destroyed near its all signs in fragmented, gradually rubber or in partial skeleton remains found at different archaeological sites (Kerley 1965). Besides these, still many complete skeletal remains are present in burial conditions, but still problematic in point of the age estimation. There are notable changes which occur at the molecular level in different tissues of the body

(Meissner and Ritz-Timme 2010), estimation and assessment of which can form the base to the prediction of age at death by molecular methods.

This article aimed at gathering the currently used molecular techniques and methodological information in relation to the human bone study and its relevance in anthropology and forensic angles. Before proceeding for the molecular analysis, one should decide wisely that which part of the bone needs to be examined, as this kind of step is more expensive, destruct the sample may affect the result of the study (Ubelaker 2018).

9.2 Conventional Tools and Techniques to Study Historic and Prehistoric Bones

The primary stage examination on old bones focuses on different inquiries like the species determination, explore facts about the age as well as sex of that remains and the time period to which the remains belong to. Both morphological analyses based upon the size and shape and molecular techniques are applied in this field. For example, different methods are useful for the sex determination of recovered skeletal remains (Bruzek 2002; McGlynn 2007). Even successful sex detection rates are found upon adult as well as approximate complete skeletal remains but fail to identify in the cases of poorly preserved skeletons, juvenile samples as well as in the case of incomplete skeletal remains (Black III 1978a, b; Brown et al. 2007; Bauer et al. 2013). Different methodology for the analysis of juvenile skeletons found to be more reliable among males as compared to the females (Weaver 1980; Schutkowski 1993; Scheuer 2002; Galdames et al. 2008; Wilson et al. 2008). Similarly, little information on past medical emergencies and mediations can be predicted from trauma signs on skeletons (Robbins Schug 2017) but not able to go through the deep into the pathophysiology. Recently, Shinde et al. (2018) used a number of morphological techniques even use for counting the age during the death (Brothwell 1981; Lovejoy et al. 1985; Brooks and Suchey 1990) of two grave skeletal remains. Hence here we can say that the morphological analysis of old skeletal remains is still used by the analyser by its uniqueness, irrespective of their drawbacks. Those sort of questions along with others likely discussed here related to archaeological skeleton remains are not solved by the morphological and metric methods quite accurately. In this regard, molecular tools are considered as an alternative but simultaneously run with different methodological as well as biasness-related issues.

9.3 Molecular Tools, Challenges and Solutions

A good number of ancient samples are available preserving little or no DNA, hence some of them are useful with respect to the aDNA analysis (Pääbo et al. 2004). Low amount of very fragmented DNA is recovered from biological remains mixed with different inhibitors related to that environment as a result of different influencing factors related to taphonomy (Matheson et al. 2009). A study result indicates that the

palaeontological and archaeological bone samples, which are unearthed freshly even not washed or treated with any solution, contain more DNA approximately sixfold and yield authentic sequence by twofold (Pruvost et al. 2007). Trace amount of DNA is preserved in different ancient biological remains with different types of continuous degradation over time (Rohland and Hofreiter 2007).

Passes of time and interaction with the surrounding environment either make fragment or modify the threads of genome made up of four-letter bases A, T, G and C (Linderholm 2016). Ancient DNA (aDNA) needs to be handled with more care because of their fragile in nature to minimize damage during downstream processing (Höss et al. 1996; Hofreiter et al. 2001; Hansen et al. 2006). aDNA-related studies face the challenging issues like most of the samples are contaminated with the outside DNA and absent of result during their analysis (Pääbo et al. 2004). A sample contains DNA which tends to be contaminated at any stage, at its finding or at the time of excavation or during the storage or at laboratory during the downstream proceedings (Skoglund et al. 2014). To avoid contamination-trained personals along with sophisticated laboratories as well as advanced instrumentation are necessary for effective study results (Cooper and Poinar 2000; Mulligan 2005). Sample contamination can easily be identified as well as quantified by modern methods (Malmström et al. 2007). Also absence of cellular assemblies in the quite old specimens due to different modifications in biochemical level may create problems to find the remained DNA from aqueous part during the extraction (Geigl 2002). Hence, the molecular methods used for the analysis of ancient bone remains are involved with many methodological as well as processing challenges. To overcome such problems, many advancements of the technologies in terms of low template specificity and/or inhibitor tolerance capability have been developed over the time which has been discussed in detail below.

9.3.1 DNA Isolation and Varied Techniques

Different types of protocols are followed for aDNA extraction like silica binding, use of membrane (Leonard et al. 2000) and alcohol-based methods like isopropanol or ethanol precipitation (Kalmár et al. 2000). Before proceeding for any analysis, a number of methods of DNA extraction from the ancient bone were performed, which involves the common steps like incubate bone powder in buffer (Pääbo 1990; Cattaneo et al. 1995; Hänni et al. 1995a, b), followed by phenol–chloroform extraction (Pääbo 1990; Hänni et al. 1995a, b; Perry et al. 1988) or either dialysed with EDTA and Tris–HCl buffer (Hänni et al. 1995a, b) and finally concentration DNA yields by isopropanol/ethanol precipitation (Hänni et al. 1995a, b; Cattaneo et al. 1997) or by microconcentrators (Pääbo 1990; Yang et al. 1998); can eluted (washed) from silica pellet after separated with silica suspension (Evison et al. 1997; Prado et al. 1997). A study indicates that during the demineralization of bone prior to extraction of DNA by different steps, some DNA is released into the supernatant (Schwarz et al. 2009). From a number of methods developed for the DNA extraction from ancient bone samples, the precipitation by ethanol or isopropanol is considered

as the most appropriate method for this purpose (Kalmár et al. 2000). Maximum aDNA extract can be performed at room temperature from ancient bone and teeth powder, treating by buffer-containing proteinase K and EDTA, where guanidinium thiocyanate used in solution at high concentration to bind the DNA to silica; this minimizes the further degradation and reduces the PCR inhibitors (Rohland and Hofreiter 2007). An improved silica-in-solution method has advantage over the previous version in respect to extracting maximized endogamous tiny and fragmented DNA up to four times due to degradation over time and method has its significance for the genome-based study of Bronze Age skeletal remains with large sample size (Allentoft et al. 2015).

Besides those manual methods, several automation techniques are used for DNA extraction of old bone samples. These minimize the time of extraction, reduce the chance of contamination of environmental DNA with the endogenous DNA present in the bone samples. AutoMate Express an automated system used by modified Full Demineralization (FD) and No Demineralization (ND) protocol for DNA extraction from old skeletal remains efficiently by minimizing the time length and contamination (Pajnič et al. 2016). A recent study indicates that an automated DNA extraction system called Chemagic Magnetic Separation Module I successfully reduces the amount of PCR inhibitor during the elution of DNA from decomposed or old bone samples in comparison with routine organic method of DNA extraction, as the amount of extraction is quite same by both methods (Hong et al. 2017).

Always the aDNA mixed with endogenous DNA as well as DNA from the surrounding environment (Shapiro and Hofreiter 2014). Territorial environment always decides how much endogamous DNA is present in the sample along with playing a vital role for its preservation (Sawyer et al. 2012). Though most of the ancient samples are recovered in degraded conditions, their mt-DNA remains in significantly intact form due to their huge number, that is, $\sim 10^2$ to 10^3 copies per living cell (Hagelberg and Clegg 1991). Perhaps, the research on molecular level of ancient sample was started on the remains of Quagga belongs to the extinct family of Equus (Higuchi et al. 1984). Later, mt-DNA was extracted successfully from different ancient human samples like skin of Egyptian mummy (Pääbo 1985) as well as bone (Horai et al. 1989; Hagelberg et al. 1989). However, the applicability of using mtDNA is limited only to matrilineal lineage analysis.

9.3.2 PCR Process and Its Challenges

Successful operation of polymerase chain reaction (PCR) (Mullis and Faloona 1987) brought an evolution in biology dealing with the nucleic acids, open new chances for the analysis of aDNA as well as fresh DNA. This amplification technique produces sufficient copy of supplied segment of nucleic acid (DNA) called as amplicon even from a trace amount of targeted template. During PCR of different extracted DNA, many challenges come out most often due to the remained inhibitors and degraded or fragmented amplicons. Hence, sometimes PCR optimization is necessary when normal procedure is not able to harvest the copy of chosen amplicon (Lorenz

2012). During short tandem repeat (STR) examination, autosomal and Y markers face different difficulties like the amount of DNA and level of degradation associated with the former, where the latter only by the degradation level (Putkonen et al. 2010).

9.3.2.1 Role of PCR Inhibitors

PCR is an in vitro reaction, which involves the action of enzymes and can be influenced by the presence or absence of chemical molecules present during this reaction. There is a diverse range of inhibitors found. These chemical compounds mostly originate either during the process of nucleic acid extraction or as a matter of exogenous contamination. Additionally, these organic or inorganic compounds may vary in their physical state and one or more types of these compounds may be found to be present in a single reaction. To overcome this problem, different protocols need to be followed according to the medium or matrix used to remove these inhibitors prior to PCR set up (Schrader et al. 2012). Majority of inhibitors involved in PCR process are organic in nature, that is urea, phenol, ethanol, different polysaccharides, sodium dodecyl sulphate, proteinases, melanin as well as different proteins, such as myoglobin, haemoglobin, collagen protein, lactoferrin, immunoglobin G, bile salts, tannic acid and humic acids (Rossen et al. 1992; Rådström et al. 2004). In the forensic, archaeological as well as palaeontological samples humic acid mainly act as PCR inhibitor as samples are submerged in soil or in complete burial condition at different deposits (Baar et al. 2011). Besides, calcium ion (Ca^{++}) is an inorganic PCR inhibitor found in the bone samples.

More amount of PCR inhibitors are present in the ancient bones as well as teeth samples than other fresh samples (Kalmár et al. 2000), which may cause problems during the amplification process post purification of DNA. The specimen which is recovered from burial conditions is associated with different common PCR inhibitors such as humic acids, fluvic acids, collagen, chondroitin, heavy metals as well as chemically modified saccharides (Monroe et al. 2013). Genome-related examination of forensic samples are considered to be more challenging due to the presence of PCR inhibitors coextracted along with DNA, creating anomalies during downstream processing such as reduced peak height of short tandem repeats (STR) markers, allele dropouts, enhanced stutter and reduced sensitivity of PCR process. Dilution of the templet is necessary (automatically inhibitor will dilute) when the presence of PCR inhibitor is expected (Wilson 1997). But in case of less templet concentration as well as more degraded sample, dilution is not acceptable as further dilution minimizes the templet quantity. In order to overcome these problems, prior enumeration of PCR inhibitors should be carried out mostly during the handling of ancient DNA (King et al. 2009). It is of huge challenge to amplify spoiled and old specimens by PCR as the rate of success is decided by the quantity of DNA present, types and amount of inhibitor associated with the amplicon and the degradation index of DNA (Putkonen et al. 2010). Additionally, when toxic metals such as copper are present as a PCR inhibitor in DNA eluted from archaeological bones and other old degraded bone sample, an extra step of purification called size exclusion chromatography is mostly carried out to purify the aDNA fragments by removing ions as well as other inhibitors (Matheson et al. 2009). Recently, an engineered

polymerase enzyme 2D9 has been reported to show promising result in amplifying DNA samples in presence of some known inhibitor compounds, that is humic acid, bone powder, coprolite, peat extract, clay-rich soil, cave deposit and tar (Baar et al. 2011). A study result indicates that three commercially available polymerases: Ex Taq HS, PicoMaxx HF and FastStart Taq more efficiently amplify the highly degraded human aDNA amplicons containing impurities like PCR inhibitors in the reaction medium (Kim et al. 2015). Another study shows the efficient use of a plant enzyme KAPA3G for amplification of low-quality fragmented DNA samples extracted from ancient buried bone (Nilsson et al. 2016).

9.3.2.2 Degraded DNA Samples

Bones are considered to be the ideal samples for the aDNA extraction and sometimes used as the one and only means as a last preserved remains (Gorgan et al. 2013). Fragmented DNA is commonly observed from ancient biological specimens. The initial work on property study of aDNA indicates that all the aged samples dated back up to ~13,000 years contain DNA having fragmented pieces in between 40 and 500 bp due to degradation (Pääbo 1989). Different physico-chemical as well as biological parameters are responsible for the degradation of DNA in different preserved remains. The quantity as well as quality of recoverable DNA fragments from old bones mainly depend on the conditions of the bones and tooth specimens preserved since death (Adler et al. 2011), with limiting factors including soil chemistry surrounding to the buried remains, environmental exposures such as humidity, sunlight and temperature (Alaeddini et al. 2010). The rise in humidity and temperature causes damage to DNA only in dead cells soon after death (Gorgan et al. 2013). Besides, environmental interactions with humic acid and pH can reduce the DNA quantity in bone samples. A study indicates that dehydration of tissue in the presence of salt minimizes the DNA damage by reducing the speed of DNA strand breaking (Shved et al. 2014). Other biological factors contributing to DNA damage include action of microbes, soil invertebrates and intracellular enzymes (Binladen et al. 2006). After the death of organisms, osteolysis occurs by diverse group of microorganisms of both endogenous and environmental origin, which can contribute to the damage of osteoblast DNA (Booth 2016). Besides the presence of pollutants, different watery environments also affect the rapidity of degradation in bone samples (Yeap et al. 1998).

DNA gets absorbed into hydroxyapatite, an inorganic bone constituent which helps in its storage for a longer period of time mostly by formation of crystalline structures (Okazaki et al. 2001; Feng 2009). DNA is prone to various alternations in its chemistry due to oxidation and hydrolysis process (Lindahl 1993; Dabney et al. 2013), which affects the extraction and post-analysis process. Reduction of fragment size of DNA is one of the common alteration of DNA structure besides blockage of action of DNA polymerase, incorporation of unusual bases due to strand lesions. The enzyme endogenous exonuclease facilitates the nuclear DNA breakdown quickly only after the damage in the nuclear membrane (Graham 2007). Most of the studies carried out on DNA degradation correlating PCR process and aDNA primarily concerns on the damage created by polymerase bypassing the amplicon (Hofreiter

et al. 2001; Stiller et al. 2006), whereas less attention is given to DNA damage, which is non-by-passable during PCR process (Hansen et al. 2006).

Unlike nuclear DNA, mtDNA preserves its structural integrity even after elongated time period due to its circular structure that minimizes the exonuclease activity (Butler 2005), smaller size as well as numerous copies present in a single cell (Schwarz et al. 2009; Gorgan et al. 2013). mtDNA can act as the convenient sample for profile generating from highly degraded sample where the STR and mini-STR analysis are tend to be failed (Coble et al. 2004), suitable for prehistoric as well as aDNA studies from the bone samples. This method is successfully implemented during the study of bone samples related to the American war (Holland et al. 1993).

As SNPs form the 90% of human DNA variation (Collins et al. 1998), found in autosomal, X, Y and mitochondrial DNA. SNP study has a greater advantage in the forensic investigation as degraded ample contains very short length fragments and is able to initiate 45–55 bp amplicon (Kidd et al. 2006). This maximizes the loci number in different degraded samples. Different studies on nuclear SNPs were done on aDNA extracted from human bones dated to archaeological as well as prehistoric time (Burger et al. 2007; Lalueza-Fox et al. 2007; Watson and Lockwood 2009). The first multiplex autosomal SNP analysis was done on bronze and Iron Age human bones containing highly degraded aDNA successfully correlated with different traits responsible for the pigmentation and biogeographical ancestry tracing (Bouakaze et al. 2009).

9.3.2.3 PCR/Mini Filer Kit

High sample degradation creates challenges for the STR analysis. That result in deformities like the loci or allele drop out (Gill et al. 2005) and incomplete profile generation. These problems are minimized with mini-STR as well as autosomal SNP, used as most suitable markers for the highly fragmented DNA analysis (Fondevila et al. 2008). Another problem associated with the STR analysis is that it provides little or no value to the analysis if the generated STR profile is not matched with the existing databases (Bardan et al. 2018). Mini-STR analysis shows to be advanced over the STR when the trace amount of DNA is available (Welch et al. 2011) and also susceptible to hostile kind of degradation in different conditions, if having length <120 bp (Fondevila et al. 2008). The rate of successfully generating the DNA profile directly correlates with the length of the given amplicon (Bender et al. 2004). Use of mini-STR analysis as compared to STR maximizes the profile success to approximately threefold, even generates full or most completed profile (Alonso et al. 2005).

9.3.3 DNA Sequencing and/or Genotyping

The exact nitrogenous base arrangement of a piece of DNA can be determined by the sequencing where the genotyping investigates the type of genetic variations present in the studied genomic strand. Different well-advanced as well as automated genotyping/sequencing techniques were developed like the electrophoresis mediated

through the microchannel then replaced by capillaries and chip-based microfabricated format for DNA sequencing. Degraded DNA sample is suitable for the STR analysis due to small-sized fragments ~100–500 bp (Butler 2007), conventional tool for the aDNA study as well as forensic evidences. Available of multiplex PCR kit for different sized STRs increased the feasibility of genotyping by means of CE where the study of DNA is compared with different standard allelic ladder (Butler 2007). During the read of long DNA fragments >1000 by capillary electrophoresis (CE), the success rate of separation for sequencing/genotyping depends upon the suitable parameters like the temperature of the electrophoretic medium as well as the separation voltage (Karger and Guttman 2009).

9.3.3.1 NGS and aDNA

First DNA sequencing done in 1975 was able to read one sequence at a time (Sanger and Coulson 1975). Recently, the limitation of Sanger sequencing is overcome by the introduction of next-generation sequencing (NGS) advanced over the previous method due to millions of sequence which can be read simultaneously and the time as well as the coast is saved. Most of the aDNA is small or fragmented due to the high rate of degradation, where NGS needs short templet, more suitable for the analysis (Linderholm 2016). A very trace amount of endogenous DNA is found from sample of ancient sources like historical, archaeological, prehistoric sites (Shapiro and Hofreiter 2014), where elution needs more time, coast and efforts. Before proceeding for the NGS, construction of aDNA library is crucial. To yield more endogenous DNA for library construction, ultimately NGS Hybridization capture method is very useful that minimizes the proportion of DNA from different environmental sources (Ávila-Arcos et al. 2011). The major problem that arises during the aDNA analysis is the chance of contamination. Major contamination source is the modern hominid DNA, which comprises near 80% (Wall and Kim 2007). With NGS, it is possible to directly assess the contamination in aDNA from modern DNA sources by measuring the heterozygosity; only applicable to haploid genomic regions like the X chromosome on male and mitochondrial genome, not able to assess the same from both the female X chromosome and their nuclear genome (Furtwängler et al. 2018). Same study indicates that, there is a negative relationship exists between the ratio of mtDNA to nuclear DNA with the rise in amount of endogenous DNA, which manipulates the appraisal of contamination in both mtDNA and ncDNA in males (Furtwängler et al. 2018). Different ancient Neanderthal genome analyses were successfully carried out with the help of NGS (Green et al. 2008; Maricic and Pääbo 2009). Standardization of molecular clock, interspecies evolutionary relationship, their geographical origin and past demography-related information can be traced from the mtDNA where nuclear aDNA added new knowledge such as to find extinct phenotypes, measure the degree to which admixture occurred and observe the selection pressure on the species; advent of next-generation sequencing (NGS) or parallel sequencing increases the accuracy and efficiency to different investigation related to aDNA (Linderholm 2016). But in ancient sample, the endogenous DNA amount is very low; hence, re-cheque with Sanger method is necessary after sequencing with parallel method for the result

confirmation, one of the major drawbacks of NGS (Welker et al. 2014). NGS can also be useful for the fast analysis of mtDNA, which minimizes the coast as well as the time (Loreille et al. 2011), best suited for the archaeological as well as prehistoric bone specimens where the mtDNA along with highly degraded or no nDNA is preserved.

9.3.4 DNA Methylation Studies

DNA methylation is a kind of biochemical modification of CG di-bases prone to the Vth position of pyrimidine and dynamicity of the modification is decided by the environmental factors as well as the nature of strand sequence (Schneider et al. 2014). Information only relies on base sequence of DNA strand, sometimes not sufficient for phenotypic evidences, where the epigenetic modification provides phenotypic evidence by means of pattern study on methylation (Vidaki et al. 2013). Outline of aDNA methylation gives a twist to the dark mystery of human evolution. For methylation study on DNA extracts from archaic samples, it is not possible through the standard ongoing methods because aDNA is damage prone (Gokhman et al. 2014; Orlando and Willerslev 2014). A study indicates that C-methylation is inversely related to the preservation of aDNA, easy to detect the same from a wide range of quite recent old samples including human bone remains having good quantity of preserved nDNA (Smith et al. 2015). Subsequently suggested that the bisulphite sequence can be useful for the epigenetic study on ancient human bone remains (Smith et al. 2015). A study on extinct species, Neandertal and the Denisovan, by plotting their complete DNA methylation map with a new method concludes that the HOXD cluster can be able to describe the structural variation between archaic and *Homo sapiens* (Gokhman et al. 2014). A study revealed that the sequence and promoter region of gene is responsible for limb development, that is HOXD9 and HOXD10 are found to be methylated in aDNA samples and found as un-methylated form in the present population (Zhenilo et al. 2016), important in the field of evolutionary study.

9.4 Application of Molecular Tools

9.4.1 Estimation of Age

It is very crucial to calculate approximation of age in accuracy of archaeological remains like bone during the examination process to provide a good insight knowledge. Different methodologies are useful for the scene by adopting according to the material used for the investigation (Marquez-Grant 2015). Initially, analysis on bone and teeth gives an idea about the age prediction in the field of forensics (Freire-aradas et al. 2017). Morphological examinations to detect the age of a bone specimen stand with several biasness, which can be revealed by the molecular techniques. A standardized method that determines the age called carbon-14 (^{14}C) is applicable for

a variety of biological remains up to 50,000 year old (Bronk Ramsey 2008), based upon the principle that after the death of the living organism, the ^{14}C decay starts to nitrogen-14 (^{14}N) (Godwin 1962). Mutation clock analysis was proofed as useful method for the age estimation of old bones that successfully dates the Denisova finger bone, which was not effectively dated by the carbon-14 method (Meyer et al. 2012a). Similarly, another study successfully applied the concept of recombination clock to date the age as well as to find generation gap of a past Eurasian population from their available genome (dated to upper Palaeolithic period); deciphers the outstanding link between age data comes from ^{14}C and genome-based method (Moorjani et al. 2016). Most of the time a new date analysis is not considered the data related to date at different pre-studied sites, though Palaeolithic samples analysis runs simultaneously with controversy related to the correctness as well as their reliability of used methods (Higham 2011). In a study, some Neanderthal bone remains were re-dated by a method called preparative high-performance liquid chromatography (Prep-HPLC) by extracting amino acid hydroxyproline, added fact to the modern human arrival time in Eastern Europe (Devièse et al. 2017).

9.4.2 Determination of Sex of the Individual

Mystery about the past social structure and stratification can be better understood if light is shed on the biological sex of the archaeological human remains. Morphological methods for sex identification are best suited for complete or less fragmented skeletal remains of adults but fail to detect in juvenile skeletons as well as more fragmentary skeletons; molecular tools provide reliable and alter method for the same. Sex identification with aDNA analysis of humans is used in different ways: by the study of amalogenin gene (Svensson et al. 2012; Shaw et al. 2015), zinc finger protein genes (of X or Y) (Svensson et al. 2008) and sequence analysis of Y or W chromosome (Cappellini et al. 2004; Allentoft et al. 2010). Molecular tools for sex identification sometimes associated with more failure rate as aDNA from archaeological samples contains reduced amount of sex-specific markers with single copy (Alonso et al. 2004). PCR multiplex Genderplex can be useful for the sex identification of medieval bones from the preserved DNA remains in the samples (Codina et al. 2009). Amelogenin markers are dominantly used as sex-specified marker but prone to contamination by modern DNA as well as associated with high in allele dropout rate in old samples (Skoglund et al. 2013). Possibilities of unfair result assumes with the presence of trace amount of contamination with exogenous DNA, which requires more exercise like huge number of clones to be proceed for the sequence analysis (Helgason et al. 2007). Even in the presence of large amount of contamination by means of modern DNA in extremely degraded sample retrieved from the aDNA, sex can be properly defined by a specific feature, that is, the cytosine–deamination signature mark (Skoglund et al. 2013). By distinguishing amalogenin gene present on the X chromosome from the pseudo one found on the male/Y chromosome, sex can be identified (Götherström et al. 1997; Morrill et al. 2008; Gibbon et al. 2009; Quincey et al. 2013). aDNA analysis has become more

relaxed with the help of advent of high-throughput shotgun sequencing in large scale (Green et al. 2006; Meyer et al. 2012a). Chromosomal sex can be identified accurately from ~70,000 years old sample having 100,000 human shotgun sequences, that is in quite low depth by means of evaluating sequence alignment ratio between X and Y chromosome (Skoglund et al. 2013). A study on Neanderthal genome is the proposed method for sex determination by eliminating different homologous base sequence region among the sex chromosomes (X and Y) (Green et al. 2010). An initial step for onsite DNA analysis method was developed, which is able to extract aDNA with less contamination, effectively amplify <50 pg of same by means of microfluidic device and identified sex through capillary gel electrophoresis examination. That was supported by different other analyses related to the anthropological, Amelogenin and male (Y) chromosome examination (Parton et al. 2013). Another analysis of old DNA retrieved from juvenile skeletal remains dated back to medieval period with the help of PCR and STR profiling pointing towards the high male mortality rate, although unable to detect sex with high conformity in some samples (Tierney and Bird 2015).

A co-amplification method that need not use sequencing was developed recently to detect the sex of archaeological remains in Turkey. That amplify highly repetitive DNA (hrDNA) and mtDNA fragments, where the heterogametic sex, that is, female can be detected by the W chromosome amplification, while for internal positive control as well as to screen the false-negative outcome, mtDNA fragment was used (Speller and Yang 2016).

9.4.3 Migration Studies

Initial studies in genetic analysis are based on the phylogeographic method (Avise et al. 1987). This approach implemented by phylogenetic tree analysis on aDNA informative in the sense to find the dynamicity in speciation process, movement pattern of a population, and their extension-drop scenario of the interested population (Leonardi et al. 2017). Large investigations in population genetics and different complex issues like the migration route and admixture pattern can be studied from the nDNA with the help of NGS (Linderholm 2016).

9.4.4 Gene Diversity Studies

Phylogenetic as well as the phylogeographical data can be gathered with the help of past DNA, unable to retrieve the information from species presently found (Allentoft et al. 2015). Genetic difference between the two European economical groups, that is, Hunter–gatherer with farmers was studied with aDNA from past bone remains (Skoglund et al. 2012). It is possible to compare the modern human diversity with the ancient extinct population of Homo, like Denisovans and Neanderthal (Noonan et al. 2006; Meyer et al. 2012a) can be revealed by analysing whole genome or by total mtDNA genome (Fu et al. 2013). aDNA analysis by new hydrolysis

probe-based real-time PCR assay of different medieval bone remains related to the immunological functioning allele named CCR5-Δ32, which indicates that the allele frequency for the same in central Europe remained unchanged over a thousand years, which signifies towards the absence of strong positive selection pressure and earliest origin of the same (Bouwman et al. 2017).

9.4.5 Kinship Study

Using nDNA analysis, it is possible to construct kinship relationship both in matrilineal and patrilineal line (Lee et al. 2010), highly relevant in anthropological study. SNP locus characterized by the low mutation rate, hence, is suitable for kinship analysis. A study reveals that in ancient individuals, first degree of kins relationship can identify easily with HID-Ion AmpliSeqTM Identity Panel that studies nuclear SNPs, but needs a good preserved DNA for the same; where for the elimination of direct female lineage relationship, mitochondrial genome capture enrichment tool is useful (Juras et al. 2017). Lineage Y-chromosomal markers and mtDNA markers are useful in tracing to correlate relationship in both paternal and maternal line, respectively (Haak et al. 2008; Malmström et al. 2012; Meyer et al. 2012a; Lee et al. 2014; Parson et al. 2015; Rothe et al. 2015; Alt et al. 2016). aDNA examination is the most accurate and direct way to testify biological affine relationships by means of biparentally hereditary markers, that is, single tandem repeats (STR) (Gamba et al. 2011; Baca et al. 2012; Juras et al. 2017), but the use of STR is quite inefficient in result point of view as aDNA is generally found in highly fragmented form (Deguilloux et al. 2014). It is important to investigate kin analysis and burial practices of a past population in an interdisciplinary style, likely by involving the anthropological, archaeological and molecular information for an unbiased result (Juras et al. 2017).

9.4.6 Ancient Disease and Epidemiology

Different upcoming disease perspective can be best understood by means of influencing them only after understanding the ancient epidemiology (Cooper et al. 2015). Ancient mass epidemiological disasters causing microbes involved in great plague (Bos et al. 2011) and Spanish flu (Taubenberger et al. 2001), which are explained by the aDNA studies. SNP-based analysis was carried out in aDNA of *Myobacterium leprae* recovered from the medieval bone remains in Europe vision insight into the transmission pattern of leprosy and the evolution of the bacteria (Watson and Lockwood 2009). Ancient microbial DNA is well preserved at dental calculus, which provides information on their evolutionary diversity and relates them with the behaviour and diet pattern (Adler et al. 2013; Warinner et al. 2014a, b; Cooper et al. 2015). Epigenetic analysis indicates that the differentially methylated regions or DMRs are considerably additionally linked to disease (Gokhman et al. 2014).

9.4.7 Species Recognition

mtDNA obtained from modern or old bone is useful in personal and species recognition (Lorente et al. 2001). HVR-I and HVR-II region of mt-DNA is useful for Hominoid ancestry (Watson et al. 1996), which is rich in cytosine (Lutz-Bonengel et al. 2004; Chen et al. 2009). These provide hint for the unrelated person recognition (Amer et al. 2017) due to heteroplasmy nature of mtDNA-HVS (Salas et al. 2001; Mabuchi et al. 2007). An aDNA recovered from old bone remains helped anthropologist to discover new species of genus Homo called 'Denisovans' an archaic form, having distinct pattern of evolutionary history in comparison with *Homo sapiens* and *Homo Neanderthals* (Reich et al. 2010).

9.4.8 Past Demography

Sex identification by aDNA analysis at a specific archaeological site, Sand Canyon Pueblo, indicates the balanced male and female ratio of the area in the past, which sheds light on demographic history (Speller and Yang 2016). For the first time, by interpreting the linear connectivity of genetic admixture age of Neanderthal genome with carbon dating data, it was found that a generation gap of 26–30 years is associated with a prehistoric population (Moorjani et al. 2016).

9.5 Proteomics Analysis from Preserved Biomolecules in Ancient Bone

Besides the study of aDNA, a completely new trend is developed for the ancient protein (aProtein) analysis from the ancient preserved samples. Different fragmented fossilized mineralized tissues, like bone, provide information from its protein sequence where the aDNA extraction is hardly possible or absent completely (Welker et al. 2015). Different geographical locations like Near East with very hot and arid climate, where remnants are found related to the human biological and cultural evolution possesses less chance of retrieval of biomolecules due to the less probability with their survivalist (Bar-Gal et al. 2003; Larson et al. 2007; Buckley and Wadsworth 2014). These type of problems in molecular archaeology are somehow solved by ongoing trend of proteomic analysis alternation to aDNA study with respect to preserved tissue remnants, time period as well as the sample area (Cappellini et al. 2014). It is possible to use non-collagen proteins (NCPs) for phylogenetic analysis as well as identification species where the result is not obtained from the bone collagen proteins type-I as their age-old survival (Buckley and Wadsworth 2014). aProtein sequence of extinct hominins can be read by tandem mass spectroscopy (TMS) helpful in tracing their phylogenetic placement (Welker et al. 2016). Also, additive information, like food pattern, patho-history, epidemiology as well as in vivo physiology, from different samples dated back to archaeological as well as palaeontological period can be gained from the focus in sequence

study (Warinner et al. 2014a, b; Hendy et al. 2016; Procopio et al. 2017). One of the major limitation of aProtein is that provides quite less information regarding the phylogenetic analysis as compared to the whole aDNA study suitably applicable to quite long time scale (Welker 2018). Different lifestyle information of the ancient Edo citizens, like obesity, stress, disease and starvation, was gathered from the first shotgun proteomic analysis of archaeological bones (Sawafuji et al. 2017). As a whole, this provides a platform for proteomic analysis of different hominin fossil contains contaminated, less or no endogenous DNA, which focus light into dark process of their evolution. Single amino acid polymorphisms (SAPs) analysis can be possible from late Pleistocene fossilized bones, which help in differentiating Neanderthals and Denisovans as well as *Homo sapiens* having recovered aDNA (Meyer et al. 2012a, b; Castellano et al. 2014; Brown et al. 2016; Welker et al. 2016). New methods of data analysis by targeting amino acid hydroxyproline of archaeological bones tend to be more informative over ^{14}C method, also effective to eliminate the modern contamination from bone collagens (Devièse et al. 2017).

9.6 Future Prospects and Conclusion

A Miniplex system was developed by adjoining multi-SNPs of mitochondrial, Y and autosomal DNAs applied for mysterious human samples, able to decipher phenotypic feature like eye iris colour, biological sex, biogeographic ancestry as well as to calculate haplogroup based on the Y and mt-DNA having the great forensic importance (Bardan et al. 2018), useful for the aDNA analysis in a encouraged way in the near future. Recently a new branch called ‘Paleovirology’ was developed, a subfield of aDNA that gathers the information regarding the evolution of the virus (Tsangaras and Greenwood 2018). More methodological innovation is required for the fine evidences. Like aDNA, aRNA is not an explored area of researchers focus, but able to decipher the past ecology by means of describing genomic condition at the time of death of an individual possible by analysing the proportion of mRNA, rRNA and tRNA (Smith and Gilbert 2018). Some limited attempt, that is, Venanzi and Rollo (1990) and Keller et al. (2017), was done to extract aDNA from the human remains. The future aspect on Paleo-transcriptomics depends upon the methodological innovation, sample availability and preserved in favourable condition suited for their stability against the diagnosis.

References

Adler CJ, Haak W, Donlon D, Cooper A, Genographic Consortium (2011) Survival and recovery of DNA from ancient teeth and bones. J Archaeol Sci 38(5):956–964

Adler CJ, Dobney K, Weyrich LS, Kaidonis J, Walker AW, Haak W, Bradshaw CJ, Townsend G, Sołtysiak A, Alt KW and Parkhill J (2013) Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and industrial revolutions. Nat Genet 45(4):450

Alaeddini R, Walsh SJ, Abbas A (2010) Molecular studies of time-and environment-dependent effects on bone DNA survival. Aust J Forensic Sci 42(3):211–220

Allentoft ME, Bunce M, Scofield RP, Hale ML, Holdaway RN (2010) Highly skewed sex ratios and biased fossil deposition of moa: ancient DNA provides new insight on New Zealand's extinct megafauna. Quat Sci Rev 29(5–6):753–762

Allentoft ME, Sikora M, Sjögren KG, Rasmussen S, Rasmussen M, Stenderup J et al (2015) Population genomics of bronze age Eurasia. Nature 522(7555):167

Allmäe R, Limbo-Simovart J, Heapost L, Verš E (2012) The content of chemical elements in archaeological human bones as a source of nutrition research. Pap Anthropol 21:27–49

Alonso A, Martin P, Albarrán C, Garcia P, Garcia O, de Simón LF, Fernández-Piqueras J (2004) Real-time PCR designs to estimate nuclear and mitochondrial DNA copy number in forensic and ancient DNA studies. Forensic Sci Int 139(2–3):141–149

Alonso A, Martín P, Albarrán C, Garcí P, De Simón LF, Iturralde MJ et al (2005) Challenges of DNA profiling in mass disaster investigations. Croat Med J 46(4):540–548

Alt KW, Zesch S, Garrido-Pena R, Knipper C, Szécsényi-Nagy A, Roth C et al (2016) A community in life and death: the late Neolithic megalithic tomb at alto de Reinoso (Burgos, Spain). PLoS One 11(1):e0146176

Amer SA, Al-Harthi HS, Refaat AM, Bakdash A, Kassab AC (2017) Identification of human bone remains by autosomal STRs and mitochondrial DNA SNPs. J Hard Tiss Biol 26(2):111–118

Ávila-Arcos MC, Cappellini E, Romero-Navarro JA, Wales N, Moreno-Mayar JV, Rasmussen M et al (2011) Application and comparison of large-scale solution-based DNA capture-enrichment methods on ancient DNA. Sci Rep 1:74

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE et al (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annu Rev Ecol Syst 18(1):489–522

Baar C, d'Abbadie M, Vaisman A, Arana ME, Hofreiter M, Woodgate R et al (2011) Molecular breeding of polymerases for resistance to environmental inhibitors. Nucleic Acids Res 39(8): e51–e51

Baca M, Doan K, Sobczyk M, Stankovic A, Węgleński P (2012) Ancient DNA reveals kinship burial patterns of a pre-Columbian Andean community. BMC Genet 13(1):30

Bardan F, Higgins D, Austin JJ (2018) A mini-multiplex SNaPshot assay for the triage of degraded human DNA. Forensic Sci Int Genet 34:62–70

Bar-Gal GK, Ducos P, Kolska Horwitz L (2003) The application of ancient DNA analysis to identify neolithic caprinae: a case study from the site of Hatoula, Israel. Int J Osteoarchaeol 13 (3):120–131

Bauer CM, Niederstätter H, McGlynn G, Stadler H, Parson W (2013) Comparison of morphological and molecular genetic sex-typing on mediaeval human skeletal remains. Forensic Sci Int Genet 7(6):581–586

Bender K, Farfán MJ, Schneider PM (2004) Preparation of degraded human DNA under controlled conditions. Forensic Sci Int 139(2–3):135–140

Beraud-Colomb E, Roubin R, Martin J, Maroc N, Gardeisen A, Trabuchet G, Goossens M (1995) Human beta-globin gene polymorphisms characterized in DNA extracted from ancient bones 12,000 years old. Am J Hum Genet 57(6):1267

Binladen J, Wiuf C, Gilbert MTP, Bunce M, Barnett R, Larson G et al (2006) Assessing the fidelity of ancient DNA sequences amplified from nuclear genes. Genetics 172(2):733–741

Black TK III (1978a) Sexual dimorphism in the tooth-crown diameters of the deciduous teeth. Am J Phys Anthropol 48(1):77–82

Black TK III (1978b) A new method for assessing the sex of fragmentary skeletal remains: femoral shaft circumference. Am J Phys Anthropol 48(2):227–231

Booth TJ (2016) An investigation into the relationship between funerary treatment and bacterial bioerosion in European archaeological human bone. Archaeometry 58(3):484–499

Bos KI, Schuenemann VJ, Golding GB, Burbano HA, Waglechner N, Coombes BK et al (2011) A draft genome of Yersinia pestis from victims of the Black death. Nature 478(7370):506

Bouakaze C, Keyser C, Crubézy E, Montagnon D, Ludes B (2009) Pigment phenotype and biogeographical ancestry from ancient skeletal remains: inferences from multiplexed autosomal SNP analysis. Int J Leg Med 123(4):315–325
Bouwman A, Shved N, Akgül G, Rühli F, Warinner C (2017) Ancient DNA investigation of a medieval German cemetery confirms long-term stability of CCR5-Δ32 allele frequencies in Central Europe. Hum Biol 89(2):119–124
Bronk Ramsey C (2008) Radiocarbon dating: revolutions in understanding. Archaeometry 50 (2):249–275
Brooks S, Suchey JM (1990) Skeletal age determination based on the os pubis: a comparison of the Acsádi-Nemeskéri and Suchey-Brooks methods. Hum Evol 5(3):227–238
Brothwell DR (1981) Digging up bones: the excavation, treatment, and study of human skeletal remains, 3rd edn. Cornell University Press, Ithaca
Brown RP, Ubelaker DH, Schanfield MS (2007) Evaluation of Purkait's triangle method for determining sexual dimorphism. J Forensic Sci 52(3):553–556
Brown S, Higham T, Slon V, Pääbo S, Meyer M, Douka K et al (2016) Identification of a new hominin bone from Denisova cave, Siberia using collagen fingerprinting and mitochondrial DNA analysis. Sci Rep 6:23559
Bruzek J (2002) A method for visual determination of sex, using the human hip bone. Am J Phys Anthropol 117(2):157–168
Buckley M, Wadsworth C (2014) Proteome degradation in ancient bone: diagenesis and phylogenetic potential. Palaeogeogr Palaeoclimatol Palaeoecol 416:69–79
Burger J, Kirchner M, Bramanti B, Haak W, Thomas MG (2007) Absence of the lactase-persistence-associated allele in early Neolithic Europeans. Proc Natl Acad Sci U S A 104 (10):3736–3741
Butler JM (2005) Forensic DNA typing: biology, technology, and genetics of STR markers, 2nd edn. Elsevier/Academic Press, New York
Butler JM (2007) Short tandem repeat typing technologies used in human identity testing. Biotechniques 43(4):Sii–SSv
Cappellini E, Chiarelli B, Sineo L, Casoli A, Di Gioia A, Vernesi C et al (2004) Biomolecular study of the human remains from tomb 5859 in the Etruscan necropolis of Monterozzi, Tarquinia (Viterbo, Italy). J Archaeol Sci 31(5):603–612
Cappellini E, Collins MJ, Gilbert MTP (2014) Unlocking ancient protein palimpsests. Science 343 (6177):1320–1322
Castellano S, Parra G, Sánchez-Quinto FA, Racimo F, Kuhlwilm M, Kircher M et al (2014) Patterns of coding variation in the complete exomes of three Neandertals. Proc Natl Acad Sci U S A 111 (18):6666–6671
Cattaneo C, Smillie DM, Gelsthorpe K, Piccinini A, Gelsthorpe AR, Sokol RJ (1995) A simple method for extracting DNA from old skeletal material. Forensic Sci Int 74(3):167–174
Cattaneo C, Craig OE, James NT, Sokol RJ (1997) Comparison of three DNA extraction methods on bone and blood stains up to 43 years old and amplification of three different gene sequences. J Foren Sci 42(6):1126–1135
Chen F, Dang YH, Yan CX, Liu YL, Deng YJ, Fulton DJ, Chen T (2009) Sequence-length variation of mtDNA HVS-I C-stretch in Chinese ethnic groups. J Zhejiang Univ Sci B 10(10):711–720
Clementz MT, Fox-Dobbs K, Wheatley PV, Koch PL, Doak DF (2009) Revisiting old bones: coupled carbon isotope analysis of bioapatite and collagen as an ecological and palaeoecological tool. Geol J 44(5):605–620
Coble MD, Just RS, O'Callaghan JE, Letmanyi IH, Peterson CT, Irwin JA, Parsons TJ (2004) Single nucleotide polymorphisms over the entire mtDNA genome that increase the power of forensic testing in Caucasians. Int J Leg Med 118(3):137–146
Codina AE, Niederstätter H, Parson W (2009) "GenderPlex" a PCR multiplex for reliable gender determination of degraded human DNA samples and complex gender constellations. Int J Leg Med 123(6):459–464

Collins FS, Brooks LD, Chakravarti A (1998) A DNA polymorphism discovery resource for research on human genetic variation. Genome Res 8(12):1229–1231
Cooper A, Poinar HN (2000) Ancient DNA: do it right or not at all. Science 289(5482):1139–1139
Cooper A, Weyrich LS, Farrer AG (2015) The relevance of ancient DNA to contemporary disease. Pathology 47:S28
Dabney J, Meyer M, Pääbo S (2013) Ancient DNA damage. Cold Spring Harb Perspect Biol 5(7): a012567
Deguilloux MF, Pemonge MH, Mendisco F, Thibon D, Cartron I, Castex D (2014) Ancient DNA and kinship analysis of human remains deposited in Merovingian necropolis sarcophagi (Jau Dignac et Loirac, France, 7th–8th century AD). J Archaeol Sci 41:399–405
Devièse T, Karavanić I, Comeskey D, Kubiak C, Korlević P, Hajdinjak M et al (2017) Direct dating of Neanderthal remains from the site of Vindija cave and implications for the middle to upper Paleolithic transition. Proc Natl Acad Sci U S A 114(40):10606–10611
Evison MP, Smillie DM, Chamberlain AT (1997) Extraction of single-copy nuclear DNA from forensic specimens with a variety of postmortem histories. J Foren Sci 42(6):1032–1038
Feng X (2009) Chemical and biochemical basis of cell-bone matrix interaction in health and disease. Curr Chem Biol 3(2):189–196
Fondevila M, Phillips C, Naverán N, Cerezo M, Rodríguez A, Calvo R et al (2008) Challenging DNA: assessment of a range of genotyping approaches for highly degraded forensic samples. Forensic Sci Int Genet Suppl Ser 1(1):26–28
Freire-Aradas A, Phillips C, Lareu MV (2017) Forensic individual age estimation with DNA: from initial approaches to methylation tests. Forensic Sci Rev 29(2):121–144
Fu Q, Mittnik A, Johnson PL, Bos K, Lari M, Bollongino R et al (2013) A revised timescale for human evolution based on ancient mitochondrial genomes. Curr Biol 23(7):553–559
Furtwängler A, Reiter E, Neumann GU, Siebke I, Steuri N, Hafner A et al (2018) Ratio of mitochondrial to nuclear DNA affects contamination estimates in ancient DNA analysis. Sci Rep 8(1):14075
Galdames ICS, Matamala DAZ, Smith RL (2008) Blind test of mandibular morphology with sex indicator in subadult mandibles. Int J Morphol 26:845–848
Gamba C, Fernández E, Tirado M, Pastor F, Arroyo-Pardo E (2011) Brief communication: ancient nuclear DNA and kinship analysis: the case of a medieval burial in san Esteban Church in Cuellar (Segovia, Central Spain). Am J Phys Anthropol 144(3):485–491
Geigl EM (2002) On the circumstances surrounding the preservation and analysis of very old DNA. Archaeometry 44(3):337–342
Gibbon V, Paximadis M, Štrkalj G, Ruff P, Penny C (2009) Novel methods of molecular sex identification from skeletal tissue using the amelogenin gene. Forensic Sci Int Genet 3(2):74–79
Gill P, Curran J, Elliot K (2005) A graphical simulation model of the entire DNA process associated with the analysis of short tandem repeat loci. Nucleic Acids Res 33(2):632–643
Godwin H (1962) Half-life of radiocarbon. Nature 195(4845):984
Gokhman D, Lavi E, Prüfer K, Fraga MF, Riancho JA, Kelso J et al (2014) Reconstructing the DNA methylation maps of the Neandertal and the Denisovan. Science 344(6183):523–527
Gorgan L, Druică R, Ştefan A, Ciorpac M (2013) Comparative DNA degradation in bones fragments. Int J Crim Invest 3(1):33–43
Götherström A, Lidén K, Ahlström T, Källersjö M, Brown TA (1997) Osteology, DNA and sex identification: morphological and molecular sex identifications of five neolithic individuals from Ajvide, Gotland. Int J Osteoarchaeol 7(1):71–81
Graham EAM (2007) DNA reviews: ancient DNA. Forensic Sci Med Pathol 3(3):221–225
Green RE, Krause J, Ptak SE, Briggs AW, Ronan MT, Simons JF et al (2006) Analysis of one million base pairs of Neanderthal DNA. Nature 444(7117):330
Green RE, Malaspinas AS, Krause J, Briggs AW, Johnson PL, Uhler C et al (2008) A complete Neandertal mitochondrial genome sequence determined by high-throughput sequencing. Cell 134(3):416–426

Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Hansen NF (2010) A draft sequence of the Neandertal genome. Science 328(5979):710–722

Haak W, Brandt G, de Jong HN, Meyer C, Ganslmeier R, Heyd V et al (2008) Ancient DNA, strontium isotopes, and osteological analyses shed light on social and kinship organization of the later stone age. Proc Natl Acad Sci U S A 105(47):18226–18231

Hagelberg E, Clegg JB (1991) Isolation and characterization of DNA from archaeological bone. Proc R Soc Lond Ser B Biol Sci 244(1309):45–50

Hagelberg E, Sykes B, Hedges R (1989) Ancient bone DNA amplified. Nature 342(6249):485–485

Hänni C, Brousseau T, Laudet V, Stehelin D (1995a) Isopropanol precipitation removes PCR inhibitors from ancient bone extracts. Nucleic Acids Res 23(5):881

Hänni C, Begue A, Laudet V, Stéhelin D, Brousseau T, Amouyel P, Duday H (1995b) Molecular typing of Neolithic human bones. J Archaeol Sci 22(5):649–658

Hansen AJ, Mitchell DL, Wiuf C, Paniker L, Brand TB, Binladen J et al (2006) Crosslinks rather than strand breaks determine access to ancient DNA sequences from frozen sediments. Genetics 173(2):1175–1179

Helgason A, Pálsson S, Lalueza-Fox C, Ghosh S, Sigurðardóttir S, Baker A et al (2007) A statistical approach to identify ancient template DNA. J Mol Evol 65(1):92–102

Hendy J, Collins M, Teoh KY, Ashford DA, Thomas-Oates J, Donoghue HD et al (2016) The challenge of identifying tuberculosis proteins in archaeological tissues. J Archaeol Sci 66:146–153

Higham T (2011) European middle and upper Palaeolithic radiocarbon dates are often older than they look: problems with previous dates and some remedies. Antiquity 85(327):235–249

Higuchi R, Bowman B, Freiberger M, Ryder OA, Wilson AC (1984) DNA sequences from the quagga, an extinct member of the horse family. Nature 312(5991):282

Hofreiter M, Jaenicke V, Serre D, Haeseler AV, Pääbo S (2001) DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. Nucleic Acids Res 29(23):4793–4799

Holland MM, Fisher DL, Mitchell LG, Rodriquez WC, Canik JJ, Merril CR, Weedn VW (1993) Mitochondrial DNA sequence analysis of human skeletal remains: identification of remains from the Vietnam war. J Forensic Sci 38(3):542–553

Hong SB, Kim Y, Park JH (2017) High-efficiency automated DNA extraction method for degraded old skeletal samples. Forensic Sci Int Genet Suppl Ser 6:e365–e367

Horai S, Hayasaka K, Murayama K, Wate N, Koike H, Nakai N (1989) DNA amplification from ancient human skeletal remains and their sequence analysis. Proc Jpn Acad Ser B 65 (10):229–233

Höss M, Dilling A, Currant A, Pääbo S (1996) Molecular phylogeny of the extinct ground sloth Mylodon darwinii. Proc Natl Acad Sci U S A 93(1):181–185

Johnson V, Beckett S, Márquez-Grant N (2017) Differentiating human versus non-human bone by exploring the nutrient foramen: implications for forensic anthropology. Int J Leg Med 131 (6):1757–1763

Juras A, Chyleński M, Krenz-Niedbała M, Malmström H, Ehler E, Pospieszny Ł et al (2017) Investigating kinship of Neolithic post-LBK human remains from Krusza Zamkowa, Poland using ancient DNA. Forensic Sci Int Genet 26:30–39. https://doi.org/10.1016/j.fsigen.2016.10.008

Kalmár T, Bachrati CZ, Marcsik A, Raskó I (2000) A simple and efficient method for PCR amplifiable DNA extraction from ancient bones. Nucleic Acids Res 28(12):e67–e67

Karger BL, Guttman A (2009) DNA sequencing by CE. Electrophoresis 30(S1):S196–S202

Keller A, Kreis S, Leidinger P, Maixner F, Ludwig N, Backes C et al (2017) miRNAs in ancient tissue specimens of the Tyrolean iceman. Mol Biol Evol 34(4):793–801

Kerley ER (1965) The microscopic determination of age in human bone. Am J Phys Anthropol 23 (2):149–163

Kidd KK, Pakstis AJ, Speed WC, Grigorenko EL, Kajuna SL, Karoma NJ et al (2006) Developing a SNP panel for forensic identification of individuals. Forensic Sci Int 164(1):20–32

Kim K, Bazarragchaa M, Brenner CH, Choi BS, Kim KY (2015) Extensive evaluation of DNA polymerase performance for highly degraded human DNA samples. Forensic Sci Int 251:171–178

King CE, Debruyne R, Kuch M, Schwarz C, Poinar HN (2009) A quantitative approach to detect and overcome PCR inhibition in ancient DNA extracts. Biotechniques 47(5):941–949

Lalueza-Fox C, Römpler H, Caramelli D, Stäubert C, Catalano G, Hughes D et al (2007) A melanocortin 1 receptor allele suggests varying pigmentation among Neanderthals. Science 318(5855):1453–1455

Larson G, Albarella U, Dobney K, Rowley-Conwy P, Schibler J, Tresset A et al (2007) Ancient DNA, pig domestication, and the spread of the Neolithic into Europe. Proc Natl Acad Sci U S A 104(39):15276–15281

Lee HY, Park MJ, Kim NY, Sim JE, Yang WI, Shin KJ (2010) Simple and highly effective DNA extraction methods from old skeletal remains using silica columns. Forensic Sci Int Genet 4 (5):275–280

Lee EJ, Renneberg R, Harder M, Krause-Kyora B, Rinne C, Müller J et al (2014) Collective burials among agro-pastoral societies in later Neolithic Germany: perspectives from ancient DNA. J Archaeol Sci 51:174–180

Lee-Thorp JA (2008) On isotopes and old bones. Archaeometry 50(6):925–950

Leonard JA, Wayne RK, Cooper A (2000) Population genetics of ice age brown bears. Proc Natl Acad Sci U S A 97(4):1651–1654

Leonardi M, Librado P, Der Sarkissian C, Schubert M, Alfarhan AH, Alquraishi SA et al (2017) Evolutionary patterns and processes: lessons from ancient DNA. Syst Biol 66(1):e1–e29

Lindahl T (1993) Instability and decay of the primary structure of DNA. Nature 362(6422):709–715

Linderholm A (2016) Ancient DNA: the next generation–chapter and verse. Biol J Linn Soc 117 (1):150–160

Loreille O, Koshinsky H, Fofanov VY, Irwin JA (2011) Application of next generation sequencing technologies to the identification of highly degraded unknown soldiers' remains. Forensic Sci Int Genet Suppl Ser 3(1):e540–e541

Lorente JA, Entrala C, Alvarez JC, Arce B, Heinrichs B, Lorente M et al (2001) Identification of missing persons: the Spanish "phoenix" program. Croat Med J 42(3):267–270

Lorenz TC (2012) Polymerase chain reaction: basic protocol plus troubleshooting and optimization strategies. JoVE 63:e3998

Lovejoy CO, Meindl RS, Pryzbeck TR, Mensforth RP (1985) Chronological metamorphosis of the auricular surface of the ilium: a new method for the determination of adult skeletal age at death. Am J Phys Anthropol 68(1):15–28

Lundy JK (1998) Forensic anthropology: what bones can tell us. Lab Med 29(7):423–427

Lutz-Bonengel S, Sänger T, Pollak S, Szibor R (2004) Different methods to determine length heteroplasmy within the mitochondrial control region. Int J Leg Med 118(5):274–281

Mabuchi T, Susukida R, Kido A, Oya M (2007) Typing the 1.1 kb control region of human mitochondrial DNA in Japanese individuals. J Forensic Sci 52(2):355–363

Malmström H, Svensson EM, Gilbert MTP, Willerslev E, Götherström A, Holmlund G (2007) More on contamination: the use of asymmetric molecular behavior to identify authentic ancient human DNA. Mol Biol Evol 24(4):998–1004

Malmström H, Vretemark M, Tillmar A, Durling MB, Skoglund P, Gilbert MTP et al (2012) Finding the founder of Stockholm–a kinship study based on Y-chromosomal, autosomal and mitochondrial DNA. Ann Anat 194(1):138–145

Maricic T, Pääbo S (2009) Optimization of 454 sequencing library preparation from small amounts of DNA permits sequence determination of both DNA strands. Biotechniques 46(1):51–57

Marquez-Grant N (2015) An overview of age estimation in forensic anthropology: perspectives and practical considerations. Ann Hum Biol 42(4):308–322

Matheson CD, Marion TE, Hayter S, Esau N, Fratpietro R, Vernon KK (2009) Removal of metal ion inhibition encountered during DNA extraction and amplification of copper-preserved

archaeological bone using size exclusion chromatography. Am J Phys Anthropol 140 (2):384–391

McGlynn G (2007) Using 13C-, 15N-, and 18O stable isotope analysis of human bone tissue to identify transhumance, high altitude habitation and reconstruct paleodiet for the early medieval alpine population at Volders, Austria. Doctoral dissertation, Lmu

Meissner C, Ritz-Timme S (2010) Molecular pathology and age estimation. Forensic Sci Int 203 (1–3):34–43

Meyer M, Kircher M, Gansauge MT, Li H, Racimo F, Mallick S et al (2012a) A high-coverage genome sequence from an archaic Denisovan individual. Science 338(6104):222–226

Meyer C, Ganslmeier R, Dresely V, Alt KW (2012b) New approaches to the reconstruction of kinship and social structure based on bioarchaeological analysis of Neolithic multiple and collective graves. In: Kolár J, Trampota F (eds) Theoretical and methodological considerations in central European Neolithic archaeology, Oxford (BAR international series 2325). Archaeopress, Oxford, pp 11–23

Monroe C, Grier C, Kemp BM (2013) Evaluating the efficacy of various thermo-stable polymerases against co-extracted PCR inhibitors in ancient DNA samples. Forensic Sci Int 228 (1–3):142–153

Moorjani P, Sankararaman S, Fu Q, Przeworski M, Patterson N, Reich D (2016) A genetic method for dating ancient genomes provides a direct estimate of human generation interval in the last 45,000 years. Proc Natl Acad Sci U S A 113(20):5652–5657

Morrill BH, Rickords LF, Schafstall HJ (2008) Sequence length polymorphisms within primate amelogenin and amelogenin-like genes: usefulness in sex determination. Am J Primatol 70 (10):976–985

Mulligan CJ (2005) Isolation and analysis of DNA from archaeological, clinical, and natural history specimens. In: Methods in enzymology, vol 395. Academic Press, Cambridge, pp 87–103

Mullis KB, Faloona FA (1987) [21] Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. In: Ray W (ed) Recombinant DNA part F. Academic Press, Inc, pp 335–350. https://doi.org/10.1016/0076-6879(87)5503-26

Nilsson M, Grånemo J, Buś MM, Havsjö M, Allen M (2016) Comparison of DNA polymerases for improved forensic analysis of challenging samples. Forensic Sci Int Genet 24:55–59

Noonan JP, Coop G, Kudaravalli S, Smith D, Krause J, Alessi J, Rubin EM (2006) Sequencing and analysis of Neanderthal genomic DNA. Science 314(5802):1113–1118

Okazaki M, Yoshida Y, Yamaguchi S, Kaneno M, Elliott JC (2001) Affinity binding phenomena of DNA onto apatite crystals. Biomaterials 22(18):2459–2464

Orlando L, Willerslev E (2014) An epigenetic window into the past? Science 345(6196):511–512

Ortner DJ (2003) Identification of pathological conditions in human skeletal remains. Academic Press, San Diego

Pääbo S (1985) Molecular cloning of ancient Egyptian mummy DNA. Nature 314(6012):644–645

Pääbo S (1989) Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. Proc Natl Acad Sci U S A 86(6):1939–1943

Pääbo S (1990) Amplifying ancient DNA. In: Innis MA, Gelfand DH, Sninsky JJ (eds) PCR protocols: a guide to methods and applications. Academic Press, San Diego, pp 159–166

Pääbo S, Poinar H, Serre D, Jaenicke-Després V, Hebler J, Rohland N et al (2004) Genetic analyses from ancient DNA. Annu Rev Genet 38:645–679

Pajnič IZ, Debska M, Pogorelc BG, Mohorčič KV, Balažic J, Zupanc T et al (2016) Highly efficient automated extraction of DNA from old and contemporary skeletal remains. J Forensic Leg Med 37:78–86

Parson W, Berger C, Sänger T, Lutz-Bonengel S (2015) Molecular genetic analysis on the remains of the dark countess: revisiting the French royal family. Forensic Sci Int Genet 19:252–254

Parton J, Hassan NAM, Brown TA, Haswell SJ, Brown KA, Shaw KJ (2013) Sex identification of ancient DNA samples using a microfluidic device. J Archaeol Sci 40(1):705–711

Perry WL, Bass WM, Riggsby WS, Sirotkin K (1988) The autodegradation of deoxyribonucleic acid (DNA) in human rib bone and its relationship to the time interval since death. J Forensic Sci 33(1):144–153

Prado VF, Castro AKF, Oliveira CL, Souza KT, Pena SDJ (1997) Extraction of DNA from human skeletal remains: practical applications in forensic sciences. Gen Anal Biomol Eng 14(2):41–44

Procopio N, Chamberlain AT, Buckley M (2017) Intra-and Interskeletal proteome variations in fresh and buried bones. J Proteome Res 16(5):2016–2029

Pruvost M, Schwarz R, Correia VB, Champlot S, Braguier S, Morel N et al (2007) Freshly excavated fossil bones are best for amplification of ancient DNA. Proc Natl Acad Sci U S A 104(3):739–744

Putkonen MT, Palo JU, Cano JM, Hedman M, Sajantila A (2010) Factors affecting the STR amplification success in poorly preserved bone samples. Investig Genet 1(1):9

Quincey D, Carle G, Alunni V, Quatrehomme G (2013) Difficulties of sex determination from forensic bone degraded DNA: a comparison of three methods. Sci Justice 53(3):253–260

Rådström P, Knutsson R, Wolffs P, Lövenklev M, Löfström C (2004) Pre-PCR processing: strategies to generate PCR-compatible samples. Mol Biotechnol 26(2):133–146

Reich D, Green RE, Kircher M, Krause J, Patterson N, Durand EY et al (2010) Genetic history of an archaic hominin group from Denisova cave in Siberia. Nature 468(7327):1053

Robbins Schug G (2017) A hierarchy of values: the bioarchaeology of complexity, order, health, and hierarchy at Harappa. In: Klaus H, Harvey AR, Cohen MN (eds) Bones of complexity: osteological indicators of emergent Heterarchy and hierarchy. University Press of Florida, Gainesville, MD, pp 263–289

Rohland N, Hofreiter M (2007) Ancient DNA extraction from bones and teeth. Nat Protoc 2 (7):1756

Rossen L, Nørskov P, Holmstrøm K, Rasmussen OF (1992) Inhibition of PCR by components of food samples, microbial diagnostic assays and DNA-extraction solutions. Int J Food Microbiol 17(1):37–45

Rothe J, Melisch C, Powers N, Geppert M, Zander J, Purps J et al (2015) Genetic research at a fivefold children's burial from medieval Berlin. Forensic Sci Int Genet 15:90–97

Salas A, Lareu MV, Carracedo A (2001) Heteroplasmy in mtDNA and the weight of evidence in forensic mtDNA analysis: a case report. Int J Leg Med 114(3):186–190

Sanger F, Coulson AR (1975) A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J Mol Biol 94(3):441–448

Sawafuji R, Cappellini E, Nagaoka T, Fotakis AK, Jersie-Christensen RR, Olsen JV et al (2017) Proteomic profiling of archaeological human bone. R Soc Open Sci 4(6):161004

Sawyer S, Krause J, Guschanski K, Savolainen V, Pääbo S (2012) Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. PLoS One 7(3):e34131

Scheuer L (2002) A blind test of mandibular morphology for sexing mandibles in the first few years of life. Am J Phys Anthropol 119(2):189–191

Schneider E, El Hajj N, Haaf T (2014) Epigenetic information from ancient DNA provides new insights into human evolution. Brain Behav Evol 84(3):169–171

Schrader C, Schielke A, Ellerbroek L, Johne R (2012) PCR inhibitors–occurrence, properties and removal. J Appl Microbiol 113(5):1014–1026

Schutkowski H (1993) Sex determination of infant and juvenile skeletons: I. Morphognostic features. Am J Phys Anthropol 90(2):199–205

Schwarz C, Debruyne R, Kuch M, McNally E, Schwarcz H, Aubrey AD et al (2009) New insights from old bones: DNA preservation and degradation in permafrost preserved mammoth remains. Nucleic Acids Res 37(10):3215–3229

Shapiro Á, Hofreiter M (2014) A paleogenomic perspective on evolution and gene function: new insights from ancient DNA. Science 343(6169):1236573

Shaw KJ, Brown KA, Brown TA, Haswell SJ (2015) Sex identification of ancient DNA samples using a microfluidic device. In: Microchip capillary electrophoresis protocols. Humana Press, New York, pp 93–98

Shinde V, Lee H, Yadav Y, Waghmare P, Jadhav N, Hong JH et al (2018) A young couple's grave found in the Rakhigarhi cemetery of the Harappan civilization. Anat Cell Biol 51(3):200–204
Shved N, Haas C, Papageorgopoulou C, Akguel G, Paulsen K, Bouwman A et al (2014) Post mortem DNA degradation of human tissue experimentally mummified in salt. PLoS One 9(10): e110753
Skoglund P, Malmström H, Raghavan M, Storå J, Hall P, Willerslev E et al (2012) Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. Science 336(6080):466–469
Skoglund P, Storå J, Götherström A, Jakobsson M (2013) Accurate sex identification of ancient human remains using DNA shotgun sequencing. J Archaeol Sci 40(12):4477–4482
Skoglund P, Northoff BH, Shunkov MV, Derevianko AP, Pääbo S, Krause J, Jakobsson M (2014) Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. Proc Natl Acad Sci U S A 111(6):2229–2234
Smith O, Gilbert MTP (2018) Ancient RNA. In: Paleogenomics. Springer, Cham, pp 53–74
Smith RW, Monroe C, Bolnick DA (2015) Detection of cytosine methylation in ancient DNA from five native American populations using bisulfite sequencing. PLoS One 10(5):e0125344
Speller CF, Yang DY (2016) Identifying the sex of archaeological Turkey remains using ancient DNA techniques. J Archaeol Sci Rep 10:520–525
Stanojevich V (2012) The role of a forensic anthropologist in death investigation. J Forensic Res 3 (6):3
Stiller M, Green RE, Ronan M, Simons JF, Du L, He W et al (2006) Patterns of nucleotide misincorporations during enzymatic amplification and direct large-scale sequencing of ancient DNA. Proc Natl Acad Sci U S A 103(37):13578–13584
Svensson EM, Götherström A, Vretemark M (2008) A DNA test for sex identification in cattle confirms osteometric results. J Archaeol Sci 35(4):942–946
Svensson EM, Telldahl Y, Sjöling E, Sundkvist A, Hulth H, Sjøvold T, Götherström A (2012) Coat colour and sex identification in horses from iron age Sweden. Ann Anat 194(1):82–87
Taubenberger JK, Reid AH, Janczewski TA, Fanning TG (2001) Integrating historical, clinical and molecular genetic data in order to explain the origin and virulence of the 1918 Spanish influenza virus. Philos Trans R Soc Lond Ser B Biol Sci 356(1416):1829–1839
Tierney SN, Bird JM (2015) Molecular sex identification of juvenile skeletal remains from an Irish medieval population using ancient DNA analysis. J Archaeol Sci 62:27–38
Tsangaras K, Greenwood AD (2018) Paleovirology: viral sequences from historical and ancient DNA. In: Paleogenomics. Springer, Cham, pp 139–162
Ubelaker DH (2018) Recent advances in forensic anthropology. Forensic Sci Res 3(4):275–277
Venanzi FM, Rollo F (1990) Mummy RNA lasts longer. Nature 343(6253):25
Vidaki A, Daniel B, Court DS (2013) Forensic DNA methylation profiling—potential opportunities and challenges. Forensic Sci Int Genet 7(5):499–507
Wall JD, Kim SK (2007) Inconsistencies in Neanderthal genomic DNA sequences. PLoS Genet 3 (10):e175
Warinner C, Rodrigues JFM, Vyas R, Trachsel C, Shved N, Grossmann J et al (2014a) Pathogens and host immunity in the ancient human oral cavity. Nat Genet 46(4):336
Warinner C, Hendy J, Speller C, Cappellini E, Fischer R, Trachsel C et al (2014b) Direct evidence of milk consumption from ancient human dental calculus. Sci Rep 4:7104
Watson CL, Lockwood DN (2009) Single nucleotide polymorphism analysis of European archaeological M. leprae DNA. PLoS One 4(10):e7547
Watson E, Bauer K, Aman R, Weiss G, von Haeseler A, Pääbo S (1996) mtDNA sequence diversity in Africa. Am J Hum Genet 59(2):437
Weaver DS (1980) Sex differences in the ilia of a known sex and age sample of fetal and infant skeletons. Am J Phys Anthropol 52(2):191–195
Welch L, Gill P, Tucker VC, Schneider PM, Parson W, Mogensen HS, Morling N (2011) A comparison of mini-STRs versus standard STRs—results of a collaborative European (EDNAP) exercise. Forensic Sci Int Genet 5(3):257–258

Welker F (2018) Elucidation of cross-species proteomic effects in human and hominin bone proteome identification through a bioinformatics experiment. BMC Evol Biol 18(1):23
Welker F, Duijm E, van der Gaag KJ, van Geel B, de Knijff P, van Leeuwen J, Gravendeel B (2014) Analysis of coprolites from the extinct mountain goat Myotragus balearicus. Quat Res 81 (1):106–116
Welker F, Collins MJ, Thomas JA, Wadsley M, Brace S, Cappellini E et al (2015) Ancient proteins resolve the evolutionary history of Darwin's South American ungulates. Nature 522(7554):81
Welker F, Hajdinjak M, Talamo S, Jaouen K, Dannemann M, David F et al (2016) Palaeoproteomic evidence identifies archaic hominins associated with the Châtelperronian at the Grotte du Renne. Proc Natl Acad Sci U S A 113(40):11162–11167
Wilson IG (1997) Inhibition and facilitation of nucleic acid amplification. Appl Environ Microbiol 63(10):3741
Wilson LA, MacLeod N, Humphrey LT (2008) Morphometric criteria for sexing juvenile human skeletons using the ilium. J Forensic Sci 53(2):269–278
Yang DY, Eng B, Waye JS, Dudar JC, Saunders SR (1998) Improved DNA extraction from ancient bones using silica-based spin columns. Am J Phys Anthropol 105(4):539–543
Yeap SS, Beaumont M, Bennett A, Keating NA, White DA, Hosking DJ (1998) Genetic and environmental factors affecting bone mineral density in large families. Postgrad Med J 74 (872):349–354
Zhenilo SV, Sokolov AS, Prokhortchou EB (2016) Epigenetics of ancient DNA. Acta Naturae 8 (3):72–76

Usefulness of Mini-STRs in Analyzing Degraded DNA Samples and Their Forensic Relevance

10

Mukesh Thakar, Bhawana Joshi, and Pankaj Shrivastava

Abstract

The ground of individual identification has always been an integral part of forensic caseworks. DNA analysis from a range of biological material serves as a basis of human identity testing. The adaptation of recently developed techniques has reached a new level of recognition to deal with a variety of biological materials encountered in different forensic situations. The selection of markers has linked different nations with the help of globally maintained DNA databases. Often, incidental state of affairs results in poor quality or quantity of the samples which in return requires special techniques or modifications to obtain some desirable results. This chapter primarily focuses on the employment, development and validation of mini-STR markers which efficiently work with a massive amount of compromised samples. An assessment of appropriate caseworks analysis has been provided which improves the typing results to a great extent from the compromised samples.

Keywords

Mini-STRs · Degraded samples · DNA analysis · Forensic case works

M. Thakar · B. Joshi
Department of Forensic Science, Punjabi University, Patiala, Punjab, India

P. Shrivastava (✉)
DNA Fingerprinting Unit, State Forensic Science Laboratory, Department of Home (Police), Government of Madhya Pradesh, Sagar, Madhya Pradesh, India

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_10

10.1 Introduction

With the advent of forensic DNA profiling technique in 1985 by professor Sir Alec Jeffreys from Leicester University, UK, human identity testing was recognized. It was perceived that nucleotides were repeated more than a few times in a specific combination (termed as Variable Number of Tandem Repeats, VNTRs) and varied between individuals. It could be identified using Restriction Fragment Length Polymorphism (RFLP) as it focused on the use of restriction enzymes to cut the surrounding regions of VNTRs in DNA. These RFLP probes are highly variable among individuals but need more effort, time and proficiency to generate a DNA profile. Also, the call for processing large numbers of DNA samples could not be achieved using these probes since multilocus RFLP methods are not amenable to automation. These limitations have been best achieved with short tandem repeat (STR) DNA markers which comparatively have a high power of discrimination and quick analysis speed (Butler 2005). STR markers have noticeably become the authoritative tool for human identification. These are of extremely polymorphic type and vary in the length of different allele. This variability in the number of repeats makes STR markers valuable in identification process. The nomenclature is generally based on the length of the repeat unit. Trinucleotides contain three nucleotides in the repeat unit, tetranucleotides encompass four, pentanucleotides cover five, likewise hexanucleotides comprises of six repeat units in the core repeat. Among various types of STR systems being used for forensic DNA typing, tetranucleotide repeat have become more popular than di- or trinucleotides since they are less prone to slippage of polymerase during enzymatic amplification (Kimpton et al. 1994). Characteristics of STR markers like, low mutation rate, low stutter characteristic, compact size, makes it best suited for forensic samples analysis (Gill et al. 1996; Butler 2011). STRs narrow allele size range is amenable to multiplexing by means of PCR technique. This technique validated the analysis of several different loci at the same time (Kimpton et al. 1993; Lygo et al. 1994). The "first-generation multiplex" involved four loci TH01, FES/FPS, VWA and F13A1 which had a matching probability of approximately 1 in 10,000 (Kimpton et al. 1994). The matching probability has been increased remarkably with the development of second-generation multiplex. Different manufactures like AmpFℓSTR® SGM Plus™, Identifiler™ (Applied Biosystems, USA), PowerPlex® 16 System (Promega, USA) or Investigator IDplex® (Qiagen, Germany) have made these multiplex STR kits commercially available (Collins et al. 2004). The set of 13 CODIS STR loci expanded now to 20 STR loci markers has been accepted globally as a common way for communication to deal in most of the forensic situations (Hares 2015). These accessible kits can efficiently create STR amplicons in the size range of 100–450 bp (Krenke et al. 2002).

In numerous caseworks, the evidentiary value of the samples is usually lost since bodies are recovered after a longer duration of time. Forensic scientists meet up with several samples ranging from stain samples or ancient skeletal remains. DNA profiling for the aim of victim identification in mass disaster cases are always problematic due to a high degree of fragmentation of various body parts or due to

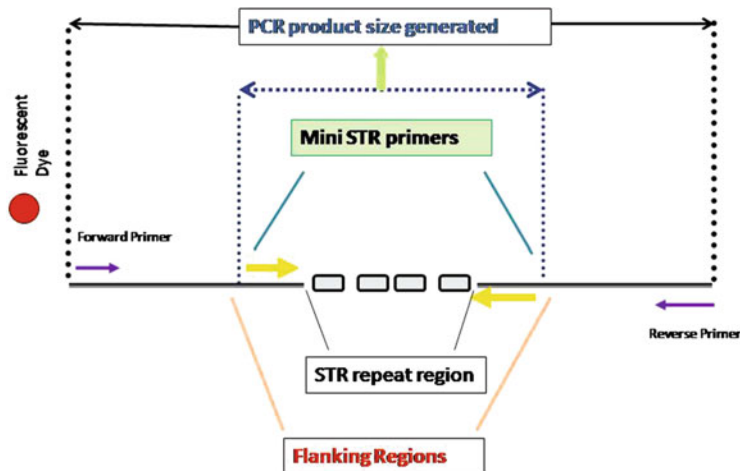


Fig. 10.1 PCR primers for mini-STR markers are redesigned to bind closer to each repeat region

a high level of intermix of remains. DNA in such samples often exists in an uneven manner. Such is the case when dealing with samples of fire cases (Heinrich et al. 2009) or bodies submerged in water for prolonged time (Phengon et al. 2008; Dhanardhono et al. 2013; Mameli et al. 2014). The reason of such breakdown includes exposure of DNA to light, humidity or fire for longer duration of time. Sometimes presence of any bacterial or microbial growth and other environmental contaminants also leads to fragmentation of DNA in the samples (Bar et al. 1988; Bender et al. 2004). Thus the quality and quantity of evidences recovered, determine the suitability of techniques to be applied for further analysis. Such low level DNA analysis is highly vulnerable to bigger risk of laboratory-based contamination. Stochastic effects such as allele dropouts, allele drop-ins, imbalance of heterozygote peak height and area and inconsistent peak size of stutter products also becomes advent in aforesaid situations (Clayton et al. 1998; Gill et al. 2000; Benschop et al. 2013). The process of cell death and DNA degradation is associated with enzymatic and nonenzymatic degradation of DNA. This further correlates with the failure of PCR amplification of degraded DNA samples (Takahashi et al. 1997; Alaeddini et al. 2010).

There has been considerable expansion in methodology used to genotype such compromised or fragmented samples. Mini-STRs have been established as a helpful tool for genotyping of samples in a degraded state (Dixon et al. 2006; Welch et al. 2011; Senge et al. 2011). The PCR amplicons length of these markers are usually compacted by moving the primers closer to the target region in order to create the smallest possible amplicons, as shown in Fig. 10.1. This provides the utmost likelihood of success when attempting to amplify highly fragmented DNA (Wiegand and Kleiber 2001; Tsukada et al. 2002; Coble and Butler 2005). A direct association between amplicon size and profile success rates has been established (Bender et al. 2004; Chung et al. 2004; Fondevila et al. 2008a, b). Additionally these mini-STRs

approved further segregation in complex paternity cases or kinship analysis missing person analysis (Goodwin et al. 2004; Butler et al. 2008).

10.2 Development of Mini-STRs

Mini-STR technology has been reported as an efficient approach to advance DNA profiling from harshly degraded casework human DNA samples that are refractory to amplification of DNA fragments larger than 200–300 bp. This approach helps in the investigation of various crime scene situations which come across low amount of degraded DNA samples. The sensitivity of typing degraded samples was improved considerably using redesigned PCR primers. It was first witnessed in investigating Waco Branch Davidian fire case's victims, where smaller STRs proved to better than larger STR amplification (Whitaker et al. 1995). The use of redesigned primer, capable of generating smaller PCR amplicons for CSF1PO showed an increased sensitivity for typing degraded specimens and better electrophoretic resolution (Yoshida et al. 1997). Amplicon length has been reduced up to 100 bp for the four systems (TH01, D10S2325, DYS19, and DYS391) to analyze degraded DNA (Wiegand and Kleiber 2001). Later, multiplex polymerase chain reaction (PCR) for the TH01, TPOX, CSF1PO and vWA loci has been carried out by a pair of newly designed primers that generated smaller fragments. These redesigned PCR primers proved to be most sensitive to evaluate even 80 pg of template DNA (Tsukada et al. 2002) (Table 10.1). A new primer design for the STR locus D2S1338 capable of generating small amplicons presented a successful STR analysis, resilient in contrast to mutation based amplification drop out (Grubwieser et al. 2003). These observations showing major benefits of typing forensic casework samples headed to redesigning of PCR primers for all the 13 CODIS STRs plus Penta D, Penta E and D2S1338 markers by work at a private laboratory (the Bode Technology Group) and NIST. These newly designed primers made each STR markers smaller in size as compared to their sizes available in STR kits (Butler et al. 2003) (Table 10.2).

10.2.1 Mini-STR Multiplex Based on Non-CODIS Loci

The conception of generating smaller size PCR amplicons by shifting the primers adjoining to STR repeat region gained potential for the analysis of non-CODIS STR loci. A range of norms for the selection of these non-CODIS markers included high heterozygosities, narrow allele size ranges and "clean" flanking regions adjacent to the STR repeat region for the primer binding. A total of 47 non-CODIS loci, capable of generating amplicon less than 125 bp in size were selected initially. Out of these, a set of two miniplexes having allele sizes less than 24 bp and heterozygosity value greater than 0.70 were selected finally. These six loci were also characterized among African Americans, Caucasians, and Hispanics from US population. These new loci successfully amplified all the degraded bone samples which, when used with the PowerPlex 16 kit gave partial profiles at 32 cycles of amplification. Along with its

Table 10.1 Timelines of designing short amplicons

Loci (primer redesigned)	Authors	Product length (bp)
CSF1PO	Yoshida et al. (1997)	150–182
D12S391	Ricci et al. (1999)	125–173
TH01, D10S2325,DYS319 and DYS19	Wiegand and Kleiber (2001)	80–100
FES, TH01 and TPOX	Hellmann et al. (2001)	<110
TH01, TPOX,CSF1PO and vWA	Tsukada et al. (2002)	74–143
D2S1338	Grubwieser et al. (2003)	114–166
13 CODIS STRs plus Penta D, Penta E and D2S1338	Butler et al. (2003)	51–281
D5S818,D8S1179, D16S539, vWA, D18S51, D13S317, TH01, CSF1PO,TPOX, FGA, D21S11,D7S820, Penta D, Penta E and D2S1338	Chung et al. 2004	125–173
D1S1677, D2S441, D4S2364, D10S1248,D14S1434 and D22S1045	Coble and Butler (2005)	67–123
D10S1248, D14S1434 and D22S1045	Dixon et al. (2006)	<170
D2S1338, D16S539, D18S51, TH01 and FGA	Grubwieser et al. (2006)	109–162
D1S1677, D2S441, D4S2364,D10S1248,D14S1434, D22S1045,TH01, CSF1PO,TPOX, FGA, D21S11,D7S820, Penta D, Penta E and D2S1338	Martin et al. (2006)	51–281
DYS522,DYS508, DYS632, DYS556, DYS570, DYS576, DYS504 and DYS540	Asamura et al. (2007a, b)	95–147
D1GATA113,D1S1627,D2S1776, D3S3053,D3S4529, D4S2408,D5S2500, D6S474,D6S1017,D8S1115,D9S1122, D9S2157,D10S1435,D11S443,D12ATA63, D17S974, D17S1301,D18S853,D20S482,D20S1082	Hill et al. (2008)	84–146
Amelogenin, SE33, D3S1358, D8S1179,D18S51, D21S11, FGA, TH01,vWA	Nastainczyk et al. (2009)	76–287
D10S1248, D1S1656, D12S391, D2S441, D22S1045	Nieuwerburgh et al. (2014)	78–253

utility in typing degraded samples, these novel mini-STRs were expected to aid in the identification of closely related individuals which often require additional genetic information (Coble and Butler 2005).

Further new primer sets were designed for another eight non-CODIS capable of generating smaller amplicons (Table 10.3). These new loci proved to highly effective in the analysis of degraded DNA samples and also in some complex kinship analysis for instance sibling verification or deficient paternity cases (Asamura et al. 2007a, b). Additional 20 non-CODIS mini-STR loci were also characterized and

Table 10.2 Comparison of amplicon sizes (bp) of mini-STRs loci in MiniFiler v/s other STR kits

	Manufacturer	D7S820 (bp)	D13S317 (bp)	D21S11 (bp)	D18S51 (bp)	D16S539 (bp)	FGA (bp)	CSF1PO (bp)
AmpFlSTR® Minifiler™	Applied Biosystems	136–176	88–132	153–211	113–193	81–121	125–281	89–129
PowerPlex® 16	Promega	215–247	176–208	203–259	290–366	264–304	322–444	321–357
PowerPlex® 21 system	Promega	269–313	302–350	203–259	134–214	84–132	265–411	318–362
PowerPlex® Fusion system	Promega							
GlobalFiler™ kit	Applied Biosystems	262–298	149–243	182–239	261–342	227–267	223–378	282–318
PowerPlex® Fusion 6C system	Promega	269–313	308–358	203–259	134–214	84–132	143–289	318–362

Table 10.3 List of non-CODIS mini-STR loci

Authors	Non-CODIS mini-STR loci	Size range (bp)
Coble and Butler (2005)	D1S1677	77–117
	D2S441	78–110
	D4S2364	67–83
	D10S1248	79–123
	D14S1434	70–102
	D22S1045	76–109
Asamura et al. (2007a, b)	D1S1171	98–138
	D2S1242	95–135
	D3S1545	75–107
	D4S2366	98–126
	D12S391	101–149
	D16S3253	79–111
	D20S161	104–132
	D21S437	104–140
Hill et al. (2008)	D1GATA113	81–105
	D1S1627	81–100
	D2S1776	127–161
	D3S3053	84–108
	D3S4529	111–139
	D4S2408	85–109
	D5S2500	85–125
	D6S474	107–135
	D6S1017	81–109
	D8S1115	63–96
	D9S1122	93–125
	D9S2157	71–107
	D10S1435	82–139
	D11S4463	88–116
	D12ATA63	76–106
	D17S974	95–123
	D17S1301	114–138
	D18S853	82–103
	D20S482	86–126
	D20S1082	73–100

optimized to create larger multiplexes to maximize the effectiveness of these markers. The heterozygosity values for all these markers were equivalent to those of 13 CODIS loci. Out of these 20, six loci contain a trinucleotide repeat unit while the remaining 14 consists of tetra nucleotide repeats. The stutter products for trinucleotide were comparatively lower than that of tetranucleotide repeats (Hill et al. 2008).

10.2.2 Mini-X STR System

In the recent past, analysis of X-STR loci is proved to be more informative over autosomal STRs, particularly in the cases of kinship analysis and parentage testing. These STR markers of X-chromosomes offer more successful analysis in some characteristic cases of kinship testing, like father–daughter duos or close blood relatives (Szibor et al. 2003; Gomes et al. 2012). Further, X-STR system also has the potential to solve siblingship status and to determine paternal grandmother–granddaughter relationships (Edelmann et al. 2004; Toni et al. 2006). To confirm the success of X-STR system for examining highly degraded DNA samples, new primer sets have been created to produce amplicons that are as small as possible. A mini-X STR multiplex system was successfully generated for all eight X-linked STR loci (DXS7423, DXS6789, DXS101, GATA165B12, DXS7133, DXS7424, DXS8378 and GATA31E08). These X-linked loci generated PCR amplicons in a range of length from 76 to 169 bp having more detection sensitivity. Also, the statistical measure of these eight loci pointed towards a high utility for analyzing degraded DNA and Japanese forensic practice (Asamura et al. 2006).

10.2.3 Mini-Y STR System

The inheritance pattern Y-chromosome establishes its role in ascertaining paternal lineages. With the accessibility of genotyping markers for the nonrecombining segment of the Y chromosome (NRY region), Y-DNA typing has turn out to be one of the basic techniques in Forensic DNA typing (Jain et al. 2016). Various Y-STR multiplex kits are available commercially, which could not detect loci with amplicon size more than 200 bp in degraded samples. In an attempt to check the efficacy of Y-STR markers for the analysis of degraded DNA samples, polymerase chain reaction (PCR) primer sets capable of generating short amplicons have been developed for DYS504, DYS508, DYS522, DYS540, DYS556, DYS570, DYS632 and DYS576.These loci do not associate with any commercially accessible Y-STR multiplex kit and are able to generate amplicons less than 150 bp. This mini-YSTR quadruplex system also generated genetic data for 224 Japanese male individuals. Their calculated value demonstrated haplotype diversity and discrimination capacity as 0.9949 and 0.7589, respectively. This system proved to be a quite effective tool for majority of forensic practises (Asamura et al. 2007a, b). The genotyping results using Y-miniplex sets produced better signal from enzymatically degraded DNA, along with a 50-year old one samples. The power of discrimination of Y-miniplexes was much higher than that of AmpFlSTR® Yfiler™ kit (Park et al. 2007).

10.3 Commercially Available Minifiler PCR Amplification Kit

The (Q8) German DNA database systems for D3S1358, FGA, TH01, SE33, D18S51,VWA, D8S1179 and D21S11 along with the gender typing marker Amelogenin has been customized with mini-STR markers. The Q8 is found to be best suited mainly for flaccid and degraded stains. The amplicon lengths of the STR markers have been significantly reduced as compared to commercially established DNA database kits. The most sensitive fluorescent dyes (FAM, JOE and TET) were used to improve the typing success for feeble stains. A higher success rate was attained with the Q8 kit in comparison with a well-established MPX2 commercial kit (Muller et al. 2007).

The AmpFlSTR® MiniFiler™ PCR Amplification Kit has been designed and made commercially available to obtain information from forensic samples where STR markers failed to give desired results. The MiniFiler™ kit effectively amplified eight of the largest sized loci (D2S1338, D7S820, D13S317, D16S539, D21S11, D18S51, CSF1PO and FGA) present in the AmpFlSTR® Identifiler® PCR Amplification Kit. An entire set of the STR loci amplified by the MiniFiler™ kit has been formerly characterized by further associations and can also be found in commercial STR multiplexes like the Identifiler and SGM Plus kits. The MiniFiler™ kit primers were constructed to achieve shorter amplified products (less than prevailing generated using other commercially accessible Applied Biosystems kits) with powerful signal strength and balanced peak heights from human samples. Initially, singleplex reactions were approved by employing redesigned primer sets to ensure locus-specific amplification. Then multiplex PCR studies were implemented by means of primer concentration adjustment and empirical attainment measure to achieve sensitive, balanced and precise signals for the eight STR loci and the sex determining marker in a single PCR reaction. The MiniFiler™ kit STR loci have a length in between 71and 250 bp. To embody the co-amplification of eight loci within this limited size range, a five-dye fragment analysis system has been developed at Applied Biosystems which are at present used in the AmpFlSTR SEfiler™, Identifiler and Yfiler kits. The five dyes (6-FAMTM, NEDTM, PET, VIC and LIZ) widen the fluorescent detection range to 660 nm, thereby implementing more loci to be multiplexed in a single PCR (Mulero et al. 2008).

10.4 Validation Studies of Mini-STRs

In conformity with the validation provisions in the SWGDAM guidelines any unique DNA typing kit must be evaluated with casework-related samples or samples from non-probative casework before it can be enforced in forensic casework. A set of nine studies were completed to validate the significance of MiniFiler™ PCR Amplification kit in forensic caseworks (Luce et al. 2009).

The use of low copy number (LCN) polymerase chain reaction (PCR) has been performed for the validation of the Q8 concept by analyzing 50 biological stains having limited and extremely degraded DNA. The sensitivity of LCN PCR is

advanced by increasing the number of cycles or via nested PCR primer. However LCN PCR is often influenced with artefacts like false alleles caused by slippage artefacts during the initial PCR cycles, allele dropouts as well as sporadic contamination. To minimize PCR errors, the number of cycles is generally accustomed to the confined amount of template DNA. The Q8 multiplex has shown full concordance with the other commercial kits. Also, Q8 multiplex has analysed more loci in comparison with the two commercial kits, MPX2 and Nonaplex (Muller et al. 2009).

10.5 Population Study on Mini-STRs

Three multiplex PCR systems (multiplex I: D10S1248, D14S1434 and D22S1045; multiplex II: D1S1677, D2S441 and D4S2364; and multiplex III: D3S3053, D6S474 and D20S482) has been used to estimate allele frequencies and forensic parameters of 9 mini-STR loci in 300 unrelated Koreans. The results recommended the use of mini-STR loci (D2S441, D6S474, D10S1248, D14S1434 and D22S1045) for forensic study of degraded DNA and are considered really informative as the CODIS STRs (Chung et al. 2007). Allele frequency data for 12 mini-STR autosomal loci (D1S1677, D2S441, D4S234, D4S2639, D6S2439, D6S1056, D9S1118, D10S1248, D14S1434, D17S1290, D20S480 and D22S1045) has been determined for 506 unrelated individuals from the Central-East Region of Argentina. The results showed no substantial deviations from Hardy–Weinberg equilibrium and suggested its suitability to provide better info for typing of degraded material (Vullo et al. 2010).

The allele frequency of six mini-STR loci, D1S1677, D4S2364, D2S441, D10S1248, D14S1434 and D22S1045, was investigated in three distinctive Singapore populations. The allele distribution of all six mini-STR loci verified that they are equally polymorphic in all three Singapore populations. No markers established any notable deviation from Hardy–Weinberg equilibrium. The heterozygosity marked for Chinese, Malay and Indian has found to be in the range of 0.600–0.816, 0.610–0.758 and 0.640–0.764, respectively. The estimated combined random match probability (RMP) of the six mini-STR found to be 4.6×10^{-6}, 3.5×10^{-6} and 4.2×10^{-6}, while the combined power of exclusion as 97.77%, 96.68% and 97.55% for Chinese, Malay and Indian, respectively (Yong et al. 2007). Allele frequencies and forensic efficiency data of eight mini-STR loci were observed in 128 unrelated individuals from Chinese Uyghur ethnic minority group. The PCR amplicons produced by these loci were less than 240 bp in size. Allele frequencies of 94 observed alleles were between 0.0039 and 0.3438. The overall results demonstrated high degree of polymorphism, higher combined power of discrimination and combined power of exclusion respectively (Wang et al. 2014). Mini-X STR multiplexes have been used to generate the population data of 366 Punjabi and 346 Sindhi individuals of Pakistan. Short amplicons were designed for a total of 11 X-STRs used in a miniplex model. The statistical parameters considered for these mini-X STR markers proved its suitability for individual identification and paternity testing (Israr et al. 2014). An analysis of eight mini-STR loci (D16S539, D7S820,

D13S317, FGA, CSF1PO, D21S11, D18S51, and D2S1338) was completed in 140 irrelative Sikh individuals (Arora, Jat and Ramgariha). All loci of mini-STRs were in accordance with Hardy–Weinberg equilibrium except at locus FGA in Ramgariha Sikh and locus D16S539 in Arora Sikh. Only 1% variation has been observed among the three studied populations, when using analysis of molecular variance (AMOVA). MDS plot generated on the basis of pair wise genetic distances that established the close genetic relationship between Jat and Ramgariha Sikh population (Dogra et al. 2015).

10.6 Concordance Studies Between Mini-STRs with STRs

The newly designed primers of mini-STRs have the potential to supplement the currently used STR systems. To access the performance of these mini-STR assays, concordant studies with globally accepted STR markers have been performed. These mini-STR primers for the 13 CODIS STRs plus the Penta D, Penta E, and D2S1338 markers afford fully concordant results with the commercial STR kits and even enhanced signal from degraded DNA exhibits (Butler et al. 2003). A full concordance has also been observed between Identifiler® and Minifiler™ kits for the analysis of about 1308 degraded samples. Full concordance in 99.7% of allele calls was observed between both of the kits (Hill et al. 2008).

A comparative study of the common markers of Identifiler® and MiniFiler™ kit for typing casework samples showed no inconsistencies in the profiles generated from both the kits. This further suggested the use of mini-STR kit to incorporate low-level DNA profiles in STR databases (Lopes et al. 2009).

10.7 Comparative Analysis of Mini-STRs with STRs, SNP and mt DNA

Fondevila et al. (2008a, b) compared the performance of standard STRs with mini-STRs and autosomal SNPs for the examination of naturally degraded samples. All submitted skeletal samples were recovered in a high rainfall and organic rich, acidic soil environment. The choice of an improved extraction procedure was considered as a significant factor for DNA recovery and typing success of degraded samples. An assessment of internal PCR control (IPC) indicated the inhabitance of PCR inhibitors. The authors suggested considering suitable evaluation of inhibition would be sufficient in choosing typing methods for degraded samples. This approach allowed successful amplification of mini-STRs and SNP multiplexes at all loci, which otherwise failed to amplify when tried with standard STRs loci. Marjanović et al. (2009) performed DNA testing for the identification of World War II skeletal remains recovered from exposed unknown graves in Slovenia. Standard autosomal STRs along with Y-STRs and mini-STRs analysis collectively used to identify missing persons after more than a half century. Out of 48 exhumed skeletal remains, 41 DNA profiles were generated that led to a positive identification of 6 remains. The

authors suggested that improvement should be made in the DNA extraction procedures to get positive outcomes for the samples that could not be profiled. Two commercially accessible DNA repair enzyme cocktails, Pre CRTM and Restorase1 were assessed against standard STR typing, mini-STRs, bi-allelic SNPs and tri-allelic SNPs. After incubation with DNA repair enzymes from PreCR™ or Restorase®, MiniFiler™ demonstrated more reproducible results and a superior percentage of detected alleles than standard STR analysis. The SNP genotyping results from GenPlex™ and the tri-allelic SNPs showed an even higher percentage of detected alleles than MiniFiler™, which recommended it to be more satisfactory for human identification cases (Westen and Sijen 2009).

A 50 nuclease assay and a mini sequencing (SNaPshot) assay has been compared with the two commercially available mini-STR kits and a set of five SNP markers. The analysis encompassed different types of stain material, along with intact and degraded DNA. The mini-STRs are found to be superior for typing old stain material with small amounts of degraded DNA than standard STR typing methods. SNP typing based on the mini sequencing (SNaPshot) assay attained a better success rate in typing of aged blood and saliva stains than standard STRs and SNP typing (Senge et al. 2011).

10.8 Recovery of Degraded DNA Samples Using Mini-STRs

Constantinescu et al. (2012) established the success of mini-STR technology in resolving challenging casework samples which were previously analyzed using commercial 16-STR multiplex kits. Incomplete DNA profiles were obtained from four problematic casework samples when amplification performed using Amp1STR®Identifiler™ and PowerPlex® 16. Array of samples submitted was bone sample, incomplete mixture from vaginal swab, and false positive mixture from a acquired strain. Presence of PCR inhibitors, extreme fragmentation or low quantity of DNA could be considered a reason for the loss of larger sized STR loci. In all the complex cases AmpFℓSTR® MiniFiler™ kit (Applied Biosystem) improved signals and generated complete DNA profile. Competency of reduced size STRs primers has been presented by creating a mini quadruplex (Penta D, D18S51, D2S1338 and Amelogenin) and a mini-pentaplex (D21S11, D7S820, CSF1PO, TH01 and FGA) that are considered as heavier loci of the commercially available kits. It is indicated that the major benefit associated with peak heights has shown up by the persistent reduced size STRs (Pizzamiglio et al. 2006). The number of DNA copies actually accessible as a template for PCR could be much lower in the cases of fire and mass fatalities. In such situations, DNA quality and quantity cannot be determined separately. Also, such low amounts of DNA (15–30 pg/μl) could not describe failures like partial amplification. Traditional STR kits resulted in stochastic variations, which make them ineffective for analyzing degraded DNA. Mini-STRs profiling was subjected to analyze such low copy number (LCN) DNA. Employment of customized mini-STR protocols within a LCN-based approach allowed a decent genotyping even from minute DNA copies (Romano et al. 2006). Martin et al.

(2006) generated genetic profiles of degraded biological samples by using MiniFiler™ kit, which provided more selective results. PCR multiplexes of mini-STR loci were used to examine formalin fixed soft tissue fragments, bone remains from mass graves of the Spanish civil war (1936–1939) and paraffin embedded biopsies. For entire case samples, mini-STR multiplexes contributed additional information than the previous partial profiles obtained by Identifiler™ kit. The changes made in the MiniFiler™ kit improved sensitivity for detection of tiny amounts of DNA. Andrade et al. (2008) effectively recovered full DNA profile with the compromised 16 samples (blood, saliva, bone tooth and hair) using MiniFiler™ PCR amplification kit. These samples when subjected to Identifiler™ kit showed partial profiles or no profiles for most of them. Decorte et al. (2008) developed a mini-STR multiplex to examine restricted DNA samples. New PCR primers were schemed by reducing the amplicon sizes for the larger STR loci (TH01, D18S51, D21S11 and FGA) of SGM Plus. These larger-sized STR loci of SGM plus indicated locus drop out in the identification of skeletal remains. When same degraded samples were tried using short amplicon size primers, it showed positive amplification results for these loci as well.

10.9 Mini-STRs Analysis for Identification of Human Skeletal Remains

Fondevila et al. (2008a, b) identified an array of charred skeletal remains recovered afterwards a major forest fire in Galicia, NW Spain. The identity of the deceased was founded by comparing its DNA profile with that of claimed daughter. A decomposition period of about 10 years and subsequent exposure to extreme temperature directed a high level degradation of genetic material. To analyze such degraded material, a new set of markers including both reduced-length amplicon STR sets and autosomal SNPs were used. After DNA extraction, typing results from two standard STR kits (Identifiler™ and Powerplex16™) were compared with mini-STRs multiplexes and SNP assays. The results recommended mini-STRS as the better approach while typing of degraded samples, at the same time SNPs could be a useful extension to traditional STRs. Opel et al. (2006) validated the performance of miniplex primers to examine human skeletal remains, exposed to different environmental conditions. Bone samples (tibia, femur) were obtained from Forensic Anthropology Centre (FAC), Tennessee and Franklin County Coroner's Office (FCCO), Columbus. To check DNA degradation and PCR inhibition, samples were subjected to three conditions (burial, surface treatment and cold storage) before extraction. The proficiency of amplification using reduced size primer sets was compared with that of PowerPlex® 16 system (Promega). A loss of intensity (50% dropout) for larger alleles due to excess fragmentation was observed with PowerPlex® 16 system. It was concluded that allele drop out due to degradation could be size related not locus related. The sets of miniplexes used in the study showed a greater likelihood (˜64%) of generating a full profile from the compromised bone samples. Parsons et al. (2007) implemented three short amplicon multiplexes (5-plex, 6-plex and 7-plex)

for the testing of degraded skeletal remains, recovered from mass graves. These graves were resulted from the conflicts during 1992 to 1999 in the former Yugoslavia. The International Commission on Missing Persons (ICMP) was set up to identify around 40,000 missing individuals. The skeletal remains were recovered after a postmortem interval of 7–15 years. Though the primer sets used to get short amplicons were selected in new combinations to primarily target the largest loci of PP16. Success rate for the mini-amplicon multiplexes was 100% particularly for the bone samples recovered from the grave enclosed by lime. There was a limited percentage recovery of alleles when same processed with PowerPlex 16 kit by Promega.

10.10 Mini-STRs for Touch DNA Samples

The efficacy of mini-STRs has been successfully fulfilled the examination of touch DNA samples originated on different substrates (cars, tools, clothing, weapons, gloves, etc.). The study has been completed by considering the success rate before and after employment of the five extra loci that were added to the ESS for amplification of low amounts of degraded DNA. The statistical and descriptive results show an increased success rate after implementation of extra loci and after the implementation of mini-STR analysis (Nieuwerburgh et al. 2014).

The recognition of a body found in the Mediterranean Sea from has been achieved by DNA profiling of mandibular body. That body was apparently exposed to seawater for several months, which yielded 200 pg of low copy DNA. Use of Minifiler™ kit along with other traditional STR kits offered a full profile (Mameli et al. 2014).

10.11 Conclusion

The use of these mini-STR markers has greatly increased the power of discrimination in cases where the existing STR markers multiplex systems often give inconclusive results. Mini-STRs have been proven as an additional tool for identification of skeletal remains, the DNA sample of whose could not be amplified when performed otherwise. The location of such biological forensic evidences offers a great challenge to further investigation which in return requires a good collection and preservation skills. It offers an alternative approach for processing a number of compromised biological specimens in forensic situations. A comparison study has been completed between standard STR profiling with DNA repair enzyme incubation, and genotyping with mini-STR and single nucleotide polymorphisms (SNPs) in progressively degraded, UV-irradiated DNA samples. The study demonstrated that most of the standard STR markers fail to amplify, while mini-STRs and especially (tri-allelic) SNPs showed even higher percentage of detected alleles (Westen and Sijen 2009). The choice of mini-STR markers also supplements the core STR loci

available in DNA databases of various countries (Schneider and Martin 2001). Hence database compatibility could be maintained to a great extent.

References

Alaeddini R, Walsh SJ, Abbas A (2010) Forensic implications of genetic analyses from degraded DNA—a review. Forensic Sci Int Genet 4:148–157

Andrade L, Bento AM, Serra A, Carvalho M, Gamero JJ, Oliveira C, Batista L, Lopes V, Balsa F, Corte-Real F, Anjos MJ (2008) AmpFISTR® Minifiler™ PCR amplification kit: the new miniSTR multiplex kit. FSI Genet Supp Ser 1:89–91

Asamura H, Sakai H, Kobayashi K, Ota M, Fukushima H (2006) MiniX-STR multiplex system population study in Japan and application to degraded DNA analysis. Int J Legal Med 120:174–181

Asamura H, Fujimori S, Ota M, Fukushima H (2007a) MiniSTR multiplex systems based on non-CODIS loci for analysis of degraded DNA samples. Forensic Sci Int 173:7–15

Asamura H, Sakai H, Ota M, Fukushima H (2007b) MiniY-STR quadruplex systems with short amplicon lengths for analysis of degraded DNA samples. Forensic Sci Int Genet 1:56–61

Bar W, Kratzer A, Machler M, Schmid W (1988) Postmortem stability of DNA. Forensic Sci Int 39 (1):59–70

Bender K, Farfan MJ, Schneider PM (2004) Preparation of degraded human DNA under controlled conditions. Forensic Sci Int 139:135–140

Benschop C, Haned H, Sijen T (2013) Consensus and pool profiles to assist in the analysis and interpretation of complex low template DNA mixtures. Int J Legal Med 127(1):11–23

Butler JM (2005) Forensic DNA typing: biology, technology, and genetics of STR markers, 2nd edn. Elsevier Academic Press, Burlington. ISBN: 0-12-147952-8

Butler JM (2011) Advanced topics in forensic DNA typing: methodology, 1st edn. Academic Press, San Diego. eBook ISBN: 9780123878236. Hardcover ISBN: 9780123745132

Butler JM, Shen Y, McCord BR (2003) The development of reduced size STR amplicons as tools for analysis of degraded DNA. J Forensic Sci 48(5):1054–1064

Butler JM, Hill CR, Decker AE, Kline MC, Vallone PM (2008) Characterization of new autosomal STR loci. FSI Genet Suppl Ser 1:95–96

Chung DT, Drabek J, Opel KL, Butler JM, McCord BR (2004) A study on the effects of degradation and template concentration on the amplification efficiency of the STR Miniplex primer sets. J Forensic Sci 49(4):733–740

Chung U, Shin KJ, Park MJ, Kim NY, Yang WI, Cho SH, Lee HY (2007) Population data of nine miniSTR loci in Koreans. Forensic Sci Int 168:e51–e53

Clayton TM, Whitaker JP, Sparkes R, Gill P (1998) Analysis and interpretation of mixed forensic stains using DNA STR profiling. Forensic Sci Int 91:55–70

Coble MD, Butler JM (2005) Characterization of new MiniSTR loci to aid analysis of degraded DNA. J Forensic Sci 50(1):43–53

Collins PJ, Hennessy LK, Leibelt CS, Roby RK, Reeder DJ, Foxall PA (2004) Developmental validation of a single-tube amplification of the 13 CODIS STR loci, D2S1338, D19S433, and amelogenin: the AmpFℓSTR® Identifiler PCR amplification kit. J Forensic Sci 49 (6):1265–1277

Constantinescu CM, Barbarii LE, Iancu CB, Constantinescu A, Neagu E, Iancu D, Girbea G (2012) Challenging DNA samples solved with MiniSTR analysis. Rom J Leg Med 20:51–56

Decorte R, Liu CF, Vanderheyden N, Cassiman JJ (2008) Development of a novel miniSTR multiplex assay for typing degraded DNA samples. FSI Genet Suppl Ser 1:112–114

Dhanardhono T, Wulandri N, Bhima SKL, Ahmad HJ, Widodo PT (2013) DNA profiling of disaster victim identification in Trenggalek shipwreck case. FSI Genet Supp Ser 4:e5–e6

Dixon LA, Dobbins AE, Pulker HK, Butler JM, Vallone PM, Coble MD, Parson W, Berger B, Grubwieser P, Mogensen HS, Morling N, Nielsen K, Sanchez JJ, Petkovski E, Carracedo A, Sanchez-Diz P, Ramos-Luis E, Brion M, Irwin JA, Just RS, Loreille O, Parsons TJ, Syndercombe-Court D, Schmitter H, Stradmann-Bellinghausen B, Bender K, Gill P (2006) Analysis of artificially degraded DNA using STRs and SNPs—results of a collaborative European (EDNAP) exercise. Forensic Sci Int 164:33–44

Dogra D, Shrivastava P, Chaudhary R, Gupta U, Jain T (2015) Population genetics for autosomal STR loci in Sikh population of central India. Hereditary Genet 4(1):1–8

Edelmann J, Lessig R, Klintschar M, Szibor R (2004) Advantages of X-chromosomal microsatellites in deficiency paternity testing: presentation of cases. Int Congr Ser 1261:257–259

Fondevila M, Phillips C, Naveran N, Cerezo M, Rodriguez A, Calvo R, Fernandez LM, Carracedo A, Lareu MV (2008a) Challenging DNA: assessment of a range of genotyping approaches for highly degraded forensic samples. FSI Genet Suppl Ser 1:26–28

Fondevila M, Phillips C, Naveran N, Fernandez L, Cerezo M, Salas A, Carracedo A, Lareu MV (2008b) Case report: identification of skeletal remains using short-amplicon marker analysis of severely degraded DNA extracted from a decomposed and charred femur. Forensic Sci Int Genet 2:212–218

Gill P, Urquhart A, Millican E, Oldroyd N, Watson S, Sparkes R, Kimpton CP (1996) A new method of STR interpretation using inferential logic-development of a criminal intelligence database. Int J Legal Med 109:14–22

Gill P, Whitaker J, Flaxman C, Brown N, Buckleton J (2000) An investigation of the rigor of interpretation rules for STRs derived from less than 100pg of DNA. Forensic Sci Int 112 (1):17–40

Gomes C, Magalhaes M, Alves C, Amorim A, Pinto N, Gusmao L (2012) Comparative evaluation of alternative batteries of genetic markers to complement autosomal STRs in kinship investigations: autosomal indels vs. X-chromosome STRs. Int J Legal Med 126(6):917–921

Goodwin W, Ballard D, Simpson K, Thacker C, Syndercombe Court D, Gow J (2004) Case study: paternity testing—when 21 loci are not enough. In: Doutremepuich C, Morling N (eds) Progress in forensic genetics, vol 10. Int Cong Ser. 1261, Amsterdam, pp 460–462

Grubwieser P, Mühlmann R, Parson W (2003) New sensitive amplification primers for the STR locus D2S1338 for degraded casework DNA. Int J Legal Med 117:185–188

Grubwieser P, Mühlmann R, Berger B, Niederstätter H, Pavlic M, Parson W (2006) A new "miniSTR-multiplex" displaying reduced amplicon lengths for the analysis of degraded DNA. Int J Legal Med 120:115–120

Hares DR (2015) Selection and implementation of expanded CODIS core loci in the United States. Forensic Sci Int Genet 17:33–34

Heinrich A, Schwark T, Simeoni E, Wurmb-Schwark NV (2009) Genetic identifications of fire deaths. FSI Genet Suppl Ser 2:253–254

Hellmann A, Rohleder U, Schmitter H, Wittig M (2001) STR typing of human telogen hairs-a new approach. Int J Legal Med 114:269–273

Hill CR, Kline MC, Coble MD, Butler JM (2008) Characterization of 26 MiniSTR loci for improved analysis of degraded DNA samples. J Forensic Sci 53(1):73–80

Israr M, Shahid AA, Rahman Z, Zar MS, Shahzad MS, Husnain T, Pfeifer C, Wiegand P (2014) Development and characterization of a new 12-plex ChrX miniSTR system. Int J Legal Med 128:595–598

Jain T, Shrivastava P, Bansal DD, Dash HR, Trivedi VB (2016) PowerPlex Y23 system: a fast, sensitive and reliable Y-STR multiplex system for forensic and population genetic purpose. J Mol Biomark Diagn 7:287

Kimpton CP, Gill P, Walton A, Urquhart A, Millican ES, Adams M (1993) Automated DNA profiling employing multiplex amplification of short tandem repeat loci. PCR Methods Appl 3:13–22

Kimpton CP, Fisher D, Watson S, Adams M, Urquhart A, Lygo J, Gill P (1994) Evaluation of an automated DNA profiling system employing multiplex amplification of four tetrameric STR loci. Int J Legal Med 106:302–311

Krenke BE, Tereba A, Anderson SJ, Buel E, Culhane S, Finis CJ, Tomsey CS, Zachetti JM, Masibay A, Rabbach DR, Amjott EA, Sprecher CJ (2002) Validation of a 16-locus fluorescent multiplex system. J Forensic Sci 47(4):773–785

Lopes V, Andrade L, Carvalho M, Serra A, Balsa F, Bento AM, Batista L, Oliveira C, Corte-Real F, Anjos MJ (2009) Mini-STRs: a powerful tool to identify genetic profiles in samples with small amounts of DNA. FSI Genet Suppl Ser 2:121–122

Luce C, Montpetit S, Gangitano D, O'Donnell P (2009) Validation of the AmpFℓSTR® MiniFilerTM PCR amplification kit for use in forensic casework. J Forensic Sci 54(5):1046–1054

Lygo JE, Johnson PE, Holdaway DJ, Woodroffe S, Whitaker JP, Clayton TM, Kimpton CP, Gill P (1994) The validation of short tandem repeat (STR) loci for use in forensic casework. Int J Legal Med 107(2):77–89. (abstract)

Mameli A, Piras G, Delogu G (2014) The successful recovery of low copy number and degraded DNA from bones exposed to seawater suitable for generating a DNA STR profile. J Forensic Sci 59(2):470–473

Marjanović D, Durmić-Pašić A, Kovačević L, Avdić J, Džehverović M, Haverić S, Ramić J, Kalamujić B, Bilela LL, Škaro V, Projić P, Bajrović K, Drobnič K, Davoren J, Primorac D (2009) Identification of skeletal remains of communist armed forces victims during and after world war II: combined Y-chromosome short tandem repeat (STR) and MiniSTR approach. Croat Med J 50:296–304

Martin P, Garcia O, Albarran C, Garcia P, Alonso A (2006) Application of mini-STR loci to severely degraded casework samples. Int Congr Ser 1288:522–525

Mulero JJ, Chang CW, Lagac RE, Wang DY, Bas JL, McMahon TP, Hennessy LK (2008) Development and validation of the AmpFℓSTR® MiniFilerTM PCR amplification kit: a MiniSTR multiplex for the analysis of degraded and/or PCR inhibited DNA. J Forensic Sci 53(4):838–852

Muller K, Klein R, Miltner E, Wiegand P (2007) Improved STR typing of telogen hair root and hair shaft DNA. Electrophoresis 28:2835–2842

Muller K, Klein R, Miltner E, Wiegand P (2009) Q8—a short amplicon multiplex including the German DNA database systems. Forensic Sci Int Genet 1(2):205–207

Nastainczyk M, Schulz S, Kleiber M, Immel UD (2009) STR analysis of degraded DNA using a miniplex. Forensic Sci Int Genet Supp Series 2:53–54

Nieuwerburgh FV, Hoofstat DV, Neste CV, Deforce D (2014) Retrospective study of the impact of miniSTRs on forensic DNA profiling of touch DNA samples. Sci Justice 54(5):369–372

Opel KL, Chung DT, Drabek J, Tatarek NE, Jantz LM, McCord BR (2006) The application of Miniplex primer sets in the analysis of degraded DNA from human skeletal remains. J Forensic Sci 51(2):351–356

Park MJ, Lee HY, Chung U, Kang SC, Jin Shin K (2007) Y-STR analysis of degraded DNA using reduced sized amplicons. Int J Legal Med 121:152–157

Parsons TJ, Huel R, Davoren J, Katzmarzyk C, Milos A, Selmanovic A, Selmanovic L, Coble MD, Rizvic A (2007) A application of novel 'mini-amplicon' STR multiplexes to high volume casework on degraded skeletal remains. Forensic Sci Int Genet 1:175–179

Phengon P, Wongwigggarn S, Panvisavas N (2008) Analysis of DNA from degraded tissue. FSI Genet Suppl Ser 1:439–441

Pizzamiglio M, Marino A, Coli A, Floris T, Garofano L (2006) The use of mini-STRs on degraded DNA samples. Int Congr Ser 1288:498–500

Ricci U, Giovannucci Uzielli ML, Klintschar M (1999) Modified primers for D12S391 and a modified silver staining technique. Int J Legal Med 112(5):342–344

Romano C, Luise ED, Martino DD, Spitaleri S, Saravo L (2006) A novel approach for genotyping of LCN-DNA recovered from highly degraded samples. Int Congr Ser 1288:577–579

Schneider PM, Martin PD (2001) Criminal DNA databases: the European situation. Forensic Sci Int 119:232–238

Senge T, Madea B, Junge A, Rothschild MA, Schneider PM (2011) STRs, mini STRs and SNPs—a comparative study for typing degraded DNA. Legal Med 13:68–74

Szibor R, Krawczak M, Hering S, Edelmann J, Kuhlisch E, Krause D (2003) Use of X-linked markers for forensic purposes. Int J Legal Med 117:67–74

Takahashi M, Kato Y, Mukoyama H, Kanaya H, Kamiyama S (1997) Evaluation of five polymorphic microsatellite markers for typing DNA from decomposed human tissues—correlation between the size of the alleles and that of the template DNA. Forensic Sci Int 90:1–9

Toni C, Presciuttini S, Spinetti I, Rocchi A, Domenici R (2006) Usefulness of X-chromosome markers in resolving relationships: report of a court case involving presumed half sisters. Int Congr Ser 1288:301–303

Tsukada K, Takayanagi K, Asamura H, Ota M, Fukushima H (2002) Multiplex short tandem repeat typing in degraded samples using newly designed primers for the TH01, TPOX, CSF1PO, and vWA loci. Legal Med 4:239–245

Vullo C, Borosky A, Romanini C, Catelli L, Yamamoto T (2010) Frequency data for 12 mini STR loci in Argentina. Forensic Sci Int Genet 4:e79–e81

Wang H, Ruan J, Shen C, Zhang Y, Liu W, Meng H, Yang X, Yan J, Liao S, Fan S, Jia X, Wei X (2014) Polymorphic distribution and forensic effectiveness study of eight miniSTR in Chinese Uyghur ethnic group. Mol Biol Rep 41:2371–2375

Welch L, Gill P, Schneider PM, Parson W, Mogensen HS, Morling N (2011) A comparison of mini-STRs versus standard STRs-results of a collaborative European (EDNAP) exercise. Forensic Sci Int Genet 5:257–258

Westen AA, Sijen T (2009) Degraded DNA sample analysis using DNA repair enzymes, mini STRs and (tri-allelic) SNPs. FSI Genet Suppl Ser 2:505–507

Whitaker JP, Clayton TM, Urquhart AJ, Millican ES, Downes TJ, Kimpton CP, Gill P (1995) Short tandem repeat typing of bodies from a mass disaster: high success rate and characteristic amplification patterns in highly degraded samples. Biotechniques 18(4):670–677

Wiegand P, Kleiber M (2001) Less is more—length reduction of STR amplicons using redesigned primers. Int J Legal Med 114:285–287

Yong RYY, Gan LSH, Coble MD, Yap EPH (2007) Allele frequencies of six miniSTR loci of three ethnic populations in Singapore. Forensic Sci Int 166:240–243

Yoshida K, Sekiguchi K, Kasai K, Sato H, Seta S, Sensabaugh GF (1997) Evaluation of new primers for CSF1PO. Int J Legal Med 110:36–38

Capillary Electrophoresis Issues in Forensic DNA Typing

11

Shuchi Kaushik and Vivek Sahajpal

Abstract

This review on capillary electrophoresis discussed about the techniques and separation mechanisms employed in Deoxyribonucleic Acid (DNA) analyses while emphasizing more on the unprecedented advantages of capillary electrophoresis (CE) over the flat gel-based technology. With the emergence of novel approaches, capillary electrophoresis techniques evolve from development stages to validated and applied methods and thus gained its popularity among the forensic research groups as a powerful method for high-speed and high-throughput DNA sequencing and method of choice for forensic analysis of limited sample volumes of short tandem repeats (STRs) of DNA fragments. The types of gel matrices for sieving DNA fragments are also compared on the basis of their separation-based resolution performance. The review also summarizes types of CE platforms available for DNA sequencing, sizing, and aptamer studies. Expanding and emerging techniques in the field are also highlighted suggesting the pivotal role of optimization of CE methods in the development of DNA analysis technologies.

Keywords

DNA analysis · Forensic analysis · Gel-based technology · High-throughput sequencing · STRs

S. Kaushik (✉)
State Forensic Science Laboratory, Sagar, Madhya Pradesh, India

V. Sahajpal
State Forensic Science Laboratory, Directorate of Forensic Services, Shimla, Himachal Pradesh, India

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_11

11.1 Introduction

The discipline dealing with civil and criminal justice and makes use of numerous techniques implied in physical and biological sciences is Forensic science. It is the amalgamation of other branches of science such as biology, chemistry, physics, toxicology, pathology, odontology, entomology, anthropology, psychology, etc. (Bell 2006; Houck and Siegel 2006; Bogusz 2000). Therefore it can be said that forensic science covers broad interdisciplinary areas and technologies, and a lot of work has been going on across the globe to improve the mechanisms of instruments already in use and also to develop new techniques for better performance and results. One of the greatest achievements in this field is the introduction of DNA (Deoxyribonucleic Acid) typing methods and instrumentation to analyze even the minute differences among the genetic makeup of two individuals.

The judiciary system has benefited immensely by the modern and advanced techniques, especially by DNA profiling. These techniques help the law enforcement agencies to differentiate between the suspect and the innocent based on match–mismatch results. The trace amount of DNA sample in biological evidence present at the scene of the crime is sufficient enough to do the examination. Thus, the actual culprit can be reliably charged and the falsely accused person can be excluded (Butler 2001; Budowle et al. 2000).

Short tandem repeats (STR) were considered to be useful markers for differentiating two individuals as these are tandem repeat sequences unique to any individual. These 2–7 base pair repeats are repeatedly present adjacent to each other in variable number and thus generate an exclusive DNA profile for any person. The high utility of these STR loci in DNA typing was first adopted by the forensic and judicial system in the early 1990s and since then continued to be considered as the most accepted and admissible biological evidence in the court of justice (Butler 2006). Since then, rigorous efforts have been made by a forensic human identity testing committee to standardize the use of STRs in forensic studies (Butler et al. 2004; Budowle et al. 2000; Edwards et al. 1991; Kimpton et al. 1993). It has been more than two decades since the 13 STR markers, including CSF1PO, FGA, TH01, TPOX, VWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, and D21S11, were selected by the core of the FBI Laboratory's Combined DNA Index System (CODIS) in November 1997 (Budowle et al. 1999). In 2015, the original 13 CODIS Core Loci was expanded and seven new loci were introduced into it after reviewing the validation data obtained from two available PCR amplification kits: D1S1656, D2S441, D2S1338, D10S1248, D12S391, D19S433, and D22S1045. Multiplex PCR amplification is required to analyze all these markers in a single test. This idea is made possible by the availability of commercial STR multiplex systems, which make use of chemical dyes producing spectra of different fluorescence (Table 11.1).

The major advantage of using STR marker kits over the other DNA typing methods is the ease in their use, rapid analysis, and very low amount of DNA can also be analyzed as in case of degraded samples. As reported by the National Commission on the Future of DNA Evidence, these kits will continue to not only

Table 11.1 Commercially available multiplex kits

S. No.	Name of commercially available STR kit	Manufacturer	Markers and Dye set
1	Identifiler™, Identifiler Plus™, and Indentifiler Direct™	Life Technologies	Amelogenin +15 markers
2	PowerPlex® 16 and PowerPlex® 16 HS	Promega Corporation	Amelogenin +13 markers
3	PowerPlex® 18D	Promega Corporation	Amelogenin +15 markers
4	NGM™	Life Technologies	Amelogenin +15 markers, 5-dye chemistry
5	NGM SElect™ and NGM SElect™ express	Life Technologies	Amelogenin +16 markers, 5-dye chemistry
6	PowerPlex® ESX-16 and PowerPlex®ESI-16	Promega Corporation	Amelogenin +15 markers, 5-dye chemistry
7	PowerPlex® ESX-17 and PowerPlex®ESI-17	Promega Corporation	Amelogenin +16 markers, 5-dye chemistry
8	ESSplex and ESSplex SE	Qiagen, Hilden, Germany	Amelogenin +15 markers, 5-dye chemistry
9	PowerPlex® ES	Promega Corporation	Amelogenin +08 markers (including highly polymorphic SE33 locus)
10	SEfiler™ and SEfilerPlus™	Qiagen, Hilden, Germany	Amelogenin +10 markers
11	PowerPlex® 21	Promega Corporation	Amelogenin +20 markers
12	Globalfiler™ Express	Life Technologies	Amelogenin +23 markers
13	PowerPlex® Fusion	Promega Corporation	Amelogenin +23 markers
14	PowerPlex® CS7 System	Promega Corporation	07 short tandem repeat loci
15	Investigator HDplex kit	Qiagen, Hilden, Germany	Amelogenin +13 highly polymorphic markers
16	MiniFiler™	Life Technologies	Amelogenin +08 markers (first commercial kit to amplify miniSTRs)
17	PowerPlex® Y	Promega Corporation	12 Y-STR loci (first kit to test male lineage)
18	Yfiler™	Applied Biosystems	17 Y-STR loci
19	Argus Y-12 QS	Qiagen, Hilden, Germany	12 Y-STR loci + internal control
20	PowerPlex® Y23	Promega Corporation	12 Y-STR loci +06 additional loci
21	Argus X-12 kit	Qiagen, Hilden, Germany	12 X-STR loci

simplify the use of STRs in coming 5–10 years but also help in the enlargement of large and valuable DNA databases (Gill 2002; National Institute of Justice 2000). The present review discusses about the role of capillary electrophoresis in DNA typing methods and issues pertaining to its performance.

11.2 Basics of CE Along with Advantages over Flat Gel Bed-Based Technology

As the demand for rapid, high-throughput DNA diagnostic and screening applications has tremendously increased in the past few years, there is a dire need to develop new and effective analytical methods. In the early 1990s, the real question was how to move from a gel to a more advanced form, such as a capillary. Several pieces of research have been conducted to fulfill this requirement. X-linked acrylamide gel-filled capillaries were tried, but they showed some limitations in terms of their reusability, thermal degradation, and bubble formation. The alternate method was considered not to use a gel, but the question of the resolution was the main issue. Then during 1998–1990, Barry Karger's group utilized gel-filled capillaries to separate single-stranded DNA. They also introduced sieving polymers in the form of linear polyacrylamide to separate restriction digests. In 1992, commercial CE systems with lasers were introduced as Beckman P/ACE. This lead to a number of research works and the results were also very promising in terms of following areas: (1) microfabricated devices in which the electrophoresis is miniaturized in small capillaries; (2) matrix-assisted laser desorption/ionization time of flight-mode mass spectrometry; and (3) chip-based DNA sequencing and hybridization (Righetti et al. 2002).

Capillary electrophoresis (CE), a micro-separation technique, was first introduced by Hjerten et al. in the year 1967 for ions, micro and macromolecules separation, and chiral analysis (Ahuja and Jimidar 2008). CE techniques can also be used to separate and quantify drugs, their corresponding impurities, and metabolites (Holzgrabe et al. 2006; Holzgrabe 2008; Wätzig 2002). Since its introduction, tremendous progress has been achieved to make it more advance and useful in forensic analyses (Righetti et al. 2002). ABI 310 (Applied Biosystems) was introduced in July 1995 as the first commercially available multicolor CE. Capillary zone electrophoresis (CZE) is the best-modified version of CE, which is now available in a fully automated form to analyze the DNA length and polymorphism.

The mechanism of capillary electrophoresis is based on electrokinetic injection. The negatively charged DNA fragment is forced to move under the influence of a high voltage. The factors that influence the shape of the DNA molecule and thus migration of DNA fragments through the capillary include the electric field strength (E), injection time (t), the concentration of DNA in the sample, the area of the capillary opening ($\pi r2$), and the ionic strength of the medium.

The separation of DNA fragments during electrophoresis is influenced by the concentration and conductivity of the buffer solution. Resolution is negatively affected by the presence of highly concentrated DNA and high conductance. Buffer

solution charges the current molecules during electrophoresis as it dissolves and stabilizes the DNA fragment. Electrophoresis buffer enriched in urea, pyrrolidinone, EDTA, and formamide is used for consistent pH. Urea can be used for denaturing and viscosity, pyrolidinone for denaturing, and EDTA for stability and chelating metals. Formamide denatures DNA into single-stranded DNA due to its substantial denaturant property. The use of equal volumes of PCR product and high-quality formamide (<100 μS/cm) improves sensitivity. The run temperature of 60 °C helps reduce the secondary structure of DNA and enhances precision.

Band broadening effect due to electro-osmotic flow is one of the reasons for the lower resolution of DNA fragments running in capillaries. This can be resolved by using samples of lower ionic strength. This method is known as field amplified injection or stacking. Due to the absence of charge carriers in the sample zone, it abruptly ends the strong electric field at the interface by acting as a wall between the sample zone and buffer inside the capillary. The negatively charged DNA fragments will move quickly into the sample zone and start pre-concentrated at the interface and minimize the band broadening effect of the peak separation. However, the resolution problems can be avoided by the replacement of the capillary at regular intervals of time.

The air bubble formation in cross-linked polyacrylamide gels used earlier for size separation during capillary electrophoresis was the major disadvantage of these gels. The availability of uncross-linked polyacrylamide such as poly-dimethyl acrylamide matrices overcomes the two significant issues of sieving gels, i.e., viscosity and coating ability. Nowadays, performance-optimized polymers matrices are used as a separation medium in capillary electrophoresis with different trade names, like POP 4, POP 6, POP 7 (where POP stands for performance-optimized polymer) (Durney et al. 2015a). In a forensic DNA study, it has been demonstrated that POP-4 can resolve the DNA fragments having a single base difference up to 250 bases, and two-base resolution can be achieved up to 350 bases within a 31-min separation (Wenz et al. 1998).

The length and the concentration of the separation media affect the separation characteristics. Thus to improve the resolution, field strength and sample ionic strength should be lowered, and increasing the gel concentration and the polymer length enhances the resolution.

The advantages of capillary gel electrophoresis over the traditional gel electrophoresis are: gel used in capillary electrophoresis is not attached to the capillary wall and thus flows during the run; actually, it can be considered as not a gel, whereas, in slab gel, the gel attaches to the support slab. DNA molecules in CE interact with entangled polymers which are not cross-linked as in slab gels. The samples had to be loaded manually into the separation medium in traditional gel electrophoresis, whereas, in the automated CE systems, the same is automated and carried out through electrokinetic injection. This leads to precision and homogeneity in injection volumes across different runs. The gel in CE is pumpable, i.e., can be replaced after each run, and capillary can be refilled, but the same cannot be done in slab gel. Further, only a small quantity of DNA samples is needed in each injection.

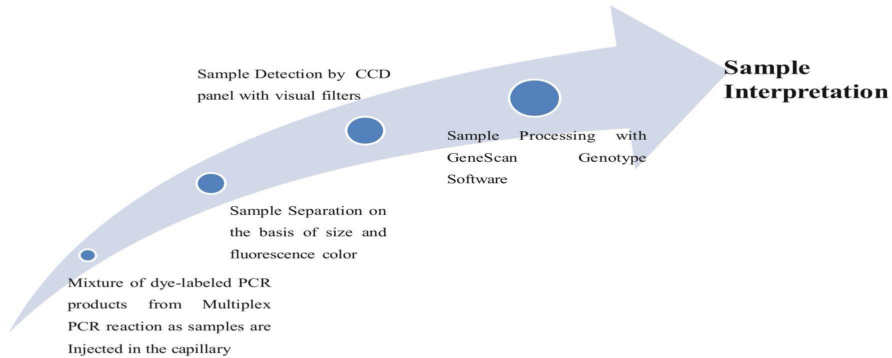


Fig. 11.1 Schematic illustration of separation and detection of STR alleles with genetic analyzer

The different sized bands of amplified DNA are obtained upon electrophoresis, which is subjected to either ultraviolet (UV) absorbance or laser-induced fluorescence (LIF) detector, so as to separate the polymerase chain reaction (PCR) products. The sensitivity of the UV detector is not up to the mark, and LIF has a unique requirement of fluorescence tag. Further, both the detector systems lack in providing the information required to determine the structure of the detected solute directly. The internal size standard with a different dye color is also usually co-electrophoresed to calibrate sizes. The specialized software based on the internal size standard's curve then analyzes the electropherogram of multicolored data and determines the size of STR alleles (Fig. 11.1).

The alleles are compared to the previously sequenced allelic ladder to generate the STR genotype profile (Butler et al. 2004). In the earlier capillary electrophoresis (CE) system such as ABI 310, the allelic ladder runs along with the internal size standard in one injection, and sample alleles with the same internal size standard are run in subsequent injections on the capillary in a sequential fashion (Lazaruk et al. 1998). However, the current technology uses capillary arrays, each array comprising of 4–24 capillaries with respect to different models of Sequencers. Hence the ladder can be run with samples simultaneously in a single injection, thus taking care of variations that may creep up during different injections.

There are specific characteristics required to be exhibited by separation and detection techniques used for accurate DNA typing of fluorescently labeled STR markers and these include: (1) reliable sizing methods for size range between 75 and 500 bp; (2) highly precise run-to-run comparison of allelic ladders with the sequentially processed STR samples; (3) efficient color differentiation of different dye sets used; (4) high resolution to detect microvariant alleles which are usually varying by 1 bp.

11.3 Applications of Forensic DNA Testing

The use of CE systems in analyzing forensic DNA typing leads to primarily two applications of this technology: forensic casework and the DNA database. Both applications have their issues and challenges.

11.3.1 Forensic Casework

It is mandatory to evaluate the DNA profiles generated by using the CE system through a comprehensive validation program of assessing the STR peaks (Moretti et al. 2001a, b). The DNA Advisory Board has set the operational criteria for this validation, which every forensic laboratory has to follow (Hares 2015; Gettings et al. 2019). These peaks must be assessed on a number of parameters such as peak imbalance, stochastic effects, stutter, and $n - 1$ peaks while considering the limitations of CE such as the sensitivity of the instrument, the threshold value of peak, heterozygote peak ratio, matrix assignment in pull-up issues pertaining to the spectral overlap of fluorescent dyes, etc. Similarly, the artifacts like spikes and dye blobs observed in an electrophoretic run should be reviewed and correctly identified, not as true allelic peaks. Thus with a thorough understanding of the CE system and its validation, DNA profiles of even mixed or degraded samples may be appropriately interpreted.

11.3.2 DNA Databasing

The database of DNA profiles generated during forensic studies is a useful forensic tool as it facilitates the rapid comparison of profiles. The updating of the database with new DNA profiles regularly increases the probability of a case-to-case match or case to convicted offender match. The most common problems related to forensic DNA database are (1) the backlog of convicted offender samples, which are not yet processed and included in the database and (2) second review of electropherograms of generated profiles is very time-consuming. It delays the uploading of convicted offender profiles into the database. However, many DNA laboratories have now developed highly automated and “expert” systems to handle this situation, which not only generates the DNA profile in the rapid time frame but also evaluates the electropherogram using specific criteria detailed by the examining laboratory.

11.4 Types of CE Platforms Available and Issues Related to Their Use

There have been rigorous attempts by the researchers to develop rapid, automated, and sensitive instrumentation for capillary electrophoresis. Earlier, the most common platform used globally for STR testing was the ABI 310 Genetic Analyzer. In

this single capillary instrument, about 48 DNA samples are processed in a serial fashion in a 24 h time period. There were many assumptions from the use of ABI 310 genetic analyzer such as DNA is a sphere that is not actually. The conditions for an unknown run are the same as the ladder run, but they are not. The ROX dye (carboxyrhodamine, an inert or passive fluorescent dye) migrates relatively the same as the FAM (fluorescein amidite) dye, but it does not. A calibration for one ladder is suitable for an entire run, which is accurate sometimes but not always. Temperature is constant, but to what precision it is unknown. Thus at this rate of approximately one sample per 30-min processing and other inappropriate assumptions, it cannot match the parallel processing potential throughput of a multilane slab-gel system. Further, in the 310 system, the autosampler calibration needs to be carried out prior to initiating a new run. This was indeed, time-consuming. None the less, the ABI 310 remains the most important model for training purposes for learning automated capillary electrophoresis.

However, the search for updated and more advanced instruments for rapid, automated systems lead to the availability of an updated version of the genetic analyzer in 2001. It was a 16-capillary ABI 3100, which soon dominated many laboratories as being the choice of instrument for the next level of throughput. Its four capillary versions, ABI 3100 Avant, were also available. The polymer injection in ABI 310 and ABI 3100 was based on a syringe system, wherein the polymer was filled into a syringe and the syringe was mounted onto the instrument. The polymer injections were, however, carried out by the instrument from the syringe. These models gave the flexibility to the user to fill polymer into the syringe as per the requirement. Hence if a limited number of runs were to be carried out, the user could fill a limited volume of polymer into the syringe. ABI 3100 was followed by ABI 3130 and ABI 3130 XL models with 4 and 16 capillaries, respectively. The models were based on ABI 3100 platform; however, in these models, the syringe system for loading polymer was replaced by a polymer delivery pump. Hence a standard vial containing POP was attached to the polymer delivery pump assembly and filling of the unit and capillaries became automated. ABI 3100 and ABI 3130 are currently in use, but manufacturing is discontinued. These models were followed by ABI 3500 series (ABI 3500 and ABI 3500 XL, 8 and 24 capillary, respectively). These machines have further made automated CE more user-friendly, as the user is not required to prepare dilutions of anode or cathode buffer; instead, the buffers come as ready to load trays, where the concentrations have been highly optimized, thus bringing in more precision to the CE work. Further, the reagents come with RFID (Radio Frequency Identification) tags, such that the machine keeps the log of loaded reagents, including their life/expiry. This makes the system highly reliable and precise that provides an assurance of quality, which is vital in forensic DNA analysis. Nowadays several commercial capillary array electrophoresis (CAE) instruments are available, such as ABI 3700 or ABI 3730 (96-capillary), SpectrumMedix (96–384 capillary), which increases the sample throughput capabilities by offering parallel run of up to 384 capillaries in an independent environment and also demonstrates the potential of multi-capillary CE systems to give reliable results (Butler et al. 2004; Durney et al. 2015a). However, as far as

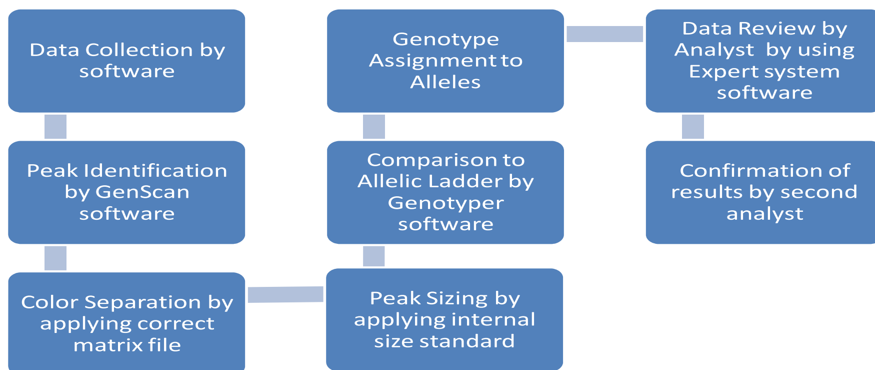


Fig. 11.2 Schematic representation of steps involved in data processing by genetic analyzer

forensic DNA typing for HID (Human Identity) is concerned, the mainstay instruments remain ABI 3100, 3130, and 3500 series. These systems work in a simple generalized systematic stepwise process, which is graphically summarized in Fig. 11.2.

These instruments were developed by using the prototypes of Huang et al. (1992), Kambara and Takahashi (1993), and Ueno and Yeung (1994), respectively. The molecular dynamics instrumentation includes confocal detector system with four photomultiplier tubes having appropriate filters and dichroic beam splitters to focus the laser beam, which excites the fluorescent molecules in the liquid sheath, cool CCD (charge-coupled device) camera for accumulating and imaging the emitted light, and a scanner to assemble the data from all the capillaries. Research has been advanced in order to improve the detection system. Kambara et al. intercalated a glass rod of the same external diameter with each capillary and thus ameliorated the light transmission with side illumination (Schmalzing et al. 1997). Similarly, another group used a single beam to illuminate and collect light from all the capillaries (Woolley and Mathies 1994). Further development in this direction continued in the form of using capillaries in a wheel-like fashion (Woolley et al. 1997; Medintz et al. 2001). The most advanced kind of CAE instrument makes use of a low power laser line generator, holographic transmission grating, and a back-illuminated CCD chip (Schmalzing et al. 1999).

The common problem associated with most of the CAE instruments is the limitation of reading length to 500 bp only, while there is a requirement to sequence at least 800 bp with 99% accuracy. Thus to combat this problem, Endo et al. proposed a solution of using gradients of electric field strength in an acceleration, plateau, decreased, and a constant cycle ranging from 90–130 V/cm up to 220 V/cm at a column temperature of 60 °C (Carey and Mitnik 2002).

11.4.1 Microfabricated Capillary Arrays

Microfabricated or micromachined or microchip devices are another platform of capillary array electrophoresis (Kovacs et al. 1996). These can be prepared by either using photolithography on glass or silicon substrates (Jacobson et al. 1994a) or making use of molded plastic for a micromachined template (Duffy et al. 1998). Several researchers have used various approaches to evaluate the relationship between different read lengths and applied voltage systems (Simpson et al. 1998; Schmalzing et al. 1998; Liu et al. 1999). Their studies revealed that decreased voltage results in an increased resolution.

11.4.2 Microfluidic Chips for DNA Sizing and Quantitation

Ueda et al. adopted microfabricated chips with 6-mm separation channels and highly sensitive silicon intensified target camera for rapid analysis of triplet repeat DNA fragments in 12 s, which was 20 times faster than other conventional separation capillary zone electrophoresis (CZE) instrumentations (Ueda et al. 2000).

The higher density capillary arrays can be generated by microfabrication, which facilitates physical assemblies and makes a direct link between capillaries and other devices on the chip. There is continuous progress in the development of microfluidic systems, which gives the researcher a liberation to assay a wide array of samples such as buccal cells (Gelfi et al. 1998) and whole blood from multiple donors in a very cost-effective and rapid manner (Fisher and Lerman 1983; Gelfi et al. 1994, 1996).

Although several modifications have also been done in microfabrication technique such as the inclusion of gas or liquid chromatography to the instrument (Terry et al. 1979; Manz et al. 1990) or fabrication of individual CE chips to separate fluorescent dyes (Manz et al. 1992; Jacobson et al. 1994b) or fluorescently labeled amino acids (Harrison et al. 1993; Effenhauser et al. 1993; Fan and Harrison 1994), yet questions pertaining to usefulness of these techniques for resolving DNA at a higher level or fabrication of multiple separation channels in a single chip remained answered since a long time. It was reported in one study that glass substrates can be modified for excellent high-speed resolution of DNA fragments by using photography and chemical etching methods on microfabricated CE separation channels, thus establishing the feasibility of integrated devices for electrophoretic DNA analysis.

11.5 Future Directions of Emerging and Expanding Technology

The limitations of existing capillary electrophoresis systems either single-capillary ABI 310 or multi-capillary ABI 3100, ABI 3130, or ABI 3500 instruments need to be rectified so that more rapid and easier STR typing technologies can be developed. Some factors which must be considered in exceeding the capabilities of current analysis systems include: (1) simultaneous analysis of PCR reactions with high

resolution of at least five fluorescent dyes without any pull-ups, (2) maintenance of single-base resolution ranging from 50 to 250 or 500 bp, and (3) adoption of time-of-flight mass spectrometry approaches.

The innovative steps taken for the development of capillary electrophoresis in advancing DNA separation methods paved to new scientific discoveries in terms of generating novel sieving gels, artificial gels, aptamers, or novel mechanisms of DNA separations. A recent report suggested an alternative analysis method for stability and conformation analyses. In this study, the use of monophosphate nucleosides supports the sequence-based selectivity of DNA fragments (Zhang and McGown 2013). Capillary electrophoresis may be modified by the use of metal ions, which influence the conformational changes in DNA (Li et al. 2009). Indirect molecular detection can be done by using DNA aptamers because the aptamer displacement is associated with conformational changes. Thus, it can be used as a molecular label in multiplexed separation based assays (Perrier et al. 2014).

11.5.1 Transformed Gels for Improved Analyses

Phospholipid nanogels may be used as a suitable material for DNA analysis as it is responsive to thermal variations and can accommodate sieving separation in different viscosities (Durney et al. 2013, 2015b). Examples of such matrices are triblock copolymers of poly ethylene glycol (PEG) and poly propylene glycol (PPG), commonly known as Pluronics, A-B-A, where A= Poly Ethylene Oxide (PEO) and B= Poly Propylene Oxide (PPO) (Kim et al. 2010; http://www.sigmaaldrich.com/catalog/product/aldrich/542342?lang=en®ion=US).

The major advantage of using these hydrophilic micelles containing matrices is the availability of dynamic surface coating, which provides a chance to DNA analytes to interact with each other favorably and thus can be resolved on the basis of their minute secondary structural conformations (Choi et al. 2014; Hwang et al. 2014).

11.5.2 Artificial Matrices

The matrices can be prepared artificially by incorporating electrically insulating materials within the separation channels at micrometer to sub-micrometer level. Few examples include pillar arrays, self-assembling colloidal crystals, and nanoscale channels. Different parameters have been involved in studies using pillar arrays, such as the effects of geometry (Chan et al. 2009), size (Ou et al. 2011), order (Olson and Dorfman 2012), and space (Ou et al. 2010; Park et al. 2012) on separation performance. Research has also been done on obstacles in fabrication, which provides insight regarding Ogston transport and the effect of the relationship between channel dimensions and DNA length on mobility (Wang et al. 2013). The knowledge of mobility shift and its relationship with entropy, degree of interaction between DNA and channel wall, and ionic strength of buffer may be harnessed to

adjust the size of the fabricated matrix according to the available DNA fragments in the experiment (Kounovsky-Shafer et al. 2013; Russell et al. 2014).

The matrix made up of monodispersed colloidal crystals is a high-performance fabricated microfluidic choice for obtaining optimal resolution in DNA analyses at low cost when compared with a matrix composed of different size colloidal particles (Nazemifard et al. 2012).

11.6 Conclusion

DNA typing technologies are playing a crucial role in forensic analysis and human identification studies. Many pieces of research are conducted globally to modernize and ease this process. The factors which played a critical role in the advancement of DNA typing include the use of STRs, multiplex PCR, multichannel laser-induced fluorescence, and use of entangled polymer buffers in capillary electrophoresis. DNA analysis technologies have been developed a lot with the advent of capillary electrophoresis, which is not only used for sequencing and sizing but also the discovery of aptamers and affinity studies. There are lot many advantages of capillary electrophoresis in separating macromolecules like DNA, which makes it an excellent choice for DNA analysis such as the injection, separation, detection are automated in advanced capillary electrophoresis techniques with ease to retrieve the peak information from the automatically stored data. It has excellent sensitivity and resolving power and the use of this technique makes rapid separations possible in real time.

Although with the emergence of novel approaches, capillary electrophoresis techniques have evolved from development stages to validated and applied methods, yet further optimization of capillary electrophoresis methods will continue to be pivotal in expanding the field of DNA analysis.

Thus the development of techniques to better understand separations in capillary, such as artificial matrices, suggests that better results can be obtained by reviewing following factors: (1) precision and resolution of DNA typing systems are must for reliable identification of STR, (2) effect of temperature results in a differential response, (3) for reproducibility denaturation plays a significant role, (4) polymeric buffer and injection media composition control the separation efficiency and precision. The automation of testing procedures, including PCR, robustness, and long-term stability of these systems, may enhance the sample throughput and thus must be taken under consideration while doing complex assays.

References

Ahuja S, Jimidar MI (2008) Capillary electrophoresis methods for pharmaceutical analysis, 1st edn. Elsevier, Amsterdam

Bell S (2006) Forensic chemistry. Pearson-Prentice Hall, Upper Saddle River

Bogusz MJ (ed) (2000) Handbook of analytical separations. Forensic science, vol 2. Elsevier, Amsterdam
Budowle B, Moretti TR, Baumstark AL, Defenbaugh DA, Keys KM (1999) Population data on the thirteen CODIS core short tandem repeat loci in African Americans, U.S. Caucasians, Hispanics, Bahamians, Jamaicans, and Trinidadians. J Forensic Sci 44:1277–1286
Budowle B, Smith J, Moretti T, DiZinno JA (2000) DNA typing protocols: molecular biology and forensic analysis. Eaton Publishing, Natick, MA, pp 157–188
Butler JM (2001) Forensic DNA typing: biology and technology behind STR markers. Academic Press, San Diego, CA
Butler JM (2006) Genetics and genomics of core short tandem repeat loci used in human identity testing. J Forensic Sci 51:253–265
Butler JM, Buel E, Crivellente F, McCord BR (2004) Forensic DNA typing by capillary electrophoresis using the ABI prism 310 and 3100 genetic analyzers for STR analysis. Electrophoresis 25:1397–1412
Carey L, Mitnik L (2002) Trends in DNA forensic analysis. Electrophoresis 23:1386–1397
Chan YC, Zohar Y, Lee Y-K (2009) Effects of embedded submicron pillar arrays in microfluidic channels on large DNA electrophoresis. Electrophoresis 30(18):3242–3249
Choi W, Shin GW, Hwang HS, Pack SP, Jung GY, Jung GY (2014) A multiplex single nucleotide polymorphism genotyping method using ligase-based mismatch discrimination and CESSCP. Electrophoresis 35(8):1196–1203
Duffy DC, McDonald JC, Schueller OJA, Whitesides GM (1998) Rapid prototyping of microfluidic systems in poly(dimethylsiloxane). Anal Chem 70:4974–4984
Durney BC, Lounsbury JA, Poe BL, Landers JP, Holland LA (2013) A thermally responsive phospholipid pseudogel: tunable DNA sieving with capillary electrophoresis. Anal Chem 85 (14):6617–6625
Durney BC, Crihfield CL, Holland LA (2015a) Capillary electrophoresis applied to DNA: determining and harnessing sequence and structure to advance bioanalyses (2009–2014). Anal Bioanal Chem 407:6923–6938. https://doi.org/10.1007/s00216-015-8703-5
Durney BC, Bachert BA, Sloane HS, Lukomski S, Landers JP, Holland LA (2015b) Reversible phospholipid nanogels for DNA fragment size determinations up to 1,500 base pairs and integrated sample stacking. Anal Chim Acta 880:136. https://doi.org/10.1016/j.aca.2015.03.009
Edwards A, Civitello A, Hammond HA, Caskey CT (1991) DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am J Hum Genet 49:746–756
Effenhauser CS, Manz A, Widmer HM (1993) Glass chips for high-speed capillary electrophoresis separations with submicrometer plate heights. Anal Chem 65:2637–2642
Fan ZH, Harrison DJ (1994) Micromachining of capillary electrophoresis injectors and separators on glass chips and evaluation of flow at capillary intersections. Anal Chem 66:177–184
Fisher SG, Lerman LS (1983) DNA fragments differing by single base-pair substitutions are separated in denaturing gradient gels: correspondence with melting theory. Proc Natl Acad Sci U S A 80:1579–1583
Gelfi C, Righetti PG, Cremonesi L, Ferrari M (1994) Detection of point mutations by capillary electrophoresis in liquid polymers in temporal thermal gradients. Electrophoresis 15:1506–1511
Gelfi C, Cremonesi L, Ferrari M, Righetti PG (1996) Temperature-programmed capillary electrophoresis for detection of DNA point mutations. BioTechniques 21:926–932
Gelfi C, Curcio M, Righetti PG, Sebastiano S, Citterio A, Ahmadzadeh H, Dovichi N (1998) Surface modification based on Si-O, Si-C sublayers and a series of *n*-substituted acrylamide top layers for capillary electrophoresis. Electrophoresis 19:1677–1682
Gettings KB, Ballard D, Bodnerc M, Borsuk LA, King JL, Parsonc W, Phillips C (2019) Report from the STRAND Working Group on the 2019 STR sequence nomenclature meeting. Forensic Sci Int Genet 43:102165
Gill P (2002) Role of short tandem repeat DNA in forensic casework in the UK—past, present, and future perspectives. BioTechniques 32:366–372

Hares DR (2015) Selection and implementation of expanded CODIS core loci in the United States. Forensic Sci Int Genet 17:33–34

Harrison DJ, Fluri K, Seiler K, Fan Z, Effenhauser CS, Manz A (1993) Micromachining a miniaturized capillary electrophoresis-based chemical analysis system on a chip. Science 261:895–897

Holzgrabe U (2008) The need for CE methods in pharmaceutical monographs. In: Ahuja S, Jimidar MI (eds) Capillary electrophoresis methods for pharmaceutical analysis, 1st edn. Elsevier, Amsterdam, pp 245–258

Holzgrabe U, Brinz D, Kopec S, Weber C, Bitar Y (2006) Why not using capillary electrophoresis in drug analysis? Electrophoresis 27:2283–2292

Houck MM, Siegel JA (2006) Fundamentals of forensic science. Elsevier/Academic Press, Burlington

http://www.sigmaaldrich.com/catalog/product/aldrich/542342?lang=en®ion=US

Huang XC, Quesada MA, Mathies RA (1992) DNA sequencing using capillary array electrophoresis. Anal Chem 64:967–972

Hwang HS, Shin GW, Jung GY, Jung GY (2014) A simple and precise diagnostic method for spinal muscular atrophy using a quantitative SNP analysis system. Electrophoresis 35(23):3402–3407

Jacobson SC, Hergenroder R, Koutny LB, Warmack RJ, Ramsey JM (1994a) Effects of injection schemes and column geometry on the performance of microchip electrophoresis devices. Anal Chem 66:1107–1113

Jacobson SC, Hergenroeder R, Koutny LB, Ramsey JM (1994b) High-speed separations on a microchip. Anal Chem 66(7):1114–1118

Kambara H, Takahashi S (1993) Multiple-sheath flow capillary array DNA analyser. Nature 361:565–566

Kim SJ, Shin GW, Choi SJ, Hwang HS, Jung GY, Seo TS (2010) Triblock copolymer matrix-based capillary electrophoretic microdevice for high-resolution multiplex pathogen detection. Electrophoresis 31(6):1108–1115

Kimpton CP, Gill P, Walton A, Urquhart A, Millican ES, Adams M (1993) Automated DNA profiling employing multiplex amplification of short tandem repeat loci. PCR Methods Appl 3:13–22

Kounovsky-Shafer KL, Hernández-Ortiz JP, Jo K, Odijk T, de Pablo JJ, Schwartz DC (2013) Presentation of large DNA molecules for analysis as nanoconfined dumbbells. Macromolecules 46(20):8356–8368

Kovacs GTA, Petersen K, Albin M (1996) Fabricating large arrays of microwells with arbitrary dimensions and filling them using discontinuous Dewetting. Anal Chem 68:407A–412A

Lazaruk K, Walsh PS, Oaks F, Gilbert D, Rosenblum BB, Menchen S, Scheibler D, Wenz HM, Holt C, Wallin J (1998) Genotyping of forensic short tandem repeat (STR) systems based on sizing precision in a capillary electrophoresis instrument. Electrophoresis 19:86–93

Li T, Zhang D, Luo W, Lu M, Wang Z, Song Y, Wang H (2009) Metal cation mediated-capillary electrophoresis of nucleic acids. Anal Chem 82(2):487–490

Liu S, Shi Y, Ja WW, Mathies RA (1999) Optimization of high-speed DNA sequencing on microfabricated capillary electrophoresis channels. Anal Chem 71:566–573

Manz A, Miyahara Y, Miura J, Watanabe Y, Miyagi H, Sato K (1990) Design of an open-tubular column liquid chromatograph using silicon chip technology. Sens Actuators B Chem 1:249–255

Manz A, Harrison DJ, Verpoorte EMJ, Fettinger JC, Paulus A, Ludi H, Widmer HM (1992) Planar chips technology for miniaturization and integration of separation techniques into monitoring systems. J Chromatogr 593:253–258

Medintz IL, Berti L, Emrich CA, Tom J, Scherer JR, Mathies RA (2001) Genotyping energy-transfer-cassette-labeled short-tandem-repeat amplicons with capillary array electrophoresis microchannel plates. Clin Chem 47:1614–1621

Moretti TR, Baumstark AL, Defenbaugh DA, Keys KM, Brown AL, Budowle B (2001a) Validation of STR typing by capillary electrophoresis. J Forensic Sci 46:661–676

Moretti TR, Baumstark AL, Defenbaugh DA, Keys KM, Smerick JB, Budowle B (2001b) Validation of short tandem repeats (STRs) for forensic usage: performance testing of fluorescent multiplex STR systems and analysis of authentic and simulated forensic samples. J Forensic Sci 46:647–660

National Institute of Justice (2000) The future of forensic dnatesting: predictions of the research and development Working Group of the National Commission on the Future of DNA Evidence, Washington, DC. http://www.ojp.usdoj.gov/nij/pubs-sum/183697.htm

Nazemifard N, Wang L, Ye W, Bhattacharjee S, Masliyah JH, Harrison DJ (2012) A systematic evaluation of the role of crystalline order in nanoporous materials on DNA separation. Lab Chip 12(1):146–152

Olson DW, Dorfman KD (2012) Experimental study of the effect of disorder on DNA dynamics in post arrays during electrophoresis. Phys Rev E 86(4):041909

Ou J, Carpenter SJ, Dorfman KD (2010) Onset of channeling during DNA electrophoresis in a sparse ordered post array. Biomicrofluidics 4(1):013203

Ou J, Joswiak MN, Carpenter SJ, Dorfman KD (2011) Plasma thinned nanopost arrays for DNA electrophoresis. J Vac Sci Technol A 29(1):011025

Park S-G, Olson DW, Dorfman KD (2012) DNA electrophoresis in a nanofence array. Lab Chip 12 (8):1463–1470

Perrier S, Zhu Z, Fiore E, Ravelet C, Guieu V, Peyrin E (2014) Capillary gel electrophoresis-coupled aptamer enzymatic cleavage protection strategy for the simultaneous detection of multiple small analytes. Anal Chem 86(9):4233–4240

Righetti PG, Gelfi C, D’Acunto MR (2002) Recent progress in DNA analysis by capillary electrophoresis. Electrophoresis 23:1361–1374

Russell AJ, Del Bonis-O’Donnell JT, Wynne TM, Napoli MT, Pennathur S (2014) Separation behavior of short single- and double-stranded DNA in 1 micron and 100 nm glass channels. Electrophoresis 35(2–3):412–418

Schmalzing D, Koutny L, Adourian A, Belgrader P, Matsudaira P, Ehrlich D (1997) DNA typing in thirty seconds with a microfabricated device. Proc Natl Acad Sci U S A 94:10273–10278

Schmalzing D, Adourian M, Koutny L, Ziaugra L, Matsudaira P, Ehrlich D (1998) DNA sequencing on microfabricated electrophoretic devices. Anal Chem 70:2303–2310

Schmalzing D, Koutny L, Chisholm D, Adourian A, Matsudaira P, Ehrlich D (1999) Two-color multiplexed analysis of eight short tandem repeat loci with an electrophoretic microdevice. Anal Biochem 270:148–152

Simpson PC, Roach D, Woolley AT, Thirsen T, Johnston R, Sensabaugh GF, Mathies RA (1998) High-throughput genetic analysis using microfabricated 96-sample capillary array electrophoresis microplates. Proc Natl Acad Sci U S A 95:2256–2261

Terry SC, Jerman JH, Angell JB (1979) A gas chromatographic air analyzer fabricated on a silicon wafer. IEEE Trans Electron Devices 26:1880–1886

Ueda M, Kiba Y, Abe H, Arai A, Nakanishi H, Baba Y (2000) Fast separation of oligonucleotide and triplet-repeat DNA on a microfabricated capillary electrophoresis device and capillary electrophoresis. Electrophoresis 21:176–180

Ueno K, Yeung ES (1994) Simultaneous monitoring of DNA fragments separated by electrophoresis in a multiplexed array of 100 capillaries. Anal Chem 66:1424–1431

Wang CH, Ho CH, Chou YC (2013) Sieving DNA molecules by length dependence in artificial nano-channel matrices. J Appl Phys 113(2):024701–024707

Wätzig HK (2002) Kapill are lektrophorese in der pharmazeutischen Industrie. Pharm Ind 64:1183–1185

Wenz H, Robertson JM, Menchen S, Oaks F, Demorest DM, Scheibler D, Rosenblum BB, Wike C, Gilbert DA, Efcavitch JW (1998) High-precision genotyping by denaturing capillary electrophoresis. Genome Res 8(1):69–80

Woolley AT, Mathies RA (1994) Ultra-high-speed DNA fragment separations using microfabricated capillary array electrophoresis chips. Proc Natl Acad Sci U S A 91:11348–11352

Woolley AT, Sensabaugh GF, Mathies RA (1997) High-speed DNA genotyping using microfabricated capillary array electrophoresis chips. Anal Chem 69:2181–2186

Zhang X, McGown LB (2013) Sequence-based separation of single-stranded DNA using nucleotides in capillary electrophoresis: focus on phosphate. Electrophoresis 34(12):1778–1786

Human Trafficking and DNA Analysis

12

María Saiz, Maria Jesus Alvarez-Cubero, Begona Lorente-Remon, Luis Javier Martínez-Gonzalez, and Jose A. Lorente

Abstract

One of the most impact and widespread violations of human rights is human trafficking. Historical events and laws defining human trafficking have changed along recent years. Here we report the principal consequences of human trafficking indicating differences among sex and countries all over the world. Here we highlight the increasing rates of human trafficking in the cases of people escaping from war and persecutions, which have become in the recent years one of the most vulnerable events for human trafficking. Moreover, it is reported the reinforced role of universal and global laws for unifying this problem and victims. Finally, forensic strategies such as DNA typing have been one of the most relevant options for having a specific human identification in relevant cases such as human trafficking.

M. Saiz · B. Lorente-Remon
Laboratory of Genetic Identification, Faculty of Medicine, University of Granada, Granada, Spain

M. J. Alvarez-Cubero
Department of Biochemistry and Molecular Biology III, Faculty of Medicine, University of Granada, Granada, Spain

GENYO (University of Granada-Andalusian Government—Pfizer, Centre for Genomics and Oncological Research), Granada, Spain

L. J. Martínez-Gonzalez
GENYO (University of Granada-Andalusian Government—Pfizer, Centre for Genomics and Oncological Research), Granada, Spain

J. A. Lorente (✉)
Laboratory of Genetic Identification, Faculty of Medicine, University of Granada, Granada, Spain

GENYO (University of Granada-Andalusian Government—Pfizer, Centre for Genomics and Oncological Research), Granada, Spain
e-mail: jose.lorente@genyo.es

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_12

Keywords

DNA identification · Human trafficking · Databases · Missing persons · Criminalistic

12.1 Introduction

The United Nations Office on Drugs and Crime (UNODC) has reported that a total of 63,251 victims were detected in 106 countries and territories between 2012 and 2014. A clear majority (around 70%) of these victims were females; moreover, the trends in men are increasing these late years; however, with clear differences among countries. Children remain being the second most commonly detected group of victims (after women) of trafficking globally, ranging from 25 to 30% of the total during the period of 2012–2014 (UNODC 2016).

The main purposes of human trafficking are wide; ranking the main aims the sexual exploitation and forced labour. But there are also victims used as beggars, for forced marriages, pornography or organ trafficking, etc. Around 72% of the cases are related to sexual exploitation, contrasting to the 20–30% for forced labour in women trafficking victims; around 85.7% for forced labour and a 6.8% in trafficking for sexual exploitation in men trafficking victims. There are also differences among the countries in the human trafficking aims, for example in the Western and Southern Europe comparing to the Central and the South-Eastern Europe (UNODC 2016).

Recent circumstances of people fugitives from war and persecutions are also becoming trafficking vulnerable victims. A recent situation like this has happened in Syria with a large amount of refugees victims of the latest conflict in this region (UNODC 2016).

Moreover, one of the most increasing cases of current human trafficking is focused on organ illegal transplants, making organ trafficking a relevant problem. According to the WHO (World Health Organization) around 11,000 human organs were obtained on the black market in 2010. It is a reality in many parts of the world, like Indonesia, China, India, South Africa, Brazil, among others. This is mainly due to the increasing demand of organ transplants, especially kidney (What You Need to Know about Illegal Human Organ Trafficking | Big Think n.d.).

Therefore, this problem affects the whole world and multiple projects and institutions act to minimize it (Más Allá de las Fronteras: Explorando los Enlaces entre Trata, Globalización y Seguridad 2010). FBI provides victims with assistance benefits and improves current cases, mainly by the Office for Victim Assistance with victim specialists nationwide (Human Trafficking/Involuntary Servitude—FBI n.d.).

Many of the current projects, which are focused on human trafficking, include as the main tool for making human identification the use of DNA typing. It is currently widely used in immigration, with high relevance in important cases such as whether to advance or stop with the requests for immigration or international adoptions. Moreover, biological tests can also raise information on unidentified human remains and missing persons. Taking these two applications into account, genetic

technologies can be used to identify missing individuals in cases of human trafficking (Palmbach et al. 2014).

12.2 Worldwide Initiatives Against Human Trafficking and Exploitation

We have to guarantee children to be treated kindly and following the national and international standards. Many initiatives have been created to prove that governments know the importance of the problem of human trafficking and the need to exterminate the oppression of children. International agreements and policies must be in place to follow and detect the signs, to transfer the data and to interconnect all this information for having an active prosecution of committers as well as for the reintegration of child victims. Global coordination plays a critical role, but legal and social difficulties exist by data inclusion entree and disclosure among countries. This considerably minimizes the progress of establishing a comprehensive and efficient counter-trafficking system (Saiz et al. 2016).

12.2.1 UNODC–UN.GIFT

In exertion to protect children (and women) and to combat human trafficking, the United Nations initiated the *UN-Global Initiative to Fight Human Trafficking* (UN.GIFT) in March 2007 http://www.unodc.org/ (United Nations Office on Drugs and Crime n.d.). The UN.GIFT was launched to encourage the whole fighting against human trafficking based on international agreements reached by the United Nations. So far, around 150 groups have signed the procedure for the prevention, suppression and punishment against human trafficking which ended in the Palermo Convention against the transnational organized crime.

In March 2007, the International Labour Organization (ILO) launched the UN.GIFT by the Office of the High Commissioner for Human Rights in Human Nations (OHCHR), the Organization for Safety and Cooperation in Europe (OSCE), the International Organization for Migrations (IOM), the United Nations Office on Drugs and Crime (UNODC) and the United Nations Children's Fund (UNICEF).

The Global Initiative is established on a not laborious standard: human trafficking is a crime of such scale and barbarism that it cannot be handled with effectively by governments not working together. This problem needs a universal, multi-stakeholder approach based on national efforts all over the world. By promoting and facilitating collaboration and organization, UN.GIFT proposes to do synergies among the anti-trafficking activities of UN agencies, international organizations and other participants to create the most active and cost-effective strategies and good practices (UNODD 2009). UN.GIFT aims to activate state and non-state actors to get rid of human trafficking through the reduction of both the vulnerability of possible victims and the claim for exploitation in all its forms. These guarantee a satisfactory assurance and support to those who fall victim and encouraging the effective

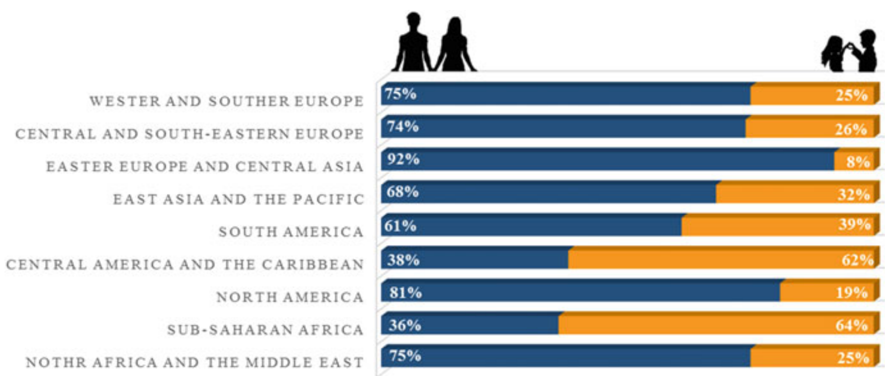


Fig. 12.1 Shares of adults and children among detected victims of human trafficking, by regions. (Figure edited from the Global Report on Trafficking in Persons 2016 (UNODC 2016))

prosecution of the criminals implicated, whilst maintaining the essential all persons human rights.

For the execution of its mission, UN.GIFT grows the knowledge and social conscience on human trafficking; promotes operative rights-based responses and inspires the relation of corporations for combined action against human trafficking (Fig. 12.1).

12.2.2 International Organization for Migrations (IOM)

IOM, firstly identified as Provisional Intergovernmental Committee for the Movements of Migrants from Europe (PICMME), was created in 1951 in response to the chaos and migrations to occidental Europe after 2nd World War. IOM's history tracks the tragedies caused by humans and natural disasters in the past half-century: Hungary in 1956, Czechoslovakia in 1968, Chile in 1973, refugees from Vietnam (known as "Boat people") in 1975, Kuwait in 1990, Kosovo and Timor in 1999, the Asiatic tsunami and the earthquake of Pakistan in 2004 and 2005. IOM is faithful to the belief that humanitarian and arranged migration benefits migrants and society.

Since IOM was created as a logistic society, it has extended its effort to be the main international organization that deals with governments and civil society to endorse the comprehension of migrations, rise the socio-economically development through migrations and defend the honour and security of migrants.

The growing impact of its actions goes next to IOM expansion that has evolved from a little organism to an Organization with a budget around 1.3 millions of Dollars (USA) and a staff of 9000 people in more than a 150 of countries all over the world.

IOM is the benchmark in the world discussion of social, economic and political repercussions of migrations in the twenty-first century. The IOM has been working

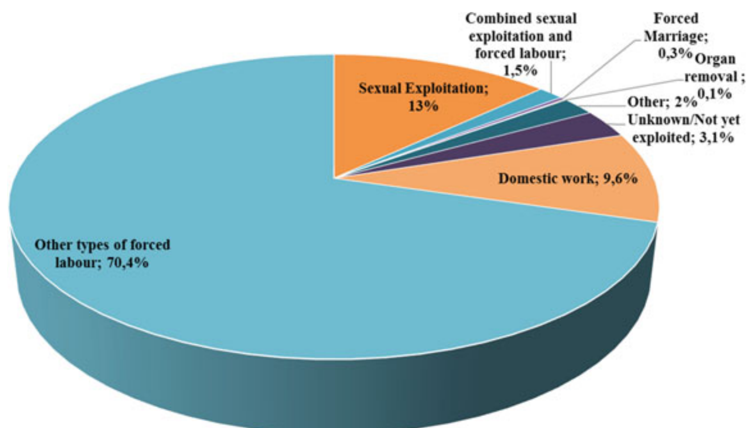


Fig. 12.2 Main causes of human trafficking reported by IOM in 2016 (International Organization for Migration 2017)

to prevent persons trafficking since 1994. By the moment, it has put into effect more than 500 projects in more than 85 countries and has provided support to nearly 90,000 trafficked persons. Its main objectives are to counter persons trafficking and to give protection to victims from the trade as well as offering options of harmless and bearable reintegration and/or return to their home countries.

IOM has assisted about 8500 trafficked persons in 2016. Eighty-six percent of these persons were adults and 14% were under 18 years old. Of all these cases, 43% were women and 57% were men (International Organization for Migration 2017). The main purposes of human trafficking are shown in Fig. 12.2.

This initiative shows that administrations are conscious of the real problem of human trafficking and that several footsteps should be done to fight against child trafficking. However, there are many difficulties to overcome in human trafficking. International agreements to recognize, search for, communicate and share data, as well as stop the criminals and support the victims have to be well known. Worldwide legal and political management plays a critical role, but legal and social difficulties hamper the development of the application of a system against human trafficking.

12.2.3 International Labour Organization (ILO): International Programme on the Elimination of Child Labour

The concept "child labour" is often defined as work that denies children of their childhood, their potential and their dignity and that is damaging to physical and mental growth. Children separated from their families, being enslaved, exposed to serious dangers and illnesses and/or abandoned to shift for themselves on the streets of huge cities—often at a very early age are the most extreme forms of child labour.

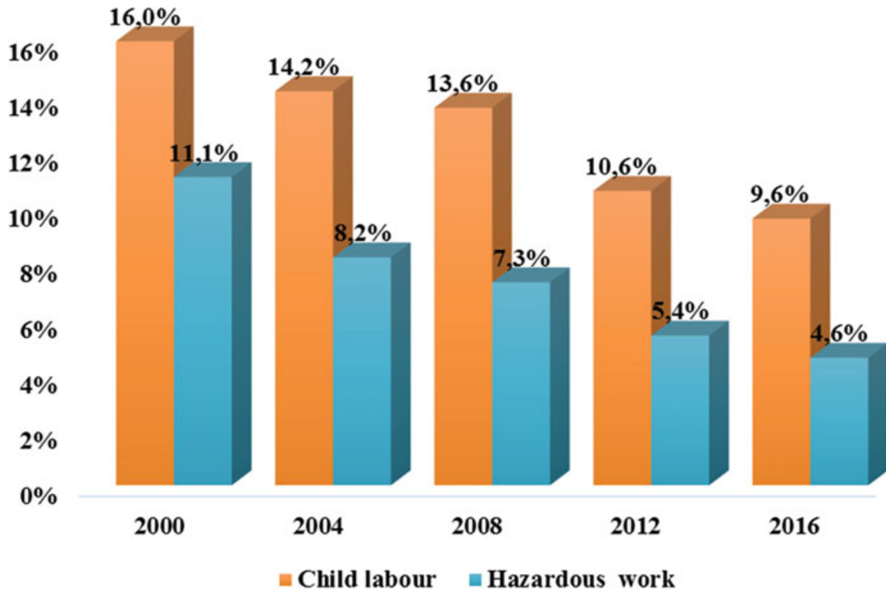


Fig. 12.3 Percentage of children's involvement in child labour and hazardous work (5–17 years age range), 2000–2016 (Internation Labour Office 2017)

The ILO's International Programme on the Elimination of Child Labour (IPEC) was created in 1992 with the whole goal of the abolition of child labour through consolidating the ability of countries to deal with the problem and encouraging a worldwide programme to fight child labour. At present, IPEC operates in 107 countries, with an annual spending on technical collaboration projects that reached over US$61 million in 2008. It is the biggest programme of its kind globally and the largest single operational programme of the ILO. Since its beginning in 1992, IPEC has worked to eliminate child labour in numerous ways: through country-based programmes which endorse policy improvement, building institutional capability and setting up concrete measures to end child labour and through mindfulness raising and mobilization intended to change social attitudes and promote ratification and effective implementation of ILO Child Labour Conventions. From projects funded in the period 1995–2010, a total of 985,698 children were withdrawn or prevented from getting up child labour through the providing of education and other opportunities (Fig. 12.3) (Office 2012).

Paired to this direct action throughout has been substantial in-depth statistical and qualitative research, child labour monitoring, programme evaluation and policy and legal analysis, which have led to the gathering of vast knowledge base of guidelines, statistical data and methodologies, good practices, thematic studies and training materials.

12.2.4 Organization for Security and Co-operation in Europe

OSCE participating States as source, transit or destination countries are virtually affected by human trafficking. This modern form of slavery is an injury to human self-esteem, often involving psychological dread and physical violence. The OSCE addresses many subjects relevant to human trafficking: economic, labour and migration policies; corruption and crime control; human rights and rule of law; discrimination and inequality (Organization for Security and Co-operation in Europe 2017).

In 2003, the Organization set up the Office of Special Representative and Coordinator for Combating Trafficking in Human Beings to support participating countries to develop and implement effective policies for combating human trafficking. The Office of the Special Representative endorses the protection of victims with a method centred and based on victim and human rights.

The OSCE Action Plan to Combat Trafficking in Human Beings offers the structure for OSCE activities supporting the anti-trafficking efforts of OSCE participating States. To date, the OSCE carries out field operation in South-Eastern Europe (Albania, Bosnia and Herzegovina, Kosovo, Montenegro, Serbia and Skopje), Eastern Europe (Moldova and Ukraine), South Caucasus (Yerevan) and Central Asia (Ashgabat, Bishkek, Astana, Dushanbe and Uzbekistan) (Organization for Security and Co-operation in Europe 2017).

12.2.5 National Centre for Missing and Exploited Children

The *National Center for Missing & Exploited Children*® (NCMEC) was established in 1984. It is the main nonprofit organization in the U.S. working with law enforcement, families and the specialists who help them on questions related to missing and sexually exploited children. This organization provides information and resources to minor victims of abuses, families and authorities as well as to other kinds of professionals.

Thirty years ago, police could add information about stolen cars, stolen guns, even stolen horses into the FBI's crime database—but not stolen children. Several tragic cases began to awaken the nation to the problem that there was no coordinated national system for addressing missing children cases. In 1979, 6-year-old Etan Patz disappeared from a New York street on his way to school. Over the next several years, 29 children and young adults were found murdered in Atlanta. Then in 1981, 6-year-old Adam Walsh was kidnapped from a Florida shopping mall and later found brutally assassinated. That is why NCMEC was created (Sedlak et al. 2002).

In the same year, the Congress of the United States approved the law of support of missing children and the *National Resource Center and Clearinghouse on Missing and Exploited Children* was established. Currently, there is a national free phoneline for missing children available 24 h 7 days a week. Only in 2017, the operation centre of NCMEC received more than 10,093 calls regarding possible child sex trafficking. To date, more than 4.5 million of calls have been received. Around

800,000 reports of missing children each year are received. Most of them are rapidly found. The NCMEC is involved in the most severe cases in those the minor is in a high risk. In 2017, NCME received 25,000 new cases reports of missing children and has analysed more than 12,900 attempted child kidnappings to identify trends and help develop safety tips for families (NCMEC 2015). The organization works with the distribution of photos and posters of children to catch public attention through its website, www.missingkids.com and a programme of spreading the photos that gather organizations and companies handling the distribution of pictures of the missing children and millions of people all around the world (NCMEC 2015; Sedlak et al. 2002).

12.2.6 Polaris: Freedom Happens Now

Polaris is a non-profit, non-governmental organization that works to combat and prevent modern-day slavery and human trafficking, founded in 2002. It is one of the largest anti-trafficking organizations in the United States. They operate the National Human Trafficking Resource Centre which serves as a central national hotline on human trafficking, launched in December 2007. The National Human Trafficking Hotline (NHTH) maintains one of the most extensive data sets on the question of human trafficking in the United States. Its mission is to provide human trafficking victims and survivors access to critical care and facilities to get help and stay safe and to prepare the anti-trafficking community with the tools to combat effectively all forms of human trafficking. In 2013, they launched a new hotline for texting "BeFree" where victims and captives could quickly and directly send out a message. Since 2007, 162,660 signals have been registered (143,029 calls, 9826 webforms and 9802 emails). In 2017, a total of 4460 cases of human trafficking were reported, most of them related to sex trafficking (71.43%) (Fig. 12.4) (Polaris n.d.).

12.2.7 DNA-Prokids Programme: DNA to Fight Crime on Children

Due to the importance of the children trafficking problem, DNA-PROKIDS (Programme for Kids Identification with DNA Systems) was started at the University of Granada based on the experiences with the Phoenix Programme (Lorente et al. 2000). The Programme is an international cooperation to help in identifying missing children, afford help to their relatives and to fight against human trafficking. This non-profit programme is supported by the Spanish Government, Andalusian Government and donations from public and private foundations and companies (BBVA, Banco SANTANDER, CajaGranada—BMN and Life Technologies).

The initiative DNA-PROKIDS was created in 2004 by Dr. José Antonio Lorente, director of the Genetic Identification Laboratory of University of Granada. It became a worldwide action after a pilot study from 2006 to 2008 in countries from Central America and Asia. In 2009, it started the international expansion of the project with the collaboration Prof. Arthur Eisenberg, from the Human Identification Center of

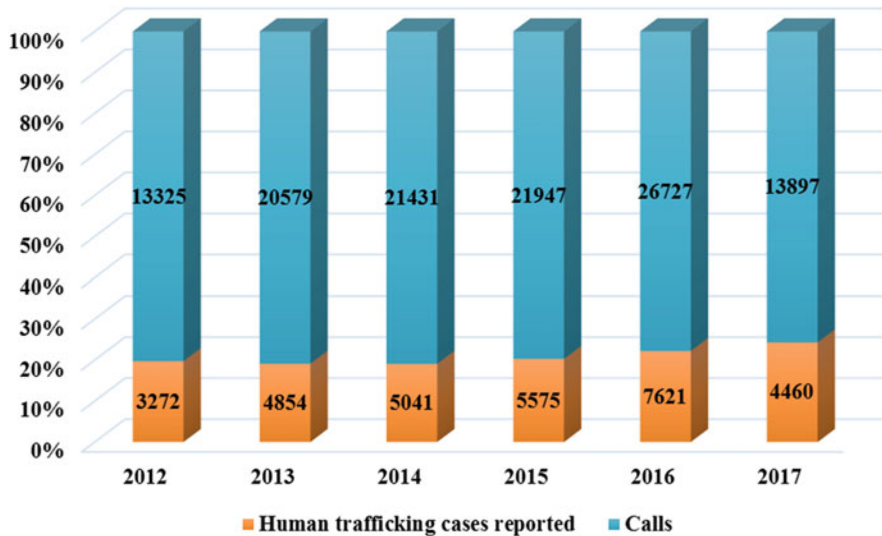


Fig. 12.4 Proportion of human trafficking cases reported in the total of calls received in the NHTH

University of North of Texas (UNT-HC) and with private financial support. Currently, the collaboration has extended to China, India, Indonesia, Philippines, Nepal, Sri Lanka, Thailand, Brazil, Guatemala and Mexico. In 2010, the "First International Congress of DNA-PROKIDS: Genetic identification against children trafficking", called together scientists, international organizations, NGOs and law enforcement agencies. The main aim was to establish a global alliance against human trafficking through the use of new technologies in genetic identification.

The main mission of DNA-PROKIDS is to identify the victims and return them to their families (reunification), to hamper traffic in human beings thanks to identification of victims and to gather information on the origins, the routes and the means of this crime and main elements for the work of law enforcement agencies.

DNA-PROKIDS is composed of three tiers. The first tier is at the national level with two genetic databases per country. One is for DNA profiles and non-genetic data obtained from children who, after proper investigation, are found in an illegal situation (e.g. not living with the natural family because these children were kidnapped or "adopted" illegally etc.). The other consists of DNA profiles and non-genetic data voluntarily provided by relatives (parents, siblings and other family members and whenever possible or "adoptive" mothers) or from personal items of reported missing children. The DNA profiles in these two databases are compared regularly to support in identifying missing children. DNA-PROKIDS first-tier pilot programmes are effectively in Guatemala and Mexico followed by Brazil and China. The power of identity testing and database searching will be most successful if DNA analysis is performed in every single case of a child being given to adoption. Moreover, when possible, the mother (or other available biological relatives) should also be tested to confirm her relationship to the child and their right to give up the

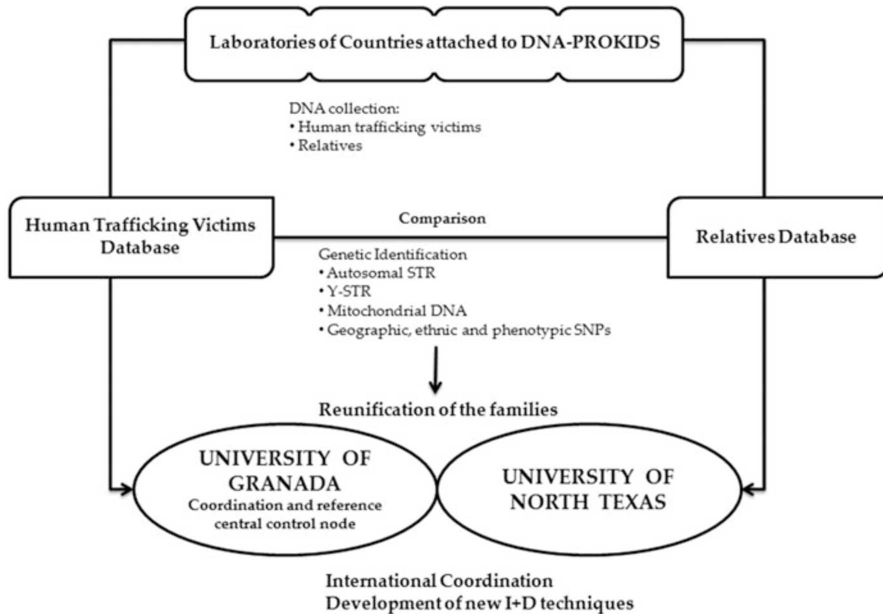


Fig. 12.5 Scheme to follow in the initiative DNA-PROKIDS

child. Employment of this act alone would decrease crimes related to children, where children who had been kidnapped, are sold for adoption.

The second tier is at international level. It will lead to the development of the infrastructure needed to share data among countries through several existing links and connections as far as through the creation of new, specific associations. Agreements from participating laboratories are required to identify an unique set of DNA markers for routine genotyping of all samples in relation to the programme; what non-genetic data that should be collected; the information that can be shared following confidentiality requirements that may differ per state and head off sufficient and sustainable financial support to establish such programmes in several countries.

The third tier is focused on data generation. There will be a universal automatic, compulsory inclusion (with signatory countries) in the database of (1) any reported child found out of his/her family (with illegitimate reasons); (2) children who are going to be adopted (previous any adoption it will be essential to confirm that the child has not been informed as missing anywhere in the world) and (3) the immediate inclusion of voluntary relatives of missing children (Fig. 12.5).

Many efforts are done to raise the database and facilitate communications between countries, like collaborations with INTERPOL in criminal investigations that will make criminals disinclined to commit these actions against children (exploitation, illegal adoptions and trafficking).

12.3 Applications and DNA Databases

12.3.1 New Tendencies in Human Identification

It is known that the most common way of human identification is focused on STRs profiles of different parts of the genome (nuclear, Y-chromosome, X-chromosome or mitochondrial one). All of them are largely validated and included in CODIS, as official loci/markers used for identification such as those of:

1. Applied Biosystems™ GlobalFiler™ PCR Amplification Kit (6-dye, 24-locus) and Promega Corporation Powerplex® Fusion 6C System (6-dye, 27-locus).
2. Applied Biosystems™ NGM Detect™ PCR Amplification Kit (6-dye and loci based on ENFSI and EDNAP requirements).
3. Applied Biosystems™ VeriFiler™ Express PCR Amplification Kit (6-dye, 25-marker multiplex assay with two gender markers, Penta D and Penta E markers).
4. There are also Y-chromosome specific kits, such as Applied Biosystems™ Yfiler™ Plus (27 Y-STR loci, 6-dye) and Promega Corporation PowerPlex® Y23 System (23 Y-STR loci, 5-dye).

Moreover, thanks to the new genomic era, related to an improvement of sequencing technologies and the reduction of the cost of sequencing, we are able to process more data in a less period of time. Currently, we are used to face with words meaning such as next-generation sequencing (NGS), whole exome sequencing or genome sequencing, with the large amount of data of all of these technologies. All of these events, suppose a big advance, but also have their limitations with problems in the analysis, presentation, format, sharing and reanalysis of genomic data (Poliakov et al. 2015).

Although several studies in NGS have been developed, it is still more developed in biomedicine field. But there are current systems with previous reports in forensic, such as MiSeq FGx System (Illumina) or the Applied Biosystems™ Precision ID NGS System (Thermo Fisher). Moreover, genomic era opens new targets of study in forensic such as the analysis of the epigenome or microbiome.

Microbiome is defined as the community of microorganisms (such as bacteria, fungi and viruses) that inhabit a particular environment and especially the collection of microorganisms living in or on the human body (Microbiome | Definition of Microbiome by Merriam-Webster n.d.). The Human Genome Project describes that we have about ten times as many microbial cells as human cells.

In forensic field, it is known that microbiome plays relevant roles such as: (1) geolocation and identification and (2) there are many corporal, internal and external sites, where microbiomes of a person can be sampled. Mainly the studies are focused on 16S rRNA cloning and comparing its diversity (Clarke et al. 2017). What makes it interesting for analysis in forensics, it is that a person's lifestyle can be a leading driver of the conformation of the microbiome communities across many body spots, including with whom they live, what they eat, where they are from,

whether they have pets, their health status, if they smoke and whom they kiss; all of these data result very informative in a forensic scenario (Clarke et al. 2017). Moreover, it is also crucial and proved that oral microbiome is heritable (Demmitt et al. 2017). So samples that serve as reservoirs for microbiome are diverse, such as human (skin, stool, hair, oral cavity); environment (soil, surfaces, air) and objects (cell phones, shoes, frames) (Clarke et al. 2017).

Concerning human trafficking, it is also of special interest that microbial communities differ in composition and function across different geographical locations, even with a deep precision of different cities. These variations are mainly due to the variations in climate, rainfall, altitude and soil as well as the metabolic properties of their host or energy sources available in the environment (Clarke et al. 2017).

Epigenome could be defined as a network of changes that modify the genome without altering the DNA sequences, having relevance in the activation of several genes or cells. One of the main molecules that play a relevant role in forensic field (as in biomedical field) are the microRNAs (miRNAs), short, non-coding RNAs of small size (19–24 nucleotides). They are stable molecules and specific for specific types of tissues and fluids, which reinforce their relevant role in forensic field. For example, it has been described that miRNA-16 and miRNA-451 are present in blood; semen with miRNA-135b and miRNA-10b; saliva with miRNA-205 and miRNA-658; vaginal secretion miRNA-124a and miRNA-372 and menstrual blood with miRNA-451 and miRNA-412. This opens a new tool with a relevant application in forensic scenes such as crimes, as a relevant way to identify the origin of the biological source (Dumache et al. 2017).

There are other regions, that although they are not new, they are highly relevant in population genetics and are often geographically oriented. The most common ones are haplogroups in Y chromosome (Y-DNA) and mitochondrial DNA (mtDNA). The Y-DNA haplogroups comprise many branches called subhaplogroups or subclades and form a phylogenetic tree of mankind. The branch lengths are delimited by the mutation rates of Y-STRs and Y-SNPs. The way used to name the different levels of the branch is defined by letters, A through T. Their subhaplogroups or subclades are expressed as letters and numbers (G2, R1b1, E3b1b, etc.). This tree is updated periodically according to new developments in the field. Currently, there are several softwares that help to classify Y-chromosome data properly such as Haplogroup Classifier, yHaplo; among others. One of the main interests of this gene in forensic field is the possibility of having an effective molecular tool to analyse the Y-STR haplotypes and know their haplogroups which allow us to discover the geographical origins of individuals (Mahal and Matsoukas 2017). All of this makes this chromosome as a crucial element of interest for human trafficking analysis.

The mtDNA haplogroups acts also as relevant markers for population genetics and it is also used the representation in clades as trees like it was explained in Y-DNA. Moreover, thanks to NGS technologies, we can improve the information of this genome, doing complete mtDNA sequencing for making a better reconstruction of difficult populations.

Similarly, as in Y-DNA there are several database and software applications that improve the use of mtDNA in forensic field such as EMPOP database (EMPOP n.d.; Parson and Dur 2007), PhyloTree software and Haplogrep2 (Arias et al. 2018).

12.3.2 Databases

Scientific and technical advances have generated and continue generating an extraordinary amount of information, which must be stored in a rational and orderly form for its later use. Databases are an organized way for data collection, which is a place where any type of information is stored. Nowadays, all databases are generated in electronica support (usually in computers or computerized) in an arranged and coherent manner, then it is automatically retrieved according to previously established parameters (Martin et al. 2001). Its access is quite regulated, depending on the importance of the stored data; those that contain information about the "privacy" of the people are the ones with more restrictions. Databases are always linked to the regulations of the Data Protection Act, 1998 (DPA) based on the principles of "good information handling". These principles show the specific rights in relation to personal information, including certain requirements on those organizations that process these data. These rules have been recently updated in Europe in the "2018 reform of EU data protection rules". Stronger rules on data protection have been established and more control over their personal data (Tikkinen-Piri et al. 2018).

In the scientific field, there are no doubts that one of the most important ways of knowledge is generated data from diverse investigations (Machado and Silva 2014). Entire and diverse fields of knowledge, such as epidemiology or biostatistics, have the base on the mathematical study of the data from different research, with the only aim of obtaining solid and objective results.

DNA analysis development allowed the establishment of genetic databases with a forensic interest. There are two types of forensic genetic databases: criminal and civil databases (Butler 2012).

- Civil DNA databases are focused on the identification of missing persons. Therefore, efforts are based on getting DNA from unidentified bones and compare them to relatives' genetic information. Biological samples from relatives are collected after getting informed consent permission and this is the major difference with criminal databases.
- Criminal DNA databases are set up to support criminal investigations. They facilitate the automatic DNA profiles comparison of different evidences, and they also allow making DNA comparisons from criminals versus forensic criminal evidences. Depending on the country, the establishment of these databases could require specific laws and all operations must be under judicial system control (Lorente et al. 1996).

At the moment there is still controversy about suspects "samples collection that in some cases are taken without their consent or without their knowledge". In some relevant investigations, some samples have been taken indirectly such as hairs, toothbrush, biological materials that remain in prison cells, other objects from the accused or indirectly from relatives. In the same way, try to obtain DNA from clinical samples that come from the suspect.. So, there should be limits to solve issues such as obtaining reference samples that may be controlled by ethical rules, not just under legal rules (Lorente et al. 1996). Obtaining the sample fraudulently could cast doubt on the entire database.

Therefore, there is no doubt about the need of accumulating data and no one can limit its free generation and their utility to society in general and to individuals in particular. However, it is essential that these actions are made in agreement with each country. Information in databases is especially sensitive and its use must be perfectly limited and restricted to law cases authorized. DNA databases have a great potential for criminal research and they prove their worth for cases with unknown offenders, serial offenders and/or co-offenders (De Moor et al. 2017).

There are some limitations in databases use. The practical benefit and usefulness of databases when fighting on human trafficking should not make us lose perspective of what they really are and their place in the criminal process. It will be equally taken into account legal and ethical problems that could be raised from the existence of excessively broad or even universal databases.

It would be necessary to consider that databases should be another instrument to help police investigation, never the only way. A complete police investigation must lead a suspect or suspects that criminalistics analyses will confirm or exclude.

At this moment, we believe that the uncontrolled databases spread (to all types of crimes or people), far from providing practical solutions; it would lead to a situation in which many people would be allegedly related to crimes. Several biological samples could be found as clues because they could be easily transported such as hair or cigarette butts.

It doesn't seem ethically correct to "make suspects" of large population groups, for example, all male inhabitants of a city or specific regions, just because a person has been sexually assaulted and semen has been found. This type of attitudes, far from being beneficial to forensic science, in our opinion, has become an obstacle and a brake on the progression and development. This situation makes that normal citizen feel more and more controlled and unprotected; the use of DNA databases in crime control cannot be allowed to compensate the lack of other actions, such as an adequate crime prevention policy and the reintegration of criminals. The "fear of being all registered" will not prevent criminal acts; however, education in this field will demonstrate that violence path only entails more violence and the lack of freedom.

In conclusion, the use of different databases combining their information, from different polices and DNA databases create new opportunities for networking analysis but always respecting the approved laws by all governments sharing their data.

References

Arias L, Barbieri C, Barreto G, Stoneking M, Pakendorf B (2018) High-resolution mitochondrial DNA analysis sheds light on human diversity, cultural interactions, and population mobility in northwestern Amazonia. Am J Phys Anthropol 165(2):238–255. https://doi.org/10.1002/ajpa.23345

Butler JM (2012) Advanced topics in forensic DNA typing: methodology. Elsevier Academic Press, San Diego

Clarke TH, Gomez A, Singh H, Nelson KE, Brinkac LM (2017) Integrating the microbiome as a resource in the forensics toolkit. Forensic Sci Int Genet 30:141–147. https://doi.org/10.1016/j.fsigen.2017.06.008

De Moor S, Vander Beken T, Van Daele S (2017) DNA databases as alternative data sources for criminological research. Eur J Crim Policy Res 23(2):175–192. https://doi.org/10.1007/s10610-016-9327-9

Demmitt BA, Corley RP, Huibregtse BM, Keller MC, Hewitt JK, McQueen MB et al (2017) Genetic influences on the human oral microbiome. BMC Genomics 18(1):659. https://doi.org/10.1186/s12864-017-4008-8

Dumache R, Rogobete AF, Sandesc D, Bedreag OH, Ciocan V, Muresan C et al (2017) Use of circulating and cellular miRNAs expression in forensic sciences. J Interdiscip Med 2(3):235. https://doi.org/10.1515/jim-2017-0074

EMPOP (n.d.). https://empop.online/. Accessed 8 June 2018

Human Trafficking/Involuntary Servitude—FBI (n.d.). https://www.fbi.gov/investigate/civil-rights/human-trafficking. Accessed 8 June 2018

Internation Labour Office (2017) Global estimates of child labour. Geneva. http://www.ilo.org/wcmsp5/groups/public/@dgreports/@dcomm/documents/publication/wcms_575499.pdf

International Organization for Migration (2017) Global trafficking trends in focus. Key trends from IOM victim of trafficking data. https://www.iom.int/sites/default/files/our_work/DMM/MAD/A4-Trafficking-External-Brief.pdf

Lorente M, Lorente JA, Budowle B, Villanueva E (1996) Science and conscience: regulation or guidelines for forensic hameogenetics? Springer, Berlin, pp 689–691. https://doi.org/10.1007/978-3-642-80029-0_214

Lorente JA, Entrala C, Alvarez JC, Lorente M, Villanueva E, Carrasco F, Budowle B (2000) Missing persons identification: genetics at work for society [2]. Science 290(5500):2257–2258

Machado H, Silva S (2014) “Would you accept having your DNA profile inserted in the National Forensic DNA database? Why?” Results of a questionnaire applied in Portugal. Forensic Sci Int Genet 8(1):132–136. https://doi.org/10.1016/j.fsigen.2013.08.014

Mahal DG, Matsoukas IG (2017) Y-STR Haplogroup diversity in the Jat population reveals several different ancient origins. Front Genet 8:121. https://doi.org/10.3389/fgene.2017.00121

Martin PD, Schmitter H, Schneider PM (2001) A brief history of the formation of DNA databases in forensic science within Europe. Forensic Sci Int 119(2):225–231. https://doi.org/10.1016/S0379-0738(00)00436-9

Más Allá de las Fronteras: Explorando los Enlaces entre Trata, Globalización y Seguridad (2010) Bangkok

Microbiome | Definition of Microbiome by Merriam-Webster (n.d.). https://www.merriam-webster.com/dictionary/microbiome. Accessed 8 June 2018

NCMEC (2015) National Center for Missing & Exploited Children annual report. http://www.missingkids.com/content/dam/ncmec/en_us/publications/ncmec2015.pdf

Office IL (2012) IPEC action against CHILD LABOUR 2010–2011: progress and future priorities. International Programme on the Elimination of Child Labour

Organization for Security and Co-operation in Europe (2017) 2016–2017 report on the OSCE special representative and co-ordinator for combating trafficking in human beings. Vienna. http://www.osce.org/secretariat/360796?download=true

Palmbach TM, Blom J, Hoynes E, Primorac D, Gaboury M (2014) Utilizing DNA analysis to combat the world wide plague of present day slavery—trafficking in persons. Croat Med J 55 (1):3–9. https://doi.org/10.3325/cmj.2014.55.3

Parson W, Dur A (2007) EMPOP-A forensic mtDNA database. Forensic Sci Int Gene 1:88–92

Polaris (n.d.). https://polarisproject.org/. Accessed 30 Jan 2018

Poliakov E, Cooper DN, Stepchenkova EI, Rogozin IB (2015) Genetics in genomic era. Genet Res Int 2015:364960. https://doi.org/10.1155/2015/364960

Saiz M, Alvarez-Cubero MJ, Alvarez JC, Lorente JA (2016) Forensic genetics against children trafficking: missing children genetic identification. In: Morewitz SJ, Colls CS (eds) Handbook of missing persons. Springer, Cham, pp 177–189. https://doi.org/10.1007/978-3-319-40199-7_13

Sedlak A, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention, NISMART (2002) National estimates of missing children: an overview. US Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention

Tikkinen-Piri C, Rohunen A, Markkula J (2018) EU general data protection regulation: changes and implications for personal data collecting companies. Comput Law Secur Rev 34(1):134–153. https://doi.org/10.1016/J.CLSR.2017.05.015

United Nations Office on Drugs and Crime (n.d.). https://www.unodc.org/. Accessed 30 Dec 2018

UNODC (2016) Global report on trafficking in persons 2016. Vienna. http://www.unodc.org/documents/data-and-analysis/glotip/2016_Global_Report_on_Trafficking_in_Persons.pdf

UNODD (2009) UN.GIFT—global initiative to fight human trafficking

What You Need to Know about Illegal Human Organ Trafficking | Big Think (n.d.). http://bigthink.com/philip-perry/what-you-need-to-know-about-human-organ-trafficking. Accessed 8 June 2018

Part III

Case Studies/Applications in Forensic DNA Typing

13 Autosomal STR Typing and Case Studies

Pankaj Shrivastava, Shivani Dixit, Varsha Srivastava,
Hirak Ranjan Dash, and R. K. Kumawat

Abstract

Autosomal STRs are the gold standard and are the most commonly used STR markers in forensic DNA typing. Most of the autosomal STRs present at the surrounding of the centromere of an autosomal chromosome are responsible for a unique DNA fingerprint in an individual. The updated information about the autosomal STR based kits is discussed along with the case studies to showcase the importance and utility of autosomal STR markers in forensic DNA typing. With the advancement in the technology, not only the number of STR markers has been increased but the inhibitor tolerance of the multiplex, reduced time of polymerase chain reaction (PCR), and the inclusion of one or more Y chromosomal markers have increased its utility in many folds.

Keywords

Autosomes · STRs · Multiplex PCR · Case study

The most common variant of DNA profiling today for criminal cases and other types of forensic uses is called autosomal "STR" (Short Tandem Repeat) analysis (Ellegren 2004; Britten and Kohne 1968). Human genome is made up of 23 pairs

P. Shrivastava (✉) · S. Dixit
DNA Fingerprinting Unit, State Forensic Science Laboratory, Department of Home (Police), Government of Madhya Pradesh, Sagar, Madhya Pradesh, India

V. Srivastava
Centre for DNA Fingerprinting and Diagnostics (CDFD), Hyderabad, India

H. R. Dash
DNA Fingerprinting Unit, Regional Forensic Science Laboratory, Bhopal, Madhya Pradesh, India

R. K. Kumawat
DNA Division, State Forensic Science Laboratory, Jaipur, Rajasthan, India

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_13

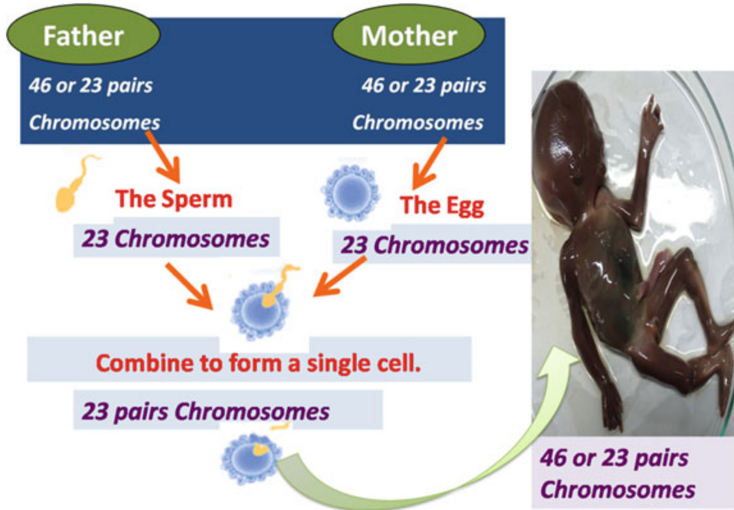


Fig. 13.1 Formation of a new body—contribution of parents to the child

of chromosomes, out of which 22 pairs are called autosomes, and last pair is sex chromosome (Broman et al. 1998; Ghebranious et al. 2003). Autosomes determine somatic characters of an individual and also contribute to the unique genetic makeup of an individual, which is inherited from parents. The pattern of this inheritance is peculiar. Every child gets 50% of autosomal genetic information from its biological father and rest 50% from its biological mother and can be genetically identified by comparing its autosomal genetic pattern with biological parents (Butler 2005) (Fig. 13.1). "Microsatellite DNA" or simple sequence repeats (SSRs) of most usually short tandem repeats (STRs) are 2–7 bp (base pair) length of a repeated sequence of DNA, mostly found at surrounding regions of the centromere of an autosomal chromosome (Kasai et al. 1990; Butler 2011). These autosomal STRs have a high degree of utility in forensic DNA applicability and admissibility in court of justice because of generating a unique DNA fingerprint of an individual (Biémont and Vieira 2006; Srivastava 2006; Butler 2007). A short tandem repeat (STR) in DNA occurs when a pattern of two or more nucleotides is repeated, and the repeated sequences are adjacent to each other (Butler 2007). These repeats are called tandem as the head to tail core sequence is the same, but the numbers of repeats are variable (Kasai et al. 1990). In a child, from a pair of STR alleles, one STR allele comes from the biological mother, and the other STR allele comes from the biological father at each of the genetic sites. Autosomal STR based DNA typing is nothing but a comparative evaluation of these STR profiles obtained from different individuals.

The STRs are popular DNA markers because these can be amplified easily in multiplex reactions by polymerase chain reaction (PCR) (Sambrook et al. 1989). These STRs exhibit the characteristics of single copy STR inheritance from each parent (Butler 2006). Furthermore, the highly polymorphic nature of STR repeats

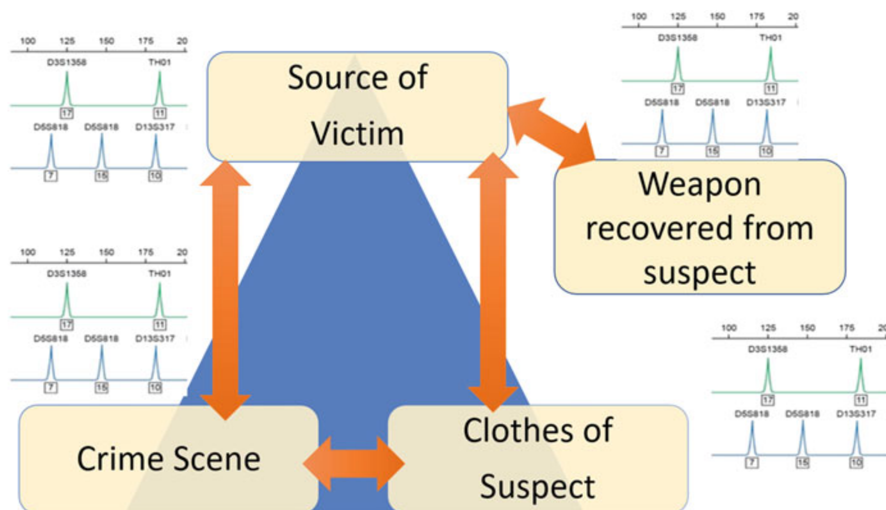


Fig. 13.2 DNA profile match linking suspect with the crime by autosomal STR

makes these markers useful for establishing identity of persons by DNA fingerprinting, which is also used as strong biological evidence in the court of law (Butler 2006, 2007). This technique is used to establish identification and provides connections between biological evidence and the suspect in a criminal investigation. The DNA profile generated from the biological samples received from the crime scene is compared with the DNA profile of the samples from suspect. If the two DNA profiles match, then it is concluded that the evidence came from that suspected person (Fig. 13.2). This technique is also used to identify bodies, track down blood relatives, and also for diagnosis of different diseases (Butler 2007).

DNA is the "blueprint" for the proteins in a human body as the sequences present in it code for a variety of proteins. DNA that actually codes for proteins cannot vary much; therefore close to 99.7% of our DNA is the same for all (Butler 2011). Usage of DNA for differentiation among people, therefore, is a tricky matter. We have a junk region in our genome, which does not code for any protein (Biémont and Vieira 2006). They are rich in tandem repeat regions called variable number tandem repeats (VNTRs) (Kasai et al. 1990). Useful differences and polymorphism in the DNA must be found in these remaining 1% of non-coding DNA regions, which is not known to code for anything specific. Because this part of DNA is non-coding, it is quite variable, which makes it an attractive tool for possible usage in distinguishing among individuals. These regions are the basis of STR analysis. A collection of these can give excellent evidence of a person's identity statistically because the likelihood of two unrelated people having the same number of repeated sequences in these regions is highly remote.

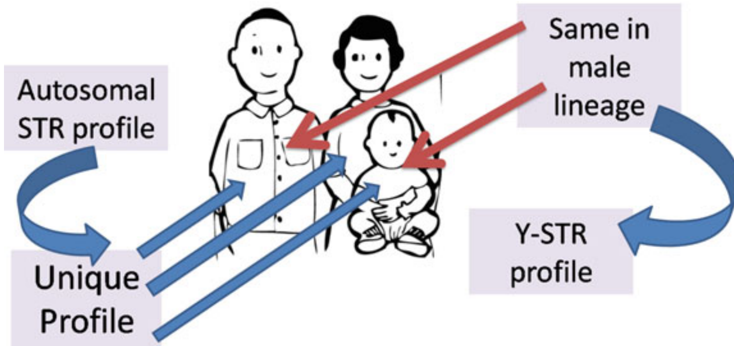


Fig. 13.3 Power and potential of autosomal STRs

The need for common nomenclature of STRs was felt considering the need for inter-laboratory reproducibility and for data comparison. Therefore designation or guidelines for nomenclature of STRs were framed by the International Society of Forensic Genetics (ISFG) (Bar et al. 1997). It was decided that

– The coding strand should be used for STRs within protein-coding genes (as well as in the intron of the genes). One such example is vWA (GenBank: M25716) (Bar et al. 1997).
– For repetitive sequences like many of the D#S## loci, the sequence originally described in the literature or the first public database entry shall become the standard reference for nomenclature (Bar et al. 1997).

It was also decided by the forum that if any nomenclature is already established and being used, but if does not follow the above guideline, the existing nomenclature should be used to avoid any confusion (Bar et al. 1997).

Autosomal chromosomes are not involved in gender determination, and STRs on these chromosomes are called autosomal STRs. Another type of STRs used for forensic purposes is called Y-STRs, which are present on the male sex-determining Y chromosome and X-STRs, which are present on the X chromosome. Profiles generated by autosomal STRs are statistically more relevant than profiles generated by Y-STRs and X-STRs. Since autosomal DNA undergoes recombination, chances of having the same profile by two individuals are more than the world population (Kayser 2017). The only exception is identical twins that develop from the same cell, so share the same genetic composition. Only males have a Y chromosome, and all males get these from their fathers, i.e., Y chromosomes are paternally inherited (Fig. 13.3). A Y-STR profile can offer substantial power of exclusion, but this type of profile certainly can't be used for establishing identity.

13.1 The Start of Multiplex Autosomal STR Typing

The 1990s was the start when the use of STR markers was first described as a useful tool for forensic human identification (Edwards et al. 1991, 1992). Just after this, the FSS (Forensic Science Services in the United Kingdom) started a massive drive for searching new loci and their population testing for the forensic purpose (Kimpton et al. 1993) followed by few other groups, including Europeans (Frégeau and Fourney 1993). The first multiplex, which was developed and tested for the forensic use by FSS, was a multiplex of four STR markers, namely TH01, VWA, FES/FPS, and F13A1, which is commonly known as "first-generation multiplex" (Kimpton et al. 1994). This was followed by second-generation multiplex (SGM) again from FSS, which included TH01, VWA, FGA, D8S1179, D18S51, and D21S11 loci along with amelogenin as a sex-determining marker. In 1994, Promega Co. launched its first STR kit, which was a three marker multiplex known as "CTT." This first commercially available STR multiplex included CSF1PO, TPOX, and TH01 and was based on silver stain analysis (Fig. 13.4).

13.2 The CODIS, ESS, and Expanded CODIS

In 1996, FBI Laboratory initiated the establishment of core STR loci to initiate a national DNA database. This was named as CODIS (Combined DNA Index System). Twenty-two working DNA laboratories participated in this initiative of FBI, and this exercise continued from April 1996 to November 1997. In this exercise, the 17 STR loci, namely CSF1PO, F13A01, F13B, FES/FPS, FGA, LPL,TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, andD21S11, were evaluated for inclusion in CODIS database establishment. Finally, after long interactions and collaborative exercises on the final STR Project, which was held on November 13–14, 1997, 13 core STR loci were adopted, which were declared as CODIS national DNA database (Olaisen et al. 1998). The 13 finally

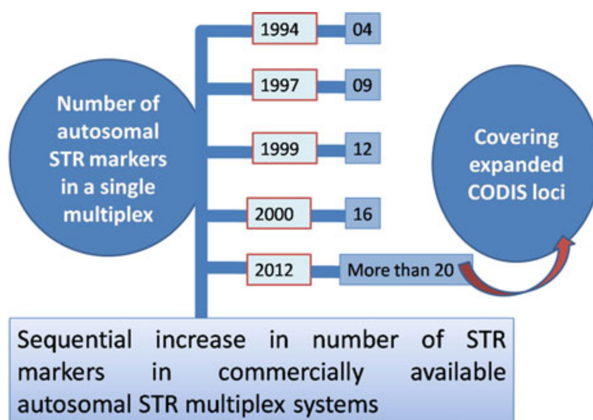


Fig. 13.4 Development in autosomal STR based multiplex kits

selected CODIS core loci were CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, and D21S11 (Butler 2011). Based on the population-based studies globally out of these 13 CODIS STR loci, FGA, D18S51, and D21S11 have been observed as the three most polymorphic markers. In contrast, TPOX, CSF1PO, and TH01 have been comparatively observed as low polymorphic markers. These three markers were the same, which were part of the first commercially available multiplex CTT, hence despite their low polymorphic character, these three made their place in CODIS 13 because, by that time, this CTT multiplex was extensively used in the United States.

13.3 The European Standard Set

In 1999, forensic DNA setups in European countries adopted the European Standard Set (ESS) of seven STR loci (FGA, TH01, vWA, D3S1358, D8S1179, D18S51, and D21S11) (Schneider 2009). Out of these seven markers, six markers were part of the SGM marker panel, and D3S1358 was additionally included. Five additional markers were accepted to be the part of ESS, namely D2S441, D10S1248, D22S1045, D1S1656, and D12S391 in April 2009, thus extending the total of 7ESS loci to 12.

In April 2011, the FBI Laboratory proposed an expanded set of core STR loci intending to avoid chances of adventitious matches for enhancing global compatibility and further to increase the power of discrimination for forensic purpose. The primary candidate loci for this expanded CODIS included the CODIS13 loci. Additionally, those loci were preferred, which were part of then commercially available multiplex kits in the United States (Butler 2006) and also which were in use globally for forensic DNA database studies (Schneider 2009). The loci with any published association with any linked medical condition or association with some decease were not taken into consideration (Hares 2015). The Working Group developed and published an implementation timeline for the public in June 2011 (Butler 2005). Based on the available three PCR amplification multiplex kits which included the candidate loci, namely GlobalFiler™ and GlobalFiler™ Express (Life Technologies now ThermoFisher Scientific) and PowerPlex® Fusion (Promega), 11 CODIS laboratories were notified as a consortium for conducting validation studies (as per FBI Quality Assurance Standards) for selection of expanded CODIS (Hares 2015). The compilation of the data generated by CODIS laboratories, along with its review and analysis, was executed with the help of the National Institute of Standards and Technology. And finally, as a result of this extensive exercise, newly expanded CODIS loci were selected, and laboratories were asked for its implementation by January 1, 2017 (Table 13.1).

Many commercial kits by different manufacturers like Promega, Applied Biosystem/Life Technologies (now ThermoFisher Scientific), Qiagen, and many others are available nowadays for autosomal STR analysis (Tables 13.2 and 13.3).

Table 13.1 The new 20 CODIS loci (Adopted since 2017)

CODIS 13 loci		Expanded loci
CSF1PO	D18S51	D1S1656
D3S1358	D21S11	D2S441
D5S818	FGA	D2S1338
D7S820	TH01	D10S1248
D8S1179	TPOX	D12S391
D13S317	vWA	D19S433
D16S539		D22S1045

These kits follow the multiplex process, which allows the co-amplification of incorporated loci in a single reaction of PCR. The reaction mix contains properly designed primers, fluorescent dyes along with Taq polymerase, dNTPs, buffer, etc.

DNA profiling today is PCR based and uses simple sequences or short tandem repeats (STR). These STR loci are targeted with sequence-specific primers and amplified using PCR. Subsequent fragment analysis is carried out to identify the allele sizes. The PCR amplicons are separated and detected using electrophoresis. Single-base differences also need to be differentiated; therefore, PCR products are resolved by automated DNA sequencing technologies with software that recognizes allele patterns by comparing them with a "ladder," which is nothing but a group of fragment sizes with known molecular weights. Repeats of a specific sequence at specific locations are identified in the genome, and a genetic profile of an individual is created.

A statistical analysis is generally made to determine the genotype frequency of the population. The Hardy–Weinberg Equilibrium (HWE) analysis provides the frequency of the observed genotype for each STR allele. The individual STR genotypes frequencies are multiplied together to assess the complete frequency profile, which gives the overall profile frequency. The generated DNA profiles from samples and suspects can be stored in a database which is used for comparative evaluation for the purpose of criminal investigation. Combined DNA Index System (CODIS), current DNA database maintained by the FBI contains case samples and samples from convicted individuals that are compared. As and when the new profile is entered, it is compared with these existing profiles in the database.

13.4 Case Studies

Autosomal STR based multiplex kits are used to decipher information in a wide variety of criminal cases (Table 13.4).

Case 1: A Suspected Case of Child Swapping in the Hospital

In a suspected case of child swapping, blood samples of father, mother, and girl child were received for forensic DNA examination. DNA was extracted by using an automated DNA extraction system EZ1 Advanced XL (Qiagen) using the automated DNA extraction kit and protocol of the manufacture. Obtained DNA was quantified

Table 13.2 Information about available autosomal STR kits

Multiplex kit	Make of the multiplex kit	Dye sets used	Launch year of the kit	Type of STR markers included in the kit					
				Autosomal STRs	Sex determination			Quality sensors	Total markers
					Amelogenin	Y Indel	Y-STRs		
PowerPlex® 16 System	Promega	4	2000	15	1	x	x	X	16
AmpFlSTR® Identifiler®	ThermoFisher Scientific	5	2001	15	1	x	x	X	16
AmpFlSTR® Profiler Plus™ ID	ThermoFisher Scientific	4	2001	9	1	x	x	X	10
AmpFlSTR® SEfiler™	ThermoFisher Scientific	5	2007	11	1	x	x	X	12
AmpFlSTR® Sinofiler™	ThermoFisher Scientific	5	2008	15	1	x	x	X	16
AmpFℓSTR™ Identifiler™ Direct	ThermoFisher Scientific	5	2009	15	1	x	x	X	16
PowerPlex® 16 HS	Promega	4	2009	15	1	x	x	X	16
PowerPlex® ESX 17 System	Promega	5	2009	16	1	x	x	X	17
AmpFlSTR® NGM™	ThermoFisher Scientific	5	2010	15	1	x	x	X	16
AmpflSTRIdentifiler Plus PCR multiplex kit	ThermoFisher Scientific	5	2010	15	1	x	x	X	16
AmpFlSTR® NGM SElect™	ThermoFisher Scientific	5	2010	16	1	x	x	X	17
Investigator Decaplex SE	Qiagen	5	2010	11	1	x	x	X	12
Investigator® IDplex Plus	Qiagen	5	2010	15	1	x	x	X	16
Investigator® ESSplex SE QS	Qiagen	5	2010	16	1	x	x	2	19
Global Filer PCR multiplex kit	ThermoFisher Scientific	6	2012	21	1	1	1	X	24

PowerPlex 21 multiplex system	Promega	5	2012	20	1	x	x	X	21
PowerPlex Fusion System	Promega	5	2012	22	1	x	1	X	24
Investigator 24plex Go and 24plex QS	Qiagen	6	2014	21	1	x	1	2	23
PowerPlex Fusion 6C System	Promega	6	2015	23	1	x	3	X	27
SureID® 23comp	Health Gene Technologies	5	2016	22	1	x	x	X	23
Sure ID PanGlobal Human DNA Identification Kit	Health Gene Technologies	6	2018	24	1	1	1	X	27
VeriFiler Plus PCR Amplification Kit	ThermoFisher Scientific	6	2018	23	1	1	x	2	25
GlobalFilerIQCPCR multiplex kit	ThermoFisher Scientific	6	2019	21	1	1	1	2	24
VersaPlex™ 27PY System	Promega	6	2019	23	1	x	3	**X**	27

Table 13.3 Comparative account of globally available STR based autosomal multiplex kits

Genetic markers	PowerPlex® 16	AmpFlSTR® Identifiler®	AmpFlSTR® Profiler Plus™ ID	AmpFlSTR® SEfiler™	AmpFlSTR® Sinofiler™	AmpFℓSTR™ Identifiler™ Direct	PowerPlex® 16 HS	PowerPlex® ESX 17 System	AmpFlSTR® NGM™	AmpflSTRIdentifiler Plus	AmpFlSTR® NGM SElect™	Investigator Decaplex SE	Investigator® IDplex Plus	Global Filer	PowerPlex 21	PowerPlex Fusion	Investigator 24plex Go and 24plex QS	SureID® 23comp	PowerPlex Fusion 6C	Sure ID PanGlobal	VeriFiler Plus	GlobalFiler IQC	VersaPlex™ 27PY
D3S1358	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
vWA	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
D16S539	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
CSF1PO	Y	Y	N	N	Y	Y	Y	Y	N	Y	N	N	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
D6S1043	N	N	N	N	Y	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	Y	Y	N	Y
D8S1179	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
D21S11	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
D18S51	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
D5S818	Y	Y	Y	N	Y	Y	Y	N	N	Y	N	N	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
D2S441	N	N	N	N	N	N	N	Y	Y	N	Y	N	N	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
D19S433	N	Y	N	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
FGA	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
D10S1248	N	N	N	N	N	N	N	N	Y	N	Y	N	N	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
D22S1045	N	N	N	N	N	N	N	Y	Y	N	Y	N	N	Y	N	Y	Y	N	Y	Y	Y	Y	Y
D1S1656	N	N	N	N	N	N	N	Y	Y	N	Y	N	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
D13S317	Y	Y	Y	N	Y	Y	Y	N	N	Y	N	N	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
D7S820	Y	Y	Y	N	Y	Y	Y	N	N	Y	N	N	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
PentaE	Y	N	N	N	N	N	Y	N	N	N	N	N	N	N	Y	Y	N	N	Y	Y	Y	N	Y
PentaD	Y	N	N	N	N	N	Y	N	N	N	N	N	N	N	Y	Y	N	N	Y	Y	Y	N	Y
TH01	Y	Y	N	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
D12S391	N	N	N	N	Y	N	N	Y	Y	N	Y	N	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
D2S1338	N	Y	N	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
TPOX	Y	Y	N	N	N	Y	Y	N	N	Y	N	N	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
SE33	N	N	N	Y	N	N	N	Y	N	N	Y	Y	N	Y	N	N	Y	N	Y	Y	N	N	N
D18S1364	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	N
D13S325	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	N
D5S2800	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	N
D9S1122	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	N
D4S2366	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	N
D3S1744	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	N
D11S2368	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	N
D21S2055	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	N
D20S482	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	N
D8S1132	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	N
D7S3048	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	N
D19S253	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	N
D17S1301	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	N
D22-GATA198B05	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	N
D6S474	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	N
D14S1434	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	N
D15S659	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	N
Amelogenin	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

DYS391	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	Y	Y	N	Y	Y	N	Y	Y
DYS576	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	Y
DYS570	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	Y
Y indel	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	N	Y	Y	Y	N
IQCS	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	Y	N
IQCL	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	Y	N
QS1	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	N	N
QS2	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	N	N

Table 13.4 Types of cases in which autosomal STR typing plays a vital role in forensic DNA examination

Types of cases	Questions which can be answered
Parentage disputes	• Biological relationship of the child with parents
	• Maternity of the child
	• Paternity of child
Identification of mutilated remains	• Identification of recovered biological material
Transfer of biological material	
1. Cases of sexual assault	• Transfer of male DNA to the source of the spot (clothing, condom, bed sheet, etc.)
	• Blood on males' genital in case of assault with minors
	• Transfer of female DNA on the clothing of suspect (transfer of blood in cases of assault with minors)
	• Transfer of female DNA (vaginal fluid) on the fingers of a male if case study dictates finger insertion into the female genital
	• Blood on any object used for insertion in cases of assault
2. Cases of murder	• Transfer of victim's DNA on the source of spot or accused
	• Source of any biological material in the crime scene or on any object found there, e.g., hair, cigarette stub, chewing gum
3. Cases of theft	Transfer of DNA of the suspect on the spot

on ABI 7500 Real-Time PCR (Thermo) *using Investigator Quantiplex Pro RGQ Kit* following the manufacturer's protocol. Five hundred picogram of DNA was amplified using PowerPlex$^{(R)}$ Fusion 6C multiplex system as per the recommendations of the manufacturer except reducing the reaction volume to 10 µl microliter. The amplified DNA fragments (amplicons) were separated on Genetic Analyzer ABI 3500XL using a 36 cm capillary array, POP™-4, and the obtained data were analyzed using GeneMapper$^{(R)}$ID-X software v1.5. The results in the performed DNA test showed the match of the girl child with the parents on all the tested markers included in the multiplex system (Table 13.5). The results in the case come up as false allegations on hospital administration, and the girl child was found to be the biological offspring of the tested couple.

Case 2: A Case of Paternity Exclusion

In a case of suspected paternity, blood samples of father and mother along with aborted fetus were received for forensic DNA examination. DNA was extracted by using an automated DNA extraction system AutoMateExpress™ (ThermoFisher Scientific) using the automated DNA extraction kit and protocol of the manufacturer. Obtained DNA was quantified on ABI 7500 Real-Time PCR (ThermoFisher Scientific) using PowerQuant$^{(R)}$ System (Promega) following the manufacturer's protocol. Five hundred picogram of DNA was amplified using PowerPlex$^{(R)}$ Fusion 6C multiplex system (Promega) as per the protocol of the manufacturer except reducing the total reaction volume to 10 µl (microliter). The amplified DNA fragments were

Table 13.5 Autosomal STR profile obtained in a suspected case of child swapping in one hospital

Genetic markers	Blood sample of father	Blood sample of child	Blood sample of mother
Amelogenin	XY	XX	XX
D3S1358	16, 17	15, 16	15
D1S1656	16, 19	13, 19	13,15
D2S441	11, 11.3	11, 12.3	11, 12.3
D10S1248	15	14, 15	14, 15
D13S317	10, 11	8, 11	8, 11
PentaE	8, 19	8, 16	14, 16
D16S539	11, 12	11, 13	11, 13
D18S51	14	14,17	16, 17
D2S1338	17, 23	17,18	18, 24
CSF1PO	10, 12	10	10, 13
PentaD	11, 12	11, 13	11, 13
THO1	6	6,9	6, 9
vWA	15, 18	17, 18	14, 17
D21S11	31.2, 33.2	29,31.2	29, 31.2
D7S820	8, 11	8, 8	8
D5S818	12	10, 12	10, 12
TPOX	10, 11	8,11	8
D8S1179	14, 15	10, 14	10
D12S391	23	19, 23	17, 19
D19S433	12, 15.2	12, 13	12, 13
SE33	16, 25.2	16, 22.2	17, 22.2
D22S1045	15, 16	14, 15	14, 15
FGA	22, 24	22, 23	19, 23
DYS391[a]	11	–	–
DYS576[a]	18	–	–
DYS570[a]	18	–	–

[a]Markers indicative of male DNA

separated on Genetic Analyzer ABI 3500XL using 36 cm capillary array, and POP™-4 and the obtained data were analyzed using GeneMapper(R) ID-X software v1.5. The results in the performed DNA test showed the match of the aborted fetus on all the tested markers with the mother while it showed inconsistencies with suspected father (Table 13.6). The result in the case concluded that the girl child was not the biological offspring of the tested couple.

Case 3: A Murder Case

In a murder case, blood-stained cotton, the weapon of offense (knife), clothes of two accused having bloodstains, a transparent tape seized from one of the accused, and clothes of deceased were seized by investigating officer. DNA was extracted by using an automated DNA extraction system AutoMateExpress™ (ThermoFisher Scientific) as per the protocol of the manufacturer. Obtained DNA was quantified

Table 13.6 Autosomal STR profile obtained in a case of paternity resulted in the exclusion

Genetic markers	Blood sample of father	Blood sample of child	Blood sample of mother
Amelogenin	X, Y	X	X
D3S1358	16, 17	18	18
D1S1656	11, 18	12	12, 16
D2S441	11, 14	11	11
D10S1248	13, 14	15, 16	13, 16
D13S317	11, 12	11	9, 11
Penta E	10, 16	12, 17	7, 17
D16S539	9, 12	8	9, 8
D18S51	12	11, 19	11, 19
D2S1338	20, 25	20, 24	20, 24
CSF1PO	10, 11	9, 12	9, 12
PentaD	10, 11	9, 11	9, 10
THO1	6, 9.3	7, 9	6, 9
vWA	15, 18	17	16, 17
D21S11	29, 31.2	28	28, 31.2
D7S820	10, 11	9	9
D5S818	12, 13	12	11, 12
TPOX	8, 11	8	8, 10
D8S1179	10, 11	14, 15	14, 15
D12S391	19, 20	17	17, 18
D19S433	14, 14.2	14, 15	14, 15
SE33	15, 31.2	15, 16	15, 16
D22S1045	15	17	16, 17
FGA	20, 21	20, 25	20, 25
DYS391[a]	10	–	–
DYS576[a]	15	–	–
DYS570[a]	18	–	–

[a]Markers indicative of male DNA

using ABI 7500 Real-Time PCR (ThermoFisher Scientific) using PowerQuant$^{(R)}$ System (Promega) according to the recommendations of the manufacturer. Five hundred picogram of DNA was amplified using ABI Global Filer™ multiplex kit (ThermoFisher Scientific) as per recommendations of the manufacturer. The amplified DNA fragments were separated on Genetic Analyzer ABI 3500XL using a 36 cm capillary array, POP™-4, and the obtained data were analyzed using GeneMapper$^{(R)}$ ID-X software v1.5 (ThermoFisher Scientific). The obtained results exhibited a single source consistent autosomal DNA profile from all the examined samples (Table 13.7). Thus the DNA profile proved the involvement of the accused into the crime.

Table 13.7 Autosomal STR profile obtained in a murder case in forensic DNA examination

Genetic markers	Article A cotton (spot)	Article C clothes (deceased)	Article D knife (accused 1)	Article F knife (accused 2)	Article G clothes (accused 1)	Article H clothes (accused 2)	Article I transparent tape (accused 1)
D3S1358	15,16	15,16	15,16	15,16	15,16	15,16	15,16
Vwa	14,18	14,18	14,18	14,18	14,18	14,18	14,18
D16S539	12,13	12,13	12,13	12,13	12,13	12,13	12,13
CSF1PO	10,12	10,12	10,12	10,12	10,12	10,12	10,12
TPOX	9	9	9	9	9	9	9
D8S1179	15	15	15	15	15	15	15
D21S11	29,32.2	29,32.2	29,32.2	29,32.2	29,32.2	29,32.2	29,32.2
D18S51	15	15	15	15	15	15	15
D2S441	11,12	11,12	11,12	11,12	11,12	11,12	11,12
D19S433	13,15	13,15	13,15	13,15	13,15	13,15	13,15
TH01	6,9	6,9	6,9	6,9	6,9	6,9	6,9
FGA	20,25	20,25	20,25	20,25	20,25	20,25	20,25
D22S1045	15	15	15	15	15	15	15
D5S818	11,12	11,12	11,12	11,12	11,12	11,12	11,12
D13S317	9,10	9,10	9,10	9,10	9,10	9,10	9,10
D7S820	11	11	11	11	11	11	11
SE33	18,32.2	18,32.2	18,32.2	18,32.2	18,32.2	18,32.2	18,32.2
D10S1248	14,16	14,16	14,16	14,16	14,16	14,16	14,16
D1S1656	13,15.2	13,15.2	13,15.2	13,15.2	13,15.2	13,15.2	13,15.2
D12S391	18,21	18,21	18,21	18,21	18,21	18,21	18,21
D2S1338	19,21	19,21	19,21	19,21	19,21	19,21	19,21
AMELOGENIN	XY	XY	XY	XY	XY	XY	XY

(continued)

Table 13.7 (continued)

Genetic markers	Article A cotton (spot)	Article C clothes (deceased)	Article D knife (accused 1)	Article F knife (accused 2)	Article G clothes (accused 1)	Article H clothes (accused 2)	Article I transparent tape (accused 1)
DYS391[a]	10	10	10	10	10	10	10
INDEL[a]	2	2	2	2	2	2	2

[a]Markers indicative of male DNA

13.5 Problems and Future Perspectives

PCR amplification becomes a challenging task with old samples which are not preserved properly. There are several factors which effect the success like the amount and quality of isolated DNA and the presence of PCR inhibitors. PCR amplification also hugely depends upon the procedure and methodology used. The following are some of the issues related to the amplification of autosomal STRs.

Analysis of degraded samples: The success of STR typing using autosomal STR based multiple kits depends upon various factors, including the amount of extracted DNA, the extent of DNA degradation, and purity of extracted DNA, i.e., the presence of inhibitors in extracted DNA (Opel et al. 2010; Kontanis and Reed 2006; Wilson 1997).

The multiplex kits being used for forensic purposes have short tandem repeats (STRs) length up to 500 bp. Forensic samples that are recovered from various adverse conditions are sometimes degraded as well. There has been the introduction of miniSTRs with reduced fragment size up to 200 bp (Mulero et al. 2008; Welch et al. 2011) with the expectation to enhance the efficacy of STR based identification. However, new generations multiplex autosomal STR kits which facilitate the usage of more loci with less than 200 base pairs size and with improved quality buffers also do not provide a satisfactory result for a wide variety of degraded samples.

Null alleles and shifted alleles: Another concern with autosomal STR kits of different commercial brands is the concordance in allele calls. Differences in primer design for the same loci incorporated in different STR based multiplex kits may result in null alleles or shifted allele calls (Westen et al. 2014).

Nucleotide variation but same allele call: The autosomal STR markers included in commercially available multiplex kits work on identifying the number of repeats. However, the variation in nucleotide sequences is not accessed by these multiplex systems (Oberacher et al. 2008).

DNA typing of touched objects: Though there have been many studies on touch DNA, still the theoretical and practical reality is different. Successful STR typing is reported on used cups, electrical switches, and used lipsticks (Webb et al. 2001; Sweet and Hildebrand 1999). Low DNA yield with touch DNA objects accounts for many STR artifacts. Some of the common issues are stutter artifacts, allele drop-in and drop-out, and peak imbalance, which makes it challenging to identify the real allele call from the sample.

Analysis of DNA typing of mixed samples: Mixed samples are another challenge for forensic DNA experts. The high content of one contributor frequently results in masking the lower contributor in the mixture. This is a common phenomenon which is observed while dealing with samples from sexual assault cases. There have been many studies on this since the inception of DNA technology in 1985 (Gill et al. 1985), but even after more than three decades, there is no standard approach to deal with mixed samples.

DNA profiling is based on comparison: Regardless of any genetic marker used for forensic DNA typing, the obtained results can be helpful only in the identification of individuals who are already known to the investigating officials (Kayser and De

Knijff 2011). This is another area which is being discussed globally for the last decade.

Ethical Statement All the samples analyzed and reported here were taken from routine casework examination performed by the author at DNA Fingerprinting Unit, State Forensic Science Laboratory, Sagar, MP, India, as per the standard guidelines for DNA fingerprinting examination in the country. The referral blood samples were collected with pre-informed written consent from the donors following the Declaration of Helsinki (Rickham 1964).

References

Bar W et al (1997) DNA recommendations further report of the DNA commission of the ISFH regarding the use of short tandem repeat systems. Int J Legal Med 110(4):175–176

Biémont C, Vieira C (2006) Junk DNA as an evolutionary force. Nature 443(7111):521–524

Britten RJ, Kohne DE (1968) Repeated sequences in DNA. Science 161(3841):529–540

Broman KW, Murray JC, Sheffield VC, White RL, Weber JL (1998) Comprehensive human genetic maps: individual and sex-specific variation in recombination. Am J Hum Genet 63 (3):861–869

Butler JM (2005) Forensic DNA typing: biology, technology, and genetics of STR markers. Elsevier, New York

Butler JM (2006) Genetics and genomics of core short tandem repeat loci used in human identity testing. J Forensic Sci 51(2):253–265

Butler JM (2007) Short tandem repeat typing technologies used in human identity testing. Biotechniques 43(4):Sii–SSv

Butler JM (2011) Advanced topics in forensic DNA typing: methodology. Academic Press, San Diego

Edwards A, Civitello A, Hammond HA, Caskey CT (1991) DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am J Hum Genet 49(4):746

Edwards AL, Hammond HA, Jin L, Caskey CT, Chakraborty R (1992) Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics 12 (2):241–253

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. Nat Rev Genet 5 (6):435–445

Frégeau CJ, Fourney RM (1993) DNA typing with fluorescently tagged short tandem repeats: a sensitive and accurate approach to human identification. BioTechniques 15(1):100–119

Ghebranious N, Vaske D, Yu A, Zhao C, Marth G, Weber JL (2003) STRP screening sets for the human genome at 5 cM density. BMC Genomics 4(1):6

Gill P, Jeffreys AJ, Werrett DJ (1985) Forensic application of DNA ‘fingerprints’. Nature 318 (6046):577

Hares DR (2015) Selection and implementation of expanded CODIS core loci in the United States. Forensic Sci Int Genet 17:33–34

Kasai K, Nakamura Y, White R (1990) Amplification of a variable number of tandem repeats (VNTR) locus (pMCT118) by the polymerase chain reaction (PCR) and its application to forensic science. J Forensic Sci 35(5):1196–1200

Kayser M (2017) Forensic use of Y-chromosome DNA: a general overview. Hum Genet 136 (5):621–635

Kayser M, De Knijff P (2011) Improving human forensics through advances in genetics, genomics and molecular biology. Nat Rev Genet 12(3):179

Kimpton CP, Gill P, Walton A, Urquhart A, Millican ES, Adams M (1993) Automated DNA profiling employing multiplex amplification of short tandem repeat loci. Genome Res 3 (1):13–22
Kimpton C et al (1994) Evaluation of an automated DNA profiling system employing multiplex amplification of four tetrameric STR loci. Int J Legal Med 106(6):302–311
Kontanis EJ, Reed FA (2006) Evaluation of real-time PCR amplification efficiencies to detect PCR inhibitors. J Forensic Sci 51(4):795–804
Mulero JJ et al (2008) Development and validation of the AmpFℓSTR® MiniFiler™ PCR amplification kit: a miniSTR multiplex for the analysis of degraded and/or PCR inhibited DNA J Forensic Sci 53(4):838–852
Oberacher H, Pitterl F, Huber G, Niederstätter H, Steinlechner M, Parson W (2008) Increased forensic efficiency of DNA fingerprints through simultaneous resolution of length and nucleotide variability by high-performance mass spectrometry. Hum Mutat 29(3):427–432
Olaisen B et al (1998) DNA recommendations 1997 of the International Society for Forensic Genetics. Vox Sang 74(1):61–63
Opel KL, Chung D, McCord BR (2010) A study of PCR inhibition mechanisms using real time PCR. J Forensic Sci 55(1):25–33
Rickham PP (1964) Human experimentation. Code of ethics of the world medical association. Declaration of Helsinki. Br Med J 2(5402):177
Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
Schneider PM (2009) Expansion of the European standard set of DNA database loci—the current situation. Profiles DNA 12:6–7
Srivastava V (2006) DNA fingerprinting in the criminal justice system: an overview. DNA Cell Biol 25(3):181–188
Sweet D, Hildebrand D (1999) Saliva from cheese bite yields DNA profile of burglar: a case report. Int J Legal Med 112(3):201–203
Webb LG, Egan SE, Turbett GR (2001) Recovery of DNA for forensic analysis from lip cosmetics. J Forensic Sci 46(6):1474–1479
Welch L et al (2011) A comparison of mini-STRs versus standard STRs-results of a collaborative European (EDNAP) exercise. Forensic Sci Int Genet 5(3):257–258
Westen AA et al (2014) Comparing six commercial autosomal STR kits in a large Dutch population sample. Forensic Sci Int Genet 10:55–63
Wilson IG (1997) Inhibition and facilitation of nucleic acid amplification. Appl Environ Microbiol 63(10):3741

Y Chromosome Short Tandem Repeats Typing

14

Rashed Alghafri

Abstract

The application of the Y chromosome short tandem repeats (Y-STR) in forensic field and lineage studies has been dramatically growing in the last few decades. In forensic cases, when male–female mixtures cannot be resolved using autosomal STR multiplexes kits, the application of Y-STR profiling will add a beneficial value by producing the male Y-STR haplotype. In forensic application, Y-STR is mostly used in sexual assault cases, paternity testing, familial searching, and male disaster victim identification (DVI). Assigning a weight to a haplotype match between evidence and reference samples is essential to make it admissible in court to help achieving justice. Because of the nature of the inheritance pattern of Y chromosome through male lineages, all loci located on Y chromosome are considered as one locus. Consequently, to assign the weight, a relevant Y-STR haplotype database should be employed to estimate the rarity of a Y-STR haplotype. This chapter explains the concept and history of Y-STR typing and its application for forensic investigation purposes. It explains all commercially available Y-STR multiplex kits such as PowerPlex® Y23 and Yfiler® Plus. Also, an illustration of in-house-developed multiplex kits for rapidly mutating Y-STRs is also presented. The discrimination capacity between male individuals, both relatives and non-relatives, and also haplotype diversity of these kits are illustrated briefly in this chapter. The interpretation guidelines of different types of loci, single-allelic and multi-allelic, are explained as per the international standard guidelines recommended by SWGDAM and ISFG. In addition, the two main Y-STR databases, YHRD and US Y-STR, are discussed along with

R. Alghafri (✉)
General Department of Forensic Sciences and Criminology, Dubai Police G.H.Q, Dubai, United Arab Emirates

Department of Biology, United Arab Emirates University, Al Ain, United Arab Emirates

Faculty of Health Science and Medicine, Bond University, Gold Coast, QLD, Australia

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_14

other databases which are used for genealogy studies. Moreover, the utilization of these databases to estimate the rarity of Y-STR haplotypes and its involved statistical uncertainties following different approaches has been discussed.

Keywords

Y-STR · Forensic · Haplotypes · DNA · Typing

14.1 Introduction

14.1.1 The Y-Chromosome

Y chromosome is considered the smallest chromosome out of the 23 pairs in human genome, which span of 60 Mb in size (Buhler 1980). The majority of the Y chromosome comprises of a region called non-recombining region of the Y chromosome (NRY). At the end of each chromatids of this chromosome, there is a region called pseudo-autosomal regions (PARs) which represent 5% of the total size of the Y chromosome (Quintana-Murci and Fellous 2001). This region comprises sequences of DNA which are similar to DNA sequences on each end of X chromosome's chromatids, and therefore, it is the only segment of Y chromosome where genetic material can be exchanged with the X chromosome during meiotic recombination (Quintana-Murci and Fellous 2001). The heterochromatin which is the tightly packed DNA and euchromatin which is the loosely packed DNA with a size of 30 Mb and 24 Mb, respectively, are making up the whole NRY. This region occurs only in a haploid state. Throughout the paternal heredities, the NRY is inherited intact and, therefore, remains untouched except when mutations are evident (Quintana-Murci et al. 2001). According to its paternal inheritance pattern, NRY segment of Y chromosomes has become extremely useful for population genetic structure researches, indeterminate paternity testing, disaster victim identification (DVI) with males individuals involvement, and anthropology studies in which identification of male lineages is needed and most importantly is for identification of male criminals in sexual assault cases (Kayser et al. 1997; Underhill et al. 2000; Hammer et al. 2001; Oota et al. 2001; Jobling and Tyler-Smith 2003; Roewer et al. 2005; Shi et al. 2010; Kayser 2017).

14.1.1.1 Y Chromosome Loci Mutations Rates

Likenesses among NRY segment between males individuals in the population represent shared ancestry among these male individuals, whereas differences among NRY segment between male individuals in the population reflect the fact that these individuals do not share close paternal lineages (Ballantyne et al. 2014). Nevertheless, to reach such conclusions, the analyst should take into consideration the mutation rate of these markers throughout the interpretation. Henceforth, mutation rates of markers in NRY plays crucial role in inclusion and exclusion of male individuals in male lineage and forensic investigations (Fan and Chu 2007).

For STR loci, which are spread all over the genome, there are two main mutation mechanisms. The first mechanism, which is presented during replication of the DNA, is the slippage of the DNA polymerase enzyme incident and incorrect reattachment to DNA strands. Whereas, the second mechanism takes place through the crossover of chromosomes during recombination which, whenever it is asymmetric, is mainly due to the configuration of the repetitive loci structure (Schlotterer and Tautz 1992; Shriver et al. 1993). In comparison, repetitive DNA sequences are more susceptible to show DNA polymerase enzyme dissociation than other sequences across the genome (Kuhner et al. 2007). The nature of repetitive structure allows increasing the probability of misaligning incidents to take place in the detached strands of DNA during the hybridization process. This process will lead to the production of secondary structures such as loop-, hairpin-, or triplex-shaped DNA strands. Consequently, an additional repeat will be added or a repeat will be skipped by the enzyme, depending on the reattachment position (Shriver et al. 1993; Pearson and Sinden 1998). Such errors can be identified and repaired during the DNA replication process using two pathways: first, by proofreading capability of exonucleases which is involved in the nascent strand synthesis, and second, it can also be achieved by the DNA mismatch repair system. Previously, it has been shown by different studies that proofreading exonucleases process is inefficient in identifying addition or loss of repeats; therefore, in STR region of DNA, there is only mismatch system left to correct any error incidents (Schlotterer and Tautz 1992). The main reason behind deficiency of proofreading exonuclease in repetitive DNA sequences is the existence of repeat dislocation of the mismatch by the 3′ terminus in the replicated DNA strand and, therefore, cannot be identified (Kroutil et al. 1996; Sia et al. 1997). DNA mismatch repair system involves a number of proteins which are critical to start the repairing process of DNA. These proteins are MutSα, MutLα, Exo1, RPA, PARP-1, PCNA, RFC, and DNA polymerase δ (Harfe and Jinks-Robertson 2000). In the mismatch repairing processes, there are two pathways, first pathway is a straightforward process directed by the break in 5′ strand. The process starts with an activation of exonuclease protein (Exo1) with an aid of MutSα and MutLα proteins. After that, RPA and PARP-1 proteins will attach to both the broken and the stabilized DNA strands, respectively, through which the PARP-1 will detect the broken strand. Next, the strand displacement process will begin with the aid of RPA, PCNA RFC, and DNA polymerase δ (Harfe and Jinks-Robertson 2000). The second pathway is complex since it starts the process of DNA dislocation after endonuclease activation in which it skips the hydrolysis step of the broken strand. During such mismatch repair pathway, the process can be started at either 5′ or 3′ strand breaks (Harfe and Jinks-Robertson 2000; Li 2008). Due to the fact that DNA mismatch repair system is too complicated as well as the fact of having only this mechanism to fix such errors within repetitive regions of DNA sequence, it is expected that this region of DNA sequence will experience a high rate of mutations compared to other sequences. This also leads to the conclusion that any defects that occur within DNA mismatch system will result in increasing the mutation rate above normal (Tiraby and Fox 1973).

Previous studies have estimated the mutation rates of STR loci to be between 10^{-5} and 10^{-2} per loci per generation (Edwards et al. 1992; Weber and Wong 1993; Brinkmann et al. 1998). Nowadays, STR loci which are mostly used in the commercial kits for forensic DNA typing are having mutation rates between 1×10^{-3} and 6.4×10^{-3}. These estimations were done by analyzing a large number of proven family members of DNA samples (Short Tandem Repeats Internet Database [Online] 2014). STR locus mutation rates are affected by different aspects, one good example is the germline differences. Due to the high number of divisions which sperm cells undergo compared to oocytes, it has been concluded that mutation rates are usually higher in male germline compared to female germline (Weber and Wong 1993; Brinkmann et al. 1998). Another aspect is the length of the repeat size in simple STR loci, where studies have shown that dinucleotide and trinucleotide repeat STR loci have higher mutation rates than tetranucleotide and pentanucleotides repeat STR loci and eventually will present higher diversity in allele sizes (Kruglyak et al. 1998; Sup Lee et al. 1999). On the other hand, in complex STR loci studies, it was determined that the factor with the most impact on mutation rates is the longest strand of homogeneous repeat number (Klintschar and Wiegand 2003), despite the following studies which have determined that the total number of repeats found within complex STR loci is more impactful on mutation rates compared to longest strand of homogeneous repeat number (Ballantyne et al. 2010). Additionally, it has also been realized that the mutation rate is increasing along with the increasing number of repeats within the same STR locus, this correlation was concluded in a study conducted using model organisms by utilizing publicly available data of human and nonhuman STR locus mutation rates extracted from previous publications. Notably, such correlation was also detected in nonhuman STR loci (Seyfert et al. 2008).

14.1.1.2 Y-STR Applications

Male Lineage and Population Genetics

Based on the rate of mutations, the genetic loci located on the non-recombinant region of Y chromosome (NRY) can be classified into two groups. The first group constitutes the loci with immensely low mutation rates such as single-nucleotide polymorphisms (SNPs) (Underhill et al. 2000; Van Oven et al. 2014). The mutation rate of Y-SNPs have been approximated to be 3×10^{-8} per nucleotide per generation (Xue et al. 2009; Poznik et al. 2013; Kayser 2017). Therefore, they are well known to be the most beneficial markers in detecting and correlating male lineages haplogroups (Underhill et al. 2000; Van Oven et al. 2014). The following group includes the loci with high mutation rates such as Y-STRs, with an average mutation rate of 10^{-3} per locus per generation (Goedbloed et al. 2009; Ballantyne et al. 2010; Kayser 2017). These loci have been determined to be effective in the analysis of paternal kinship patterns as well as in verifying the likelihood of the sample's origin (Kayser et al. 2003, 2005; Roewer et al. 2005; Coble et al. 2009).

Applications in Forensic Investigations

Y-STRs have been testified to be nearly as discriminating as autosomal STRs (Roewer et al. 1992). When autosomal STRs exhibit inadequacy of information in forensic investigations concerning male individuals, Y-STRs can aid in elevating the significance of DNA profiling. Y-STRs are mainly utilized for sexual assault cases, where the swabs would often comprise a mixture of both the victim and the offender's DNA. Despite the justification of differential extraction being effective in separating the male sperms from the female epithelial cells (Vuichard et al. 2011), the DNA of the female may yet exist in a minority of eluted male fractions. Thence, the result attained by utilizing autosomal STRs are arduous in obtaining or interpreting as mixtures. In situations where post-coital interval is lengthened, the efficacy of the differential extraction procedure was exhibited to decrease and was not beneficial post 48 h (Hall and Ballantyne 2003). However, preceding studies conducted revealed that the sperm survival time in the post-coital is extended up to 3 days followed by up to 7 days in the cervix (Willott and Allard 1982). Inflammation of the vagina, extreme douching, and administration of spermicides can lead to the loss of spermatozoa (Sibille et al. 2002). Male-specific markers including the Y-STR assays are therefore used as they can efficiently amplify the male DNA in a sample without amplifying the female DNA (Jobling et al. 1997; Gusmão et al. 2006).

In addition, there are various other applications of Y-STR kits in forensic DNA typing. This includes paternity testing for a male child and kinship analysis for male relatives. Autosomal STRs are commonly utilized for paternity testing. However, when the putative father is not present for comparison, Y-STR analysis is advantageous to a greater extent. The closest male relative of the putative father can be examined considering that the Y-STR haplotype is unchanged and shared among close male relatives (Kayser and Sajantila 2001; Stumpf and Goldstein 2001). As a result of the lack of recombination throughout the NRY region, the individual Y-STR loci are used collectively. The profile obtained is known as a haplotype. Due to the linkage, all the loci on the Y chromosome act as one locus (Gusmão et al. 2006).

14.2 Y-STR Multiplex Assays

Y-STRs were introduced to forensic practice in the mid-1990s. It has been ratified for both forensic applications and kinship analysis (Kayser et al. 1997; Schneider et al. 1998; Kayser 2017). The Europe core set of Y-STR loci known as minimum haplotype consisted of seven loci as listed in Table 14.1 (Kayser et al. 1997). The number of loci was further extended by the addition of two loci, namely DYS438 and DYS439 by the Scientific Working Group on DNA Analysis Methods (SWGDAM). The minimal and the extended haplotype was attainable by using the Y-Plex-5™ (DYS389I/II, DYS392, DYS438, DYS439) and Y-Plex 6™ (DYS19, DYS390, DYS391, DYS393, DYS389II, DYS385) multiplex Y-STR kits (Pascali et al. 1998).

Table 14.1 Currently available commercially and in-house available Y-STR assays

Minimal haplotype	Powerplex Y®	Yfiler®	PowerPlex Y23®	Yfiler Plus®	RM Y-STR
DYS19	DYS19	DYS19	DYS19	DYS19	
DYS389I	DYS389I	DYS389I	DYS389I	DYS389I	
DYS389II	DYS389II	DYS389II	DYS389II	DYS389II	
DYS390	DYS390	DYS390	DYS390	DYS390	
DYS391	DYS391	DYS391	DYS391	DYS391	
DYS392	DYS392	DYS392	DYS392	DYS392	
DYS393	DYS393	DYS393	DYS393	DYS393	
DYS385ab	DYS385ab	DYS385ab	DYS385ab	DYS385ab	
	DYS437	DYS437	DYS437	DYS437	
	DYS438	DYS438	DYS438	DYS438	
	DYS439	DYS439	DYS439	DYS439	
		DYS448	DYS448	DYS448	
		DYS456	DYS456	DYS456	
		DYS458	DYS458	DYS458	
		DYS635	DYS635	DYS635	
		YGATAH4	YGATAH4	YGATAH4	
			DYS481	DYS481	
			DYS533	DY533	
			DYS549		
			DYS570	DYS570	DYS570
			DYS576	DYS576	DYS576
			DYS643		
				DYS460	
				DYS627	DYS627
				DYS518	DYS518
				DYS449	DYS449
				DYF387S1	DYF387S1
					DYF399S1
					DYF403S1ab
					DYF404S1
					DYS612
					DYS626
					DYS526ab
					DYS547

An extensive study conducted in the year 2004, with the purpose of identifying novel Y-STR loci with the potentiality of forensic application, resulted in acquiring 166 novel Y-STR markers; totaling the markers to 219 (Kayser et al. 2004). Subsequently, this led to the development of several commercial Y-STR multiplex kits. The frequently used kits which have been approved for casework investigations include: AmpFℓSTR Yfiler® PCR Amplification Kit (Applied Biosystems™) and the Powerplex Y® Kit (Promega™), which amplifies 17 Y-STR and 12 Y-STR loci,

accordingly. Besides, in the year 2012, PowerPlex Y23® (Promega™), which amplifies 23 Y-STR loci simultaneously, was developed and validated (Thompson et al. 2013).

Moreover, the discovery of rapidly mutating Y-STRs (RM Y-STRs) has undeniably elevated the power of discrimination of Y-STRs and are therefore included in the newly released Yfiler Plus® Kit. This kit has the capability of amplifying 26 Y-STR loci simultaneously as shown in Table 14.1 (Ballantyne et al. 2014).

14.2.1 Limitations of Y-STR Kits

Although the Y-STR male specificity is certainly the utmost advantage in forensic DNA typing, it has its own drawbacks. The low effective population size of Y chromosomes means an increased susceptibility to genetic drift, inducing extreme variation of allele frequencies as a result of such random events. Such events can entirely eradicate a specific male lineage, lessening the population's genetic diversity (Charlesworth 2009).

According to previously conducted studies, Y-STRs have proven to be relatively as discriminating as autosomal STRs. Nonetheless, the Y chromosomes lack recombination, resulting in high number of shared haplotypes between male individuals. Concurrently, it was justified that autosomal and Y chromosome STRs possess an identical range of mutation rates (Roewer et al. 1992). Although presently utilized Y-STRs are capable of distinguishing between various male lineages, they are yet unable to resolve these lineages down to the individual level when determining paternal relatives (Gusmaúo et al. 2003).

Consequently, the usage of the current Y-STR kits in forensic DNA profiling will portray identical Y-STR haplotype in most of the cases of paternally related male subjects (Mulero et al. 2006; Thompson et al. 2012; Kayser 2017). Additionally, during extensive population substructure, identical haplotypes may also be found among unrelated male entities from the same community, stemming as a result of non-random mating in that subpopulation. Ultimately, this will cause haplotype frequencies to differ from estimations depending on the sum of the population. Generally, the contribution of few males in finding a subpopulation would lead to some haplotypes to be more frequent in the subpopulation in contrast to the wider population. Thereupon, highly sub-structured populations necessitate their own database to achieve the true value of a match when individuals from such populations are accused (Lewontin and Hartl 1991; Charlesworth 2009).

Since majority of the populations display differing levels of substructure, several unrelated males may share identical Y-STR haplotypes (Gusmaúo et al. 2003; Hedman et al. 2004). Hence, Y-STR evidential value is incredibly weak due to loci linkage and sub-structuring in comparison to autosomal STR loci. The evidential value of a matching Y-STR haplotype is often articulated as haplotype frequency. The Y-STR haplotype discrimination capacity (DC) can also depict the evidential power of Y-STRs in the applicable male population. The approximation of the number of unique haplotypes detected in the samples of the population is

known as discrimination capacity (Kayser et al. 1997). The current Y-STR multiplex kits have been improved by increasing the number of loci amplified in order to raise the evidential value with victory (Purps et al. 2014). Certainly, it is necessary to further enhance the Y-STR system discrimination capacity toward male individualization.

14.2.2 Enhancing Y-STR Discrimination Capacity

Providing the facts mentioned above pertaining to Y-STRs, it is coherent that there are two methods to increase the Y-STR haplotype discrimination capacity. The first method is to increase the number of loci examined in the analysis which will sequentially increase the likelihood of observing loci with different alleles, therefore aiding in the identification of male individuals. The differences in alleles, driven by mutation events, will lead to segregating a lineage into two distinct haplotypes. This approach was attempted prior to the release of Yfiler® amplification kit. A total of 20 novel single-copy Y-STR loci were analyzed by utilizing a single multiplex reaction, on 74 US population male samples. This resulted in an increase in haplotype resolution in contrast to the six Y-STR loci available in Y-Plex™ 6 kit (Butler et al. 2002). In addition, a comparative study was conducted which focused on 37 Y-STR loci, including 17 loci which are available in Yfiler® amplification kit. A total of 656 US population male samples were studied. Three of the samples had identical haplotypes which could be discriminated by the addition of a Y-STR locus: DYS522 or DYS576. Out of the additional 20 Y-STR loci analyzed, only seven loci showed a high level of differentiation which was equal to the whole set of the 20 Y-STR loci, which reflects the high resolution and discrimination imposed by these markers (Decker et al. 2007).

Another study was conducted where the authors analyzed a total of 590 male samples from 51 different populations globally, belonging to eight different regions. Novel 49 simple single-copy Y-STRs analogous to the 17 known Y-STRs present in the Yfiler® amplification kit were used. The results revealed that greater than 95% of male individuals displayed unique haplotypes in comparison to the Yfiler® kit where 90% of male individuals showed unique haplotypes. The loci with the maximum mutation rates, DYS570 and DYS576, immensely contributed in differentiating the male individuals. Although the highest number of Y-STR loci was 67 in this study, there was still 20 Y-STRs haplotypes shared across 47 male individuals, knowing that these male individuals do not share close paternal biological relationship (Vermeulen et al. 2009).

The second method in striving for improving the discrimination capacity of Y-STRs is the usage of a panel comprising of Y-STR loci with higher mutation rates than the loci that are commonly included in multiplexes such as Yfiler® or Powerplex®. In the year 2004, a study was carried out to assess the suitability of particular Y-STR loci for forensic applications. However, the mutation rates of these loci were unidentified (Kayser et al. 2004). A following in-depth study was then conducted focusing on 186 loci. The mutation rates of these loci were estimated in

2000 father–son pairs from Poland and Germany (Ballantyne et al. 2010). The results indicated that 13 loci had mutation rates of 1×10^{-2} and higher. While the remaining Y-STR loci displayed a mutation rate of 1×10^{-3} and lower. This included 12 tetranucleotide loci and 1 trinucleotide locus. Out of the 13 Y-STRs, 4 were multi-copy markers and 9 were single copy markers. Thence, a total of 21 alleles can generally be amplified by using these 13 RM Y-STR markers, termed as rapidly mutating Y-STRs (RM Y-STRs) (Ballantyne et al. 2010) (Table 14.2).

14.2.3 Rapidly Mutating Y-STRs (RM Y-STRs)

Several studies conducted prior to the discovery of the 13 RM Y-STRs in 2010 had already included one or more single-copy RM Y-STRs in an in-house developed multiplex kits (Redd et al. 2002; Hanson and Ballantyne 2007a, b; Rodig et al. 2008; D'Amato et al. 2009, 2010; Vermeulen et al. 2009; Hedman et al. 2011). However, multi-copy loci comprising DYF387S1, DYF399S1, DYF403S1, and DYF404S1 have been examined comparatively lesser than the single-copy loci. This is supposedly due to the complications involved in the genotyping and sequencing of these loci. Multi-copy loci interpretations are usually complex for two of the following reasons. First, the imbalance in the amplified peaks can alter genotyping correctly. Second, when more than one male has contributed to the sample, mixture of samples, imbalance of the peak height makes assigning specific alleles to contributors troublesome. Therefore, the usage of multi-copy Y-STR loci have been eluded in forensic DNA typing multiplex kits. Nonetheless, multi-copy Y-STRs have high discrimination capacity since each copy has an independent mutation probability, thence, increasing the mutation probability in accordance with the number of copies present in the locus.

In the year 2012, Ballantyne et al. engaged three multiplex panels amplifying 13 RM Y-STRs to examine 604 unrelated male individuals from 51 different population samples (HGDP-CEPH). Thirteen RM Y-STR displayed greater haplotype diversity (HD) and power of discrimination (PD) in comparison to the 17 Y-STRS in Yfiler® multiplex kit (Applied Biosystems) (Ballantyne et al. 2012). This set of loci was exceedingly beneficial in increasing the differentiation between close and distant male relatives by 4.4-fold of average male relative differentiation compared to the Yfiler® kit (Ballantyne et al. 2012).

Moreover, a comprehensive study conducted using 12,272 male individuals representing main populations around the world showed that 99% of them were fully discriminated using RM Y-STRs. In the substantial number of samples, the haplotype diversity was found to be exceedingly high. Amidst 111 populations globally, only six haplotypes were shared between populations within the same geographical region (Ballantyne et al. 2014). Even though the significance of RM Y-STRs has already been illustrated, two in-house developed multiplex kit which are thus far not commercialized have the potential to multiplex all the 13 RM Y-STRs (Alghafri et al. 2013, 2015; Zhang et al. 2017). These multiplexes were affirmed to be beneficial in forensic casework studies accompanied with other human

Table 14.2 Details of the 13 RM Y-STRs

Locus	Repeat type	Repeat motif (variants in bold type)	Allele ranges		Mutation rate
			Repeat number	Base pairs	
DYF387S1	Tetra, complex	$(AAAG)_3(GTAG)_1(GAAG)_4N_{16}(GAAG)_9\ \mathbf{(AAAG)_{13}}$	28–38	241–281	1.58×10^{-2}
DYF399S1	Tetra, complex	$(GAAA)_{\mathbf{3}}\mathbf{N_{7-8}(GAAA)}_{10-23}$	10–23	261–313	7.73×10^{-2}
DYF403S1a	Tetra, complex	$\mathbf{(TTCT)}_{10-17}\mathbf{N}_{2-3}\mathbf{(TTCT)}_{3-17}$	12–39	310–438	3.10×10^{-2}
DYF403S1b	Tetra, complex	$(TTCT)_{12}N_2\mathbf{(TTCT)}_8\mathbf{(TTCC)}_9\mathbf{(TTCT)}_{14}N_2(TTCT)_3$	40–59	414–490	1.18×10^{-2}
DYF404S1	Tetra, complex	$\mathbf{(TTTC)}_{10-20}N_{42}(TTTC)_3$	10–20	171–211	1.24×10^{-2}
DYS449	Tetra, complex	$\mathbf{(TTTC)}_{12-18}N_{22}(TTTC)_3N_{12}\mathbf{(TTTC)}_{12-18}$	24–37	309–361	1.22×10^{-2}
DYS518	Tetra, complex	$(AAAG)_3(GAAG)_1\mathbf{(AAAG)}_{14-22}(GGAG)_1(AAAG)_4N_6\mathbf{(AAAG)}_{11-19}$	23–35	243–291	1.84×10^{-2}
DYS526a	Tetra, complex	$(CCCT)_3N_{20}\mathbf{(CTTT)}_{11-17}\mathbf{(CCTT)}_{6-10}N_{113}\mathbf{(CCTT)}_{10-17}$	10–17	138–166	2.72×10^{-3}
DYS526b	Tetra, complex	$(CCCT)_3N_{20}\mathbf{(CTTT)}_{11-17}\mathbf{(CCTT)}_{6-10}N_{113}\mathbf{(CCTT)}_{10-17}$	29–42	345–397	1.25×10^{-2}
DYS547	Tetra, complex	$\mathbf{(CCTT)}_{9-13}T$ $\mathbf{(CTTC)}_{4-5}N_{56}\mathbf{(TTTC)}_{10-22}N_{10}(CCTT)_4(TCTC)_1\mathbf{(TTTC)}_{9-16}N_{14}(TTTC)_3$	36–48	410–458	2.36×10^{-2}
DYS570	Tetra, simple	$\mathbf{(TTTC)}_{14-24}$	10–21	246–286	1.24×10^{-2}
DYS576	Tetra, simple	$\mathbf{(AAAG)}_{13-22}$	13–23	170–210	1.43×10^{-2}

DYS612	Tri, simple	$(CCT)_5(CTT)_1(TCT)_4(CCT)_1(\mathbf{TCT})_{19–31}$	14–31	187–255	1.45×10^{-2}
DYS626	Tetra, complex	$(\mathbf{GAAA})_{14–23}N_{24}(GAAA)_3N_6(GAAA)_5\ (AAA)_1$ $(\mathbf{GAAA})_{2–3}(GAAG)_1(GAAA)_3$	11–23	221–269	1.22×10^{-2}
DYS627	Tetra, complex	$(AGAA)_3N_{16}(AGAG)_3(\mathbf{AAAG})_{12–24}N_{81}(AAGG)_3$	10–24	301–372	1.23×10^{-2}.

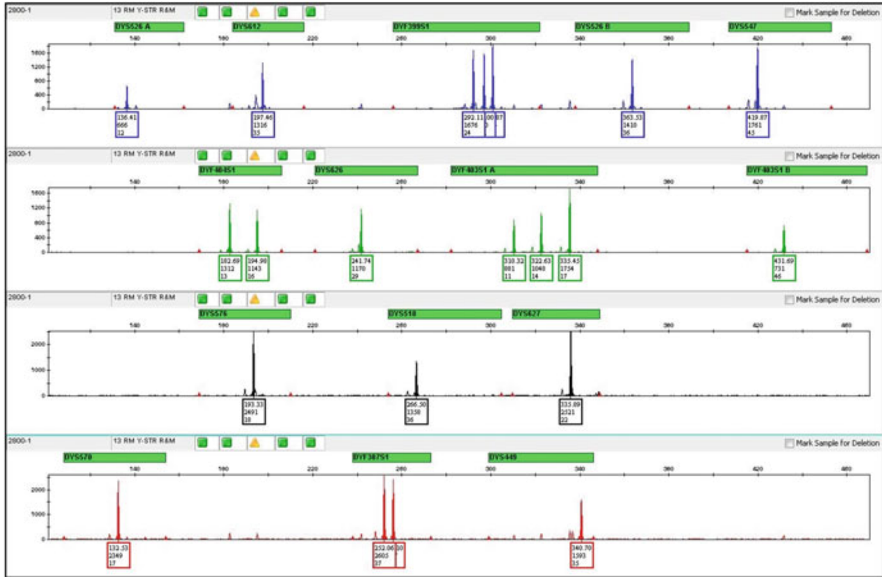


Fig. 14.1 2800M positive control haplotype analyzed by RM-Yplex (Alghafri et al. 2015)

identification kits. Besides, this will also allow scrutinizing more populations for RM Y-STRs rapidly and effortlessly. Despite the extremely high discrimination capacity of these markers, a discovery of additional 27 new rapidly mutating loci was announced in the 27th International Society of Forensic Genetics Congress, which are expected to have a tremendous discrimination among close male relatives as well as non-related male individuals (ISFG Meeting 2017, South Korea) (Figs. 14.1 and 14.2).

14.3 Y-STR Haplotype Interpretation

Y-STR interpretation differs from the interpretation of autosomal STRs in many ways. This is because Y-STR loci are considered one locus due to its inheritance pattern through male lineage. Due to the fact that Y-STRs are very specific to Y chromosome, there will always be one allele per locus on this chromosome, unless the locus being used is multi-allelic locus. In case of the later, there will be two approaches for genotyping. First approach is the conservative typing (C-Type) and the second approach is the extended typing (E-Type) (Alghafri 2015). In C-Type, peak heights for multi-allelic locus are not taken into consideration, whereas in E-Type, peak heights are taken into consideration as it is illustrated in Table 14.3. The total number of alleles present at each peak detected depends on the shortest peak at the locus where a peak which has as double height as the shortest peak will be

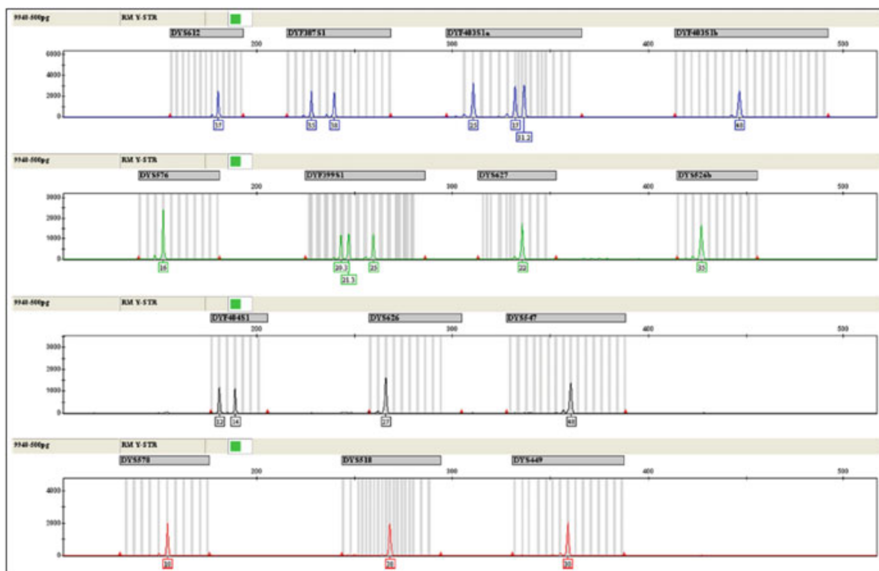


Fig. 14.2 9948 positive control haplotype analyzed by RM Y-STR assay (Zhang et al. 2017)

Table 14.3 Comparison between C-Type and E-Type genotyping interpretation methods

	Haplotype 1	Haplotype 2
Electropherogram	DYF399S1 20 2784 276.50; 21 2735 280.45; 22 2632 284.30; 23 2352 288.19; 25.1 2203 295.88; 26.1 2085 300.77	DYF399S1 20 2499 276.46; 22 4618 284.17; 23 5972 288.01; 24.1 2231 292.89; 25.1 1936 296.75
C-type genotyping	20–21–22–23–25.1–26.1	20–22–23–24.1–25.1
E-type genotyping	20–21–22–23–25.1–26.1	20–22–**22**–23–**23**–**23**–24.1–25.1

counted twice. Such effect is mostly used in the autosomal STR interpretation where heterozygote and homozygote terms are used.

14.3.1 Rarity of the Y-STR Haplotype

There are two main methods for estimating the weight of the Y-STR haplotype matches, those are counting method and likelihood ration with incorporating the potential of the co-ancestry in the same population. The later approach was achieved by different models which will be explained later in this chapter. The commonly used method is the counting method which is estimation for the rarity of a haplotype. In order to reflect the uncertainty in the point estimate which is provided by the counting method, sampling correction factors such as upper and lower limits of confidence have been used in this method. Three main steps should be followed in order to estimate the rarity of a profile. First, Y-STR haplotype should be generated. Second, searching the Y-STR haplotype against relevant database. Third, reporting the rarity of the Y-STR haplotype using genetic and statistical methods. Because of how difficult, costly, and the intensive labor work needed in analyzing all male individuals in all populations, there should at least be a haplotype database developed for each population carefully in an attempt to represent as much as possible of the male individual's haplotypes presented in the population. Therefore, as per recommendation of the International Society of Forensic Genetics (ISFG) (Gusmão et al. 2017), at least 500 random samples from each population should be collected, analyzed, and studied. The larger the number of the haplotypes studied in the population, the more realistic estimation will be obtained in estimating the rarity of a haplotype. As a solution for the rare haplotypes or those haplotypes which might not be shown in the database, it is recommended to use a minimum frequency of such haplotypes as per the equation:

$$\text{Haplotype frequency} = \frac{3}{N}$$

where, N is the total number of haplotypes in the database (Butler 2013). It is very important to choose the right wordings in reporting the result of a match between reference haplotype and evidence haplotype. This is due to the fact that such match is associated with paternal lineage instead of specific individual knowing that there are other reference samples which most likely will be matching such evidence that are coming from father, brothers, sons, and other close male relatives in addition to the random match between non-relatives when some types of Y-STR are in use.

14.3.2 Y-STR Interpretation Guidelines

In 2001, the DNA commission of the ISFG has published general recommendations regarding Y-STR analysis and interpretation (Gill et al. 2001) which then has been updated and republished in 2006 (Gill et al. 2006). In addition, the Scientific Working Group on DNA Analysis Methods (SWGDAM) has established an interpretation guideline for Y-STR haplotypes in 2009 (SWGDAM 2009) aiming at the standardization of the interpretation and analysis, which will help sharing of data

Table 14.4 Summary of different haplotype compositions from previous studies

Standardized aspect	Details
Core Y-STR loci	European 9-loci minimal haplotype (MHT) selected in 1997.
	SWGDAM Y-STR Committee selected 11-loci core haplotype (MHT + DYS438 and DYS439) in January 2003
Consistent allele nomenclature	NIST Standard Reference Material 2395 Human Y-Chromosome DNA Profiling Standard (2003)
	Kits allelic ladders
	ISFG (2006) publication
	NIST (2008) publication
Commercially available Y-STR kits	Early ReliaGene kits (2001–2003)
	PowerPlex® Y (2003)
	Yfiler® (2004)
	PowerPlex® Y23 (2012)
	Yfiler plus™ (2014)
Haplotype databases	YHRD (158,81,211-loci haplotypes for more than 1128 worldwide population)
	U.S. Y-STR (35,660 11-loci haplotypes from primary U.S. population groups)
Interpretation guidelines	SWGDAM Y-STR Interpretation Guidelines published in January 2009 (revision made in January 2014)

between worldwide laboratories. These guidelines were revised and an updated guideline has been published in 2014 (Short Tandem Repeats Internet Database [Online] 2014). Both recommendations comprise of selecting core set of Y-STR loci, availability of commercial assays, consistent alleles nomenclature, online searchable database, and interpretation guidelines as it is shown in Table 14.4.

14.3.3 Y-STR Haplotype Databases

Having these guidelines in place helped laboratories to produce the same results which can be compared to any haplotypes generated from different forensic laboratory around the world. The agreement on using a specific set of Y-STRs and continuously updating that set helped commercializing these sets and also made companies focusing on enhancing these kits to be more robust, sensitive, and precise for forensic and kinship applications. Because of analyzing the same set of Y-STRs, following the recommended nomenclature, especially with the complex types of Y-STRs, and also using single-source developed allelic ladder included in the kits, the development of easily-accessible online Y-STR databases was achieved. There are two main Y-STR haplotype databases: U.S. Y-STR Database and YHRD.org databases. All haplotyped in the databases are anonymous and can be used to estimate the frequency of specific haplotypes.

14.3.3.1 YHRD.org Database

The Y-STR Haplotype Reference Database (YHRD) is an online accessible Y-STR haplotypes and Y-SNP haplogroups databases, which was constructed by Lutz Roewer and Sascha Willuweit since 2000 (Roewer et al. 2005; Willuweit and Roewer 2007). It is considered the largest Y-STR database being used for both forensic and population genetics applications. As of October 2017 (Release 55), YHRD comprises 197,102 minimal Y-STR haplotypes representing 1128 worldwide populations. Moreover, the database comprises other commercially available kits including those mentioned in Table 14.1. It also gives the user the ability to search the maximal set of Y-STR included in this database (29 Y-STR loci). In addition, it gives user the option for searching by geographical location or populations groups.

14.3.3.2 U.S. Y-STR Database

The U.S. population-specific Y-STR database (US Y-STR) was constructed in the National Institute of Justice in December 2007 and managed by the National Center for Forensic Science (NCFS). It started with 11-loci SWGDAM recommended Y-STR haplotypes from five different US populations including Caucasian, Asian, African American, Native American, and Hispanic (Fatolitis and Ballantyne 2008). As per February 2017 (Release 4.2), it comprises 35,660 SWGDAM recommended 11-loci Y-STR haplotypes where some of these haplotypes are generated using commercially available kits which allow the comparison of up to 27-loci haplotypes generated by Yfiler Plus® kit. In U.S. Y-STR database, a duplication of the results is avoided by obtaining the autosomal profiles whenever two different samples show the same Y-STR haplotypes. Such an approach will help obtaining realistic frequency estimation of the Y-STR haplotypes.

14.3.3.3 Other Databases

There are also many companies which are holding commercial Y-STR database such as Ybase (run by genetic genealogy companies) and Ysearch (run by Family Tree DNA company) which are considered as genetic genealogy databases. These databases are not being used for forensic applications. Additionally, these databases are using different sets of Y-STR which might not allow searching haplotypes against all loci recommended for forensic application. Therefore, they are more useful for population genetics and ancestry studies.

There are also some loci such as RM Y-STRs which are not yet added to the currently online-running Y-STR database yet. Although there are more than one multiplexes developed to obtain all RM Y-STR in single multiplex reaction spontaneously (Alghafri et al. 2015; Zhang et al. 2017), these kits are not yet commercialized. Therefore, they are not available for application in the laboratories which makes it difficult to apply and use as standard method. However, up-to-date, there are more than 20,000 RM Y-STR haplotypes generated through different publications following the recommended nomenclature of the loci allowing comparison of haplotypes generated with both methods valid (Ballantyne et al. 2014; Zhang et al. 2017).

14.4 Reporting Y-STR Results

Unlike autosomal STR, Y-STR cannot be presented in courts following the power of discrimination and weight of autosomal STRs by applying "product rule" for probability determination because of the lack of recombination in amelogenin chromosomes. Therefore, counting method as mentioned above is the best approach to estimate the rarity of a specific Y-STR haplotype in a relevant database. This simply can be obtained by dividing the number of times a haplotype has been observed in the database divided by the total number of available haplotypes in database. Therefore, the larger the number of haplotypes in the database of male individuals, the better rarity estimation will be obtained.

14.4.1 Pseudo-Count Method

In 2005, David Balding has published a simple method for estimating lineage markers haplotype rarity in order to obtain a useful and conservative weight of evidence according to the following equation:

$$p = \frac{x + 2}{N + 2}$$

where p is determined by adding two observations, including evidence and suspect to the haplotype count (x) resulted from a specific haplotype database enquiry and also to the population database itself. For example, if an enquiry retained two counts in YHRD minimal haplotype database, the result will be:

$$p = \frac{x + 2}{N + 2} = \frac{2 + 2}{197,102 + 2} = \frac{4}{197,104} = 0.002\%$$

and this is more conservative than straight counting method which will be 0.001%.

14.4.2 Confidence Intervals

Confidence intervals is a range of estimate used to reflect the reliability of a statistical estimate based on the data obtained. This approach is assuring that all obtained data produced with the same procedure are falling within a given level of confidence as well as the data which might be obtained in the future. The most commonly used confidence interval level is 95%.

The 95% confidence intervals of a binomial distribution of data has been recommended by 2009 SWGDAM Y-STR Interpretation Guidelines (SWGDAM 2009); nevertheless, this approach has been reported to be problematic with small population sizes or few observations (Buckleton et al. 2011). Clopper–Pearson approach to estimate the upper limit of 95% confidence interval has been

recommended by SWGDAM Y-STR in 2014 (Short Tandem Repeats Internet Database [Online] 2014). This approach gives more conservative value compared to "normal" approach especially for the very low counts of haplotypes. The normal approach to estimate the 95% confidence interval can be obtained using the following equation:

$$p + 1.96\sqrt{\frac{(p)(1-p)}{N}}$$

where the frequency (p) is determined from the number of observations (x) in a database containing (N) number of profiles (Holland and Parsons 1999; Tully et al. 2001). To overcome the problem in the small population size or very few observations, the following Clopper–Pearson formula has been recommended:

$$\sum_{k=0}^{x} \binom{N}{k} p_0^k (1-p_0)^{N-k} = 0.05$$

where (p) is the haplotype frequency as (x) observations are expected to occur 5% of the time. (N) is the population data size, and (x) is the number of observations of the haplotype in the database, $k = 1, 2, 3, 4 \ldots x$ observations. When both the methods are applied to estimate the 95% confidence interval, a value of 0.00141% will be obtained for the normal approach and 0.00319% will be obtained for the Clopper–Pearson approach. Although both approaches were endorsed previously by SWGDAM, other authors found that confidence intervals are not useful as it is unable to approximate reality (Brenner 2013).

14.4.3 Population Substructure Effect

Estimating the frequency of haplotype following the simple counting method which depends on the total number of the population samples does not take into account population substructure effect which will eventually affect the result. Therefore, an equation for the match probability ($\acute{p}$) was derived from Balding–Nichols DNA profile match probability theory which takes into account population substructure as follows:

$$'p = \theta + p(1-\theta)$$

where (p) is the frequency of a particular Y-STR haplotype estimated using counting method based on the database such as YHRD and (θ) is a co-ancestry correction factor for substructure effect.

It has been recognized that population substructure is an evidence between populations around the world. In 2012 Cockerton et al. have published (θ) value range of 0.0000–0.0731, which has been derived from 56 published Y-STR population studies (Cockerton et al. 2012). Because of such range of values, the choice of

an appropriate and relevant (θ) value might be challenging especially in the partial Y-STR haplotypes. The more loci are used in the haplotype in question, the less (θ) value will be and vice-versa. Consequently, in partial profile with less loci, the number of shared haplotypes within and among population will increase, and eventually, (θ) value will increase and at some point, such correction will not cover the upper limit of the count proportion and as a results (θ) correction might not be truly powerful in practical basis (Budowle et al. 2009).

14.4.4 Rare Haplotypes

Taking into account the continuous increasing in the number of loci involved in the newly developed kits as well as the use of highly discriminating loci such as RM-Y-STR, it will be more often that no match is resulting when Y-STR haplotype is searched in the database. Therefore, in order to effectively assign weight to a rare haplotype, Charles Brenner has recommended an approach called "kapa model" which helps predicting the rarity of the future haplotypes which might be observed. It takes into account a fraction of singletons, once observed haplotypes, in the database. Since it has been shown previously that almost 95% of the haplotypes available in the YHRD database are singletons for 17 loci available in the Yfiler® kit, it means 95% of the time that a new Yfiler® haplotype will show no match in the database (Butler 2013). The likelihood ratio is estimated by "Kappa Model" by using the following equation:

$$\mathrm{LR} = \frac{N}{1-k}$$

where (N) is the number of samples in the database, and kappa (k) is the fraction of the singletons in the database. Taking into account the fraction of singletons above for the Yfiler® haplotypes in the YHRD which was found to be 95%, and therefore, the likelihood ratio will be estimated as follows:

$$\mathrm{LR} = \frac{N}{1-k} = \frac{N}{1-0.95} = \frac{N}{0.05} = 20 \times N$$

Therefore, the likelihood ratio is 20 times stronger for rare 17 loci haplotypes which exaggerate the value of haplotype as evidence in the forensic applications. This approach has also been endorsed by another research study in Denmark (Andersen et al. 2013). Moreover, it has been implemented in YHRD online Y-STR database, as an option, in addition to other methods for estimating the rarity of Y-STR haplotypes. When such approach will be applied for RM Y-STR database the likelihood ratio is expected to be much greater since it has been shown that these markers can differentiate males by up to 99% (Ballantyne et al. 2014). So, it will exaggerate the likelihood ratio by 100 times.

14.4.5 Mixture Analysis

One of the major advantages that Y-STR typing has in forensic DNA typing is the easiness of mixture interpretations. This is because only one allele per contributor is expected at each single-allelic locus. Therefore, it is easy to predict the number of male contributor to a biological stain. Such feature is considered very helpful especially in sexual assault cases where there is more than one preparator. Whereas multi-allelic loci are very complicated in mixture analysis especially those which has more than two alleles per locus, in spite of such loci might be helpful in differentiating close male relatives when there is only one contributor to the biological stain.

14.5 Summary

Y-STR typing is a useful tool for different application including genetic genealogy and forensic DNA typing. The researches of Y-STR haplotypes are growing rapidly aiding better understanding of population substructures and also enriching the database such as U.S. Y-STR and YHRD which will help obtaining better estimation of haplotype frequency using the approaches discussed earlier in this chapter. It will also help predicting biogeographical ancestry and even surnames with certain set of loci. Commercially available kits for Y-STR typing are improving continuously to increase the discriminating capacity of males especially close relatives. A lot of promising efforts by different researchers around the globe help increasing the value of Y-STR haplotypes in forensic cases to aid justice.

References

Alghafri R (2015) An evaluation of rapidly mutating Y-STR multi-allelic markers. Forensic Sci Int Genet Suppl Ser 5:e647–e649

Alghafri R, Goodwin W, Hadi S (2013) Rapidly mutating Y-STRs multiplex genotyping panel to investigate UAE population. Forensic Sci Int Genet Suppl Ser 4(1):e200–e201

Alghafri R, Goodwin W, Ralf A, Kayser M, Hadi S (2015) A novel multiplex assay for simultaneously analyzing 13 rapidly mutating Y-STRs. Forensic Sci Int Genet 17:91–98

Andersen MM et al (2013) Estimating Y-STR allelic drop-out rates and adjusting for interlocus balances. Forensic Sci Int Genet 7:327–336

Ballantyne KN et al (2010) Mutability of Y-chromosomal microsatellites: rates, characteristics, molecular bases, and forensic implications. Am J Hum Genet 87:341–353

Ballantyne KN et al (2012) A new future of forensic Y-chromosome analysis: rapidly mutating Y-STRs for differentiating male relatives and paternal lineages. Forensic Sci Int Genet 6:208–218

Ballantyne KN et al (2014) Toward male individualization with rapidly mutating Y-chromosomal short tandem repeats. Hum Mutat 35:1021–1032

Brenner CH (2013) Understanding Y haplotype evidence—(forget what you know from autosomal). Presentation at ISFG 2013 meeting. http://isfg2013org/wp-content/uploads/2012/06/Understanding-Y-haplotype-evidence-for-ISFG-sitepdf. Accessed 4 Apr 2017

Brinkmann B et al (1998) Population genetic diversity in relation to microsatellite heterogeneity. Hum Mutat 11:135–144
Buckleton JS et al (2011) The interpretation of lineage markers in forensic DNA testing. Forensic Sci Int Genet 5:78–83
Budowle B et al (2009) Taxes population substructure and its impact on estimating the rarity of Y-STR haplotypes from DNA evidence. J Forensic Sci 54:1016–1021
Buhler EM (1980) A synopsis of the human Y chromosome. Hum Genet 55:145–175
Butler JM (2013) Advanced topics in: forensic DNA typing: interpretation. Elsevier Academic Press, San Diego
Butler JM et al (2002) A novel multiplex for simultaneous amplification of 20 Y chromosome STR markers. Forensic Sci Int 129:10–24
Charlesworth B (2009) Effective population size and patterns of molecular evolution and variation. Nat Rev Genet 10:195–205
Coble MD et al (2009) Mystery solved: the identification of the two missing Romanov children using DNA analysis. PLoS One 4(3):e4838
Cockerton S et al (2012) Interpreting lineage markers in view of subpopulation effects. Forensic Sci Int Genet 6:393–397
D'Amato ME et al (2009) Evaluation of 21 Y-STRs for population and forensic studies. Forensic Sci Int Genet Suppl Ser 2:446–447
D'Amato ME et al (2010) Characterization of the highly discriminatory loci DYS449, DYS481, DYS518, DYS612, DYS626, DYS644 and DYS710. Forensic Sci Int Genet 4:104–110
Decker AE et al (2007) The impact of additional Y-STR loci on resolving common haplotypes and closely related individuals. Forensic Sci Int Genet 1:215–217
Edwards A et al (1992) Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics 12:241–253
Fan H, Chu JY (2007) A brief review of short tandem repeat mutation. Genom Proteom Bioinform 5:7–14
Fatolitis L, Ballantyne J (2008) The US Y-STR Database. Profiles in DNA 11(1):13–14
Gill P et al (2001) DNA Commission of the International Society of Forensic Genetics; recommendations on forensic analysis using Y-chromosome STRs. Forensic Sci Int 124:79–85
Gill P et al (2006) DNA commission of the international society of forensic genetics: recommendations on the interpretation of mixtures. Forensic Sci Int 160:90–101
Goedbloed M et al (2009) Comprehensive mutation analysis of 17 Y-chromosomal short tandem repeat polymorphisms included in the AmpFSTR® Yfiler® PCR amplification kit. Int J Legal Med 123:471–482
Gusmão L et al (2006) DNA Commission of the International Society of forensic genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis. Forensic Sci Int 157:187–197
Gusmão L et al (2017) Revised guidelines for the publication of genetic population data. Forensic Sci Int Genet 30(2017):160–163
Gusmaúo L et al (2003) Grouping of Y-STR haplotypes discloses European geographic clines. Forensic Sci Int 134:172–179
Hall A, Ballantyne J (2003) Novel Y-STR typing strategies reveal the genetic profile of the semen donor in extended interval post-coital cervicovaginal samples. Forensic Sci Int 136:58–72
Hammer MF et al (2001) Hierarchical patterns of global human Y-chromosome diversity. Mol Biol Evol 18:1189–1203
Hanson EK, Ballantyne J (2007a) Population data for 48 'Non-Core' Y chromosome STR loci. Legal Med 9:221–231
Hanson EK, Ballantyne J (2007b) An ultra-high discrimination Y chromosome short tandem repeat multiplex DNA typing system. PLoS One 2(8):e688
Harfe BD, Jinks-Robertson S (2000) DNA mismatch repair and genetic instability. Annu Rev Genet 34:359–399

Hedman M et al (2004) Analysis of 16 Y STR loci in the Finnish population reveals a local reduction in the diversity of male lineages. Forensic Sci Int 142:37–43

Hedman M, Neuvonen AM, Sajantila A, Palo JU (2011) Dissecting the Finnish male uniformity: the value of additional Y-STR loci. Forensic Sci Int Genet 5:199–201

Holland MM, Parsons TJ (1999) Mitochondrial DNA sequence analysis—validation and use for forensic casework. Forensic Sci Rev 11:21–50

Jobling MA, Tyler-Smith C (2003) The human Y chromosome: an evolutionary marker comes of age. Nat Rev Genet 4:598–612

Jobling MA, Pandya A, Tyler-Smith C (1997) The Y chromosome in forensic analysis and paternity testing. Int J Legal Med 110:118–124

Kayser M (2017) Forensic use of Y chromosome DNA: a general overview. Hum Genet 136:621–635

Kayser M, Sajantila A (2001) Mutations at Y-STR loci: implications for paternity testing and forensic analysis. Forensic Sci Int 118:116–121

Kayser M et al (1997) Evaluation of Y-chromosomal STRs: a multicenter study. Int J Legal Med 110:125

Kayser M et al (2003) Y chromosome STR haplotypes and the genetic structure of U.S. populations of African, European, and Hispanic ancestry. Genome Res 13:624–634

Kayser M et al (2004) A comprehensive survey of human Y-chromosomal microsatellites. Am J Hum Genet 74:1183–1197

Kayser M et al (2005) Significant genetic differentiation between Poland and Germany follows present-day political borders, as revealed by Y-chromosome analysis. Hum Genet 117:428–443

Klintschar M, Wiegand P (2003) Polymerase slippage in relation to the uniformity of tetrameric repeat stretches. Forensic Sci Int 135:163–166

Kroutil LC et al (1996) Exonucleolytic proofreading during replication of repetitive DNA. Biochemistry 35:1046–1053

Kruglyak S et al (1998) Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations. Proc Nat Acad Sci U S A. 95:10774–10778

Kuhner F et al (2007) Force-induced DNA slippage. Biophys J 92:2491–2497

Lewontin RC, Hartl DL (1991) Population genetics in forensic DNA typing. Science 254:1745–1750

Li G-M (2008) Mechanisms and functions of DNA mismatch repair. Cell Res 18:85–98

Mulero JJ et al (2006) Development and validation of the AmpFLSTR Yfiler(R) PCR amplification kit: a male specific, single amplification 17 Y-STR multiplex system*. J Forensic Sci 51:64–75

Oota H et al (2001) Human mtDNA and Y-chromosome variation is correlated with matrilocal versus patrilocal residence. Nat Genet 29:20–21

Pascali VL, Dobosz M, Brinkmann B (1998) Coordinating Y-chromosomal STR research for the courts. Int J Legal Med 112:1–1

Pearson CE, Sinden RR (1998) Trinucleotide repeat DNA structures: dynamic mutations from dynamic DNA. Curr Opin Struct Biol 8:321–330

Poznik GD et al (2013) Sequencing Y chromosomes resolves discrepancy in time to common ancestor of males versus females. Science 341:562–565

Purps J et al (2014) A global analysis of Y-chromosomal haplotype diversity for 23 STR loci. Forensic Sci Int Genet 12:12–23

Quintana-Murci L, Fellous M (2001) The human Y chromosome: the biological role of a "functional wasteland". J Biomed Biotechnol 2001:18–24

Quintana-Murci L, Krausz C, McElreavey K (2001) The human Y chromosome: function, evolution and disease. Forensic Sci Int 118:169–181

Redd AJ et al (2002) Forensic value of 14 novel STRs on the human Y chromosome. Forensic Sci Int 130:97–111

Rodig H et al (2008) Evaluation of haplotype discrimination capacity of 35 Y-chromosomal short tandem repeat loci. Forensic Sci Int 174:182–188

Roewer L et al (1992) Simple repeat sequences on the human Y chromosome are equally polymorphic as their autosomal counterparts. Hum Genet 89:389–394
Roewer L et al (2005) Signature of recent historical events in the European Y-chromosomal STR haplotype distribution. Hum Genet 116:279–291
Schlotterer C, Tautz D (1992) Slippage synthesis of simple sequence DNA. Nucleic Acids Res 20:211–215
Schneider PM et al (1998) Tandem repeat structure of the duplicated Y-chromosomal STR locus DYS385 and frequency studies in the German and three Asian populations. Forensic Sci Int 97:61–70
Seyfert AL et al (2008) The rate and spectrum of microsatellite mutation in Caenorhabditis elegans and Daphnia pulex. Genetics 178:2113–2121
Shi W et al (2010) A worldwide survey of human male demographic history based on Y-SNP and Y-STR data from the HGDP-CEPH populations. Mol Biol Evol 27:385–393
Short Tandem Repeats Internet Database [Online] (2014). http://www.cstl.nist.gov/strbase/mutation.htm. Accessed 7 Apr 2017
Shriver MD et al (1993) VNTR allele frequency distributions under the stepwise mutation model: a computer simulation approach. Genetics 134:983–993
Sia EA, Jinks-Robertson S, Petes TD (1997) Genetic control of microsatellite stability. Mutat Res DNA Repair 383:61–70
Sibille I et al (2002) Y-STR DNA amplification as biological evidence in sexually assaulted female victims with no cytological detection of spermatozoa. Forensic Sci Int 125:212–216
Stumpf MPH, Goldstein DB (2001) Genealogical and evolutionary inference with the human Y chromosome. Science 291:1738–1742
Sup Lee J et al (1999) Relative stabilities of dinucleotide and tetranucleotide repeats in cultured mammalian cells. Hum Mol Genet 8:2567–2572
SWGDAM (2009) Y-STR interpretation guidelines. Forensic Sci Commun 11
Thompson R, Zoppis S, McCord B (2012) An overview of DNA typing methods for human identification: past, present, and future. Humana Press, Totowa, NJ
Thompson JM et al (2013) Developmental validation of the PowerPlex- Y23 system: a single multiplex Y-STR analysis system for casework and database samples. Forensic Sci Int Genet 7:240–250
Tiraby JG, Fox MS (1973) Marker discrimination in transformation and mutation of pneumococcus. Proc Nat Acad Sci U S A 70(12):3541–3545
Tully G et al (2001) Consiideration by the European DNA profiling (EDNAP) group on the working practices, nomenclature and interpretation of mitochondrial DNA profiles. Forensic Sci Int 124:83–91
Underhill PA et al (2000) Y chromosome sequence variation and the history of human populations. Nat Genet 26:358–361
Van Oven M et al (2014) Seeing the wood for the trees: a minimal reference phylogeny for the human Y chromosome. Hum Mutat 35:187–191
Vermeulen M et al (2009) Improving global and regional resolution of male lineage differentiation by simple single-copy Y-chromosomal short tandem repeat polymorphisms. Forensic Sci Int Genet 3:205–213
Vuichard S et al (2011) Differential DNA extraction of challenging simulated sexual-assault samples: a Swiss collaborative study. Investig Genet 2:11
Weber JL, Wong C (1993) Mutation of human short tandem repeats. Hum Mol Genet 8:1123–1128
Willott GM, Allard JE (1982) Spermatozoa—their persistence after sexual intercourse. Forensic Sci Int 19:135–154

Willuweit S, Roewer L (2007) Y Chromosome Haplotype Reference Database (YHRD): update. Forensic Sci Int Gen 1(2):83–87

Xue Y et al (2009) Human Y chromosome base-substitution mutation rate measured by direct sequencing in a deep-rooting pedigree. Curr Biol 19:1453–1457

Zhang W et al (2017) Multiplex assay development and mutation rate analysis for 13 RM Y-STRs in Chinese Han population. Int J Legal Med 131:345–350

15 X-STRs: Potentials and Applications

Sonia Kakkar, R. K. Kumawat, and Pankaj Shrivastava

Abstract

The chapter describes potential and utility of X-STRs in forensic DNA analysis. The presently available X-STR multiplex systems, details of known X-STR markers, analysis of X-STR data, and its utility in forensic DNA analysis are also detailed. Besides the analysis of X-STR analysis using capillary electrophoresis system, the chapter also describes possibility of X-STR analysis using the next-generation sequencing.

Keywords

X STR · Forensic DNA analysis · Lineage markers · X chromosome

The DNA profiling or DNA fingerprinting was discovered by Sir Alec Jeffreys in 1985, while conducting an experiment Jeffreys discovered that few regions of DNA contained repetitive sequences and the repetitions were in a tandem (one after the other) fashion. The number of repetitions of these sequences was different in different individuals.

These repeated DNA sequences in human genome are present in various sizes. The repeated sequences are labeled by the length of their repeat unit and the number of continuous repeat units (Ellegren 2004). Longer repeat units usually have

S. Kakkar
Department of Forensic Medicine, Post Graduate Institute of Medical Education and Research (PGIMER), Chandigarh, India

R. K. Kumawat
DNA Division, State Forensic Science Laboratory, Jaipur, Rajasthan, India

P. Shrivastava (✉)
DNA Fingerprinting Unit, State Forensic Science Laboratory, Department of Home (Police), Government of Madhya Pradesh, Sagar, Madhya Pradesh, India

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_15

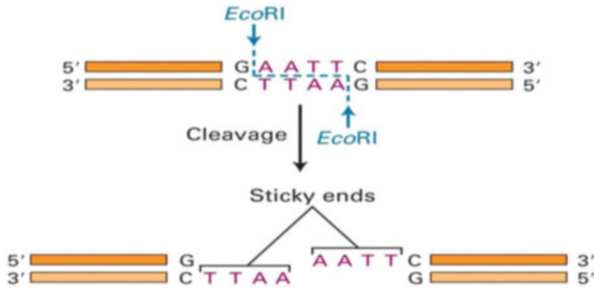


Fig. 15.1 DNA site cut by a restriction enzyme at the sticky ends

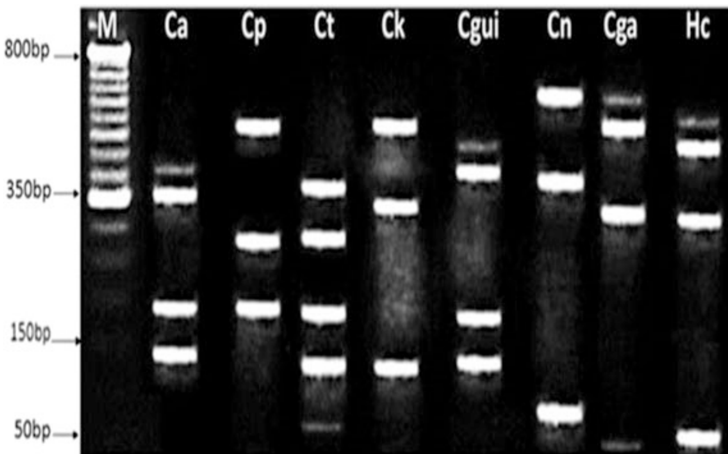


Fig. 15.2 Different length fragments produced by restriction enzymes in different individuals

100–1000 bases in one core repeat. The technique used to analyze these long repeats was known as restriction fragment length polymorphism (RFLP). The technique consisted of cutting the DNA at palindromic sequences (sticky ends) by specific restriction enzymes and created fragments of different lengths in different individuals. This formed the basis of individual identification (Figs. 15.1 and 15.2).

Drawback of RFLP analysis: The RFLP analysis required very good quality DNA as the repeat size was 100–1000 bp long. But in forensic getting a good quality evidentiary sample is not always feasible as the crime samples undergo the exposure to harsh environmental conditions and sometimes are degraded and putrefied. So after RFLPs, variable number tandem repeats (VNTRs), also known as mini-satellites came into the light. In VNTRs, repeat sequence range is from 10 to 100 bp and these are repeated 5–50 times and the number of repeat varies in each and every individual (Fig. 15.3).

DNA fragments with repeat units containing 2–10 bp in length are known as short tandem repeats (STRs). STRs are also known as microsatellite markers. More than

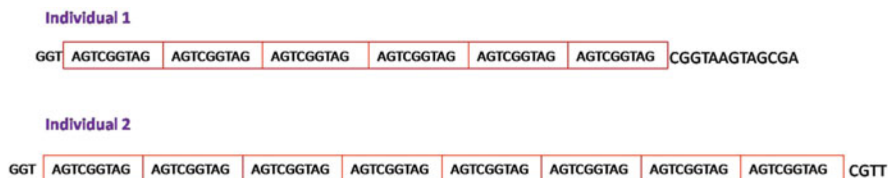


Fig. 15.3 VNTRs, variable number of repeats placed in a tandem fashion varying individual to individual

90% of these markers are found in the Junk DNA, i.e., noncoding DNA, only 8–10% are located in the coding regions (Ellegren 2000; Fan and Chu 2007). STRs are spread in the entire genome and form nearly 3% of the total human genome (Lander et al. 2001). Their distribution, however, is not consistent, they are observed less commonly in subtelomeric portions of DNA (Koreth et al. 1996). Chromosome 19 has the maximum density of STR (Subramanian et al. 2003). After every 2000 bp, on an average one STR occurs. The alleles in STR regions typically range from 50 to 300 bp. In forensic practice, the STR markers that are most commonly used consist of tetranucleotide repeats, i.e., each repeat motif contains 4 bp. The repeats are placed in a tandem fashion. Most common STR markers used for human identification are rich in adenine (A) (Nadir et al. 1996). STR segments do not code for any proteins or phenotypic characters as they lie in the junk region of DNA. The nomenclature of STRs is as, e.g., D2S1338, where D represents DNA, 2 represents the location of STR, i.e., chromosome number 2, S stands for single copy sequence, and 1338 is the unique identity number assigned to the STR. All the chromosomes including 22 autosomal pairs, X and Y sex chromosomes have STR regions.

In autosomal STR markers inheritance, a person receives one allele from father (paternal) and one from the mother (maternal). For every locus, individual with both the alleles same from mother and father is homozygous and heterozygous in case of different alleles from mother and father. While in case of X STRs inheritance daughter receives one of her two X alleles from the father (paternal) and the other one from the mother (maternal), on the other side since the son has one X chromosome, he receives only one allele (on all the X STR loci) from the maternal side only. And in case of Y-STR inheritance, son receives Y chromosome from the father as a haplotype so all the alleles from father to son are inherited as it is.

Status of STR analysis so far: Two types of STR analysis have been used so far to resolve the paternity cases: Autosomal STRs and Y-STR (El-Alfy and El-Hafez 2012; Hill et al. 2009; Jobling et al. 1997; Kayser and Sajantila 2001). The autosomal STR analysis is the most common practice in forensic laboratories, which is done on the markers present on the autosomal chromosomes. Nowadays, various autosomal STR kits are commercially available such as AmpF*l*STR® Identifiler® Plus, GlobalFiler™, VeriFiler™ Plus (Thermo Fisher Scientific), PowerPlex® 21, PowerPlex® Fusion 5C and PowerPlex® Fusion 6C (Promega), and Qiagen Investigator 24 Plex (Qiagen) and are being used in forensic DNA testing. These autosomal STR kits simultaneously co-amplify the various STR markers. For example

the AmpFlSTR® Identifiler® Plus PCR Amplification Kit is a multiplex system that amplifies autosomal15 STR loci (tetranucleotide) in a single reaction and also one gender-informing marker, Amelogenin.

- Along with D2S1338 and D19S433, all the 13 loci of Combined DNA Index System (CODIS) loci are amplified by this kit.
- These 15 loci are accepted worldwide for forensic use.

The need for new markers: The prerequisite of autosomal STR analysis was requirement of both the parents to set the identity of an individual as every individual gets one copy from father and one from mother. Moreover, the mutation rate is high in autosomal chromosomes (Kayser and Sajantila 2001). Same was the problem with Y-STR analysis as it could tell only the paternal inheritance because the son gets Y chromosome from father and daughter has no Y chromosome. For maternal inheritance, mitochondrial DNA analysis was done, as the mitochondrial DNA from mother only passes to the child (girl and boy both) and mitochondrial DNA from the father which is present in the tail of the sperm gets digested at the time of fertilization (Gyllensten et al. 1985; Schwartz and Vissing 2002). Mitochondrial DNA analysis could only be used for the study of maternal inheritance because an individual doesn't get mitochondrial genes from father as the mitochondria are present in sperm tail and tail of the sperm is vanished at the time of fertilization. Moreover, mitochondrial DNA analysis requires strict standardization protocols and the analysis is time-consuming too (Parson et al. 1998). So, in this scenario, a particular set of markers was needed which could be capable of kinship analysis in difficult situations like deficiency paternity and sibship and half sibship analysis.

15.1 X-STR in the Picture

The X chromosome is 150 million base pair (Mb) in humans, with a low gene density. The X chromosome is physically the most stable amongst all other nuclear chromosomes and this is the only chromosome which still retains syntenic loci between mouse and human (Waterston et al. 2002). According to Lyon's hypothesis, individuals with more than one copy of X chromosome have only one X active per cell. This is why Chromosome X monosomies, trisomies, and polysomies are compatible with life (Lyon 1961).

The population genetics studies of X chromosome were started after the DNA sequencing technology came into limelight and first detailed sequence of X chromosome was revealed in 1997 and 1998 (Schaffner 2004; Ziętkiewicz et al. 1997, 1998). The first phylogenetic tree of X chromosome appeared in 1999 (Jaruzelska et al. 1999; Kaessmann et al. 1999), whereas the first phylogenetic tree of mitochondrial DNA was published in 1987 (Cann et al. 1987) and that of Y was published in 1989 (Lucotte et al. 1989). X and Y both the sex chromosomes were originated from a single autosome around 300 million years ago (Lahn and Page 1999). The use of X chromosomal STRs is being explored in various forensic caseworks. The X-STRs

Table 15.1 Chromosomal mapping of Investigator® Argus X-12

Locus	Chromosomal mapping
Amelogenin X	Xp22.1–22.3
Amelogenin Y	Yp 11.2
DXS7132	Xq 11.2
DXS7423	Xq28
DXS8378	Xp22.31
DXS10074	Xq 12
DXS10079	Xq 12
DXS10101	Xq26.2
DXS10103	Xq26.2
DXS10134	Xq28
DXS10135	Xp22.31
DXS10146	Xq28
DXS10148	Xp22.31
HPRTB	Xq26.2

Source: Handbook, Investigator Argus, X-12, Qiagen, Germany

are supposed to offer additional information in various kinship cases than provided by the autosomal STRs. Any kinship situation under investigation where at least one female is involved is better studied by X-STRs, which can further be applied to cases of incest, missing person identification, immigration disputes, deficiency paternity, or other questioned kinship (Tables 15.1, 15.2, 15.3 and 15.4; Figs. 15.4, 15.5, 15.6 and 15.7). Genetic data of the most common used 12 X-STR linked markers of the Investigator® Argus X-12 kit/Investigator® Argus X-12 kit with quality sensor have been published in the handbook the abovementioned kit as follows:

Other X-STR markers explored so far: Nakamura et al. (2013) studied 18 X-STR markers in Japanese population and designed a multiplex PCR for the markers (Nakamura et al. 2013). To these 18 X-STR markers, Fukuta et al. (2018) added nine more markers which may exist in linkage disequilibrium and method for the calculation of likelihood ratios in kinship test using all the 27 markers (18 + 9) was designed (Fukuta et al. 2018).

List of all the X-STR markers explored so far:

List of 18 X-STR markers reported by Nakamura et al. (2013)	List of 9 X-STR markers added by Fukuta et al. (2018) to the previously studied markers by Nakamura et al. (2013)	12 X-STR markers available in multiplex PCR kit by QIgen
DXS6807	DXS10076	DXS7132
DXS9902	DXS10077	DXS7423
DXS6795	DXS10078	DXS8378
DXS6810	DXS10162	DXS10074
DXS7132[a]	DXS10163	DXS10079
DXS981	DXS10164	DXS10101
DXS6800	DXS6799	DXS10103
DXS6803	DXS10103[a]	DXS10134

(continued)

Table 15.2 Localization and other genetic information of 12 X-STR markers of Investigator® Argus X-12 kit on X chromosome

Locus	GenBank® Accession number	Repeat motif of the reference allele	Ref. allele	Allele range
Amelogenin X	M55418	–	–	–
Amelogenin Y	M55419	–	–	–
DXS7132	G08111	$[TCTA]_{13}$	13	8–20
DXS7423	AC109994	$[TCCA]_3$ TCTGTCCT $[TCCA]_{12}$	15	8–19
DXS8378	G08098	$[CTAT]_{12}$	12	7–15
DXS10074	AL356358	$[AAGA]_{14}$	14	4–21
DXS10079	AL049564	$[AGAG]_3$ TGAAAGAG $[AGAA]_{17}$ AGAG $[AGAA]_3$	21	14–25
DXS10101	AC004383	$[AAAG]_3$ GAAAGAAG $[GAAA]_3$ A $[GAAA]_4$ AAGA $[AAAG]_5$ AAAAAGAA $[AAAG]_{13}$ AA	28.2	24–38
DXS10103	BV680555	$[TAGA]_2$ CTGA [CAGA]$[TAGA]^{11}$ $[CAGA]^{4}$[TAGA]	19	15–21
DXS10134	AL034384	$[GAAA]_3$ GAGA $[GAAA]_4$ AA [GAAA] GAGA $[GAAA]_4$ GAGA $[GACAGA]_3$ [GAAA] GTAA $[GAAA]_3$ AAA $[GAAA]_4$ AAA $[GAAA]_{15}$	35	28–46.1
DXS10135	AC003684	$[AAGA]_3$ GAAAG $[GAAA]_{20}$	23	13–39.2
DXS10146	AL034384	$[TTCC]_3$ $[TTCC]_3$ TTCC CTCCCTTCC [TTCC] [TCCC] TTCTTCTTTC $[TTCC]_2$ TTTCTT $[CTTT]_2$ CTTC $[CTTT]_2$ T$[CTTT]_2$	26	24–46.2
DXS10148	AC003684	$[GGAA]_4[AAGA]_{12}[AAAG]_4$ N_8 $[AAGG]_2$	22	13.3–38.1
HPRTB	M26434	$[AGAT]_{12}$[a]	12	6–19

Source: Handbook, Investigator Argus, X-12, Qiagen, Germany
[a][AGAT] is the most common repeat structure; for variations, see NIST and reference 3

DXS6809	DXS10147	DXS10135
DXS6789		DXS10146
DXS7424		DXS10148
DXS101		HPRTB
DXS7133		
GATA172D05		
HPRTB[a]		
GATA31E08		
DXS8377		
DXS7423[a]		

The highlighted [a] markers are the commonly markers in three different assays

Table 15.3 Allelic ladder fragments used in Investigator® Argus X-12 kit

Locus	Dye label	Repeat numbers of allelic ladder
Amelogenin	6-FAM	X, Y
DXS10103	6-FAM	15, 16, 17, 18, 19, 20, 21
DXS8378	6-FAM	9, 10, 11, 12, 13, 14, 15
DXS7132	6-FAM	10, 11, 12, 13, 14, 15, 16, 17
DXS10134	6-FAM	28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 38.3, 39.3, 40.3, 41.3, 42.3, 43.3, 44.3
DXS10074	BTG	4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 16.2, 17, 18, 19, 20, 21
DXS10101	BTG	24, 24.2, 25, 25.2, 26, 26.2, 27, 27.2, 28, 29, 29.2, 30, 30.2, 31, 31.2, 32, 33, 33.2, 34, 34.2, 35
DXS10135	BTG	13, 14, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37.2, 39.2
DXS7423	BTY	13, 14, 15, 16, 17, 18
DXS10146	BTY	24, 25, 26, 27, 28, 29, 30, 31, 32, 32.2, 33, 33.2, 34, 34.2, 35.2, 39.2, 40.2, 41.2, 42.2, 43.2, 44.2, 45.2, 46.2
DXS10079	BTY	14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 25
HPRTB	BTR	9, 10, 11, 12, 13, 14, 15, 16, 17
DXS10148	BTR	13.3, 18, 19, 20, 21, 22.1, 23.1, 24.1, 25.1, 26.1, 27.1, 28.1, 29.1, 30.1, 31, 38.1

Source: Handbook, Investigator Argus, X-12, Qiagen, Germany

Table 15.4 Fluorescent labels of the dyes of BTO (size standard)

Color	Matrix standard
Blue (B)	6-FAM
Green (G)	BTG
Yellow (Y)	BTY
Red (R)	BTR
Orange (0)	BTO

15.2 Mutation and Recombination Rates of Various X-STR Markers

According to the 1991 report of the International Society for Forensic Genetics (ISFG) before using DNA polymorphism for paternity testing, mutation rates must be known so that mismatches can be attributed to mutation events ("1991 report concerning recommendations of the DNA Commission of the International Society for Forensic Haemogenetics relating to the use of DNA polymorphisms" 1992).

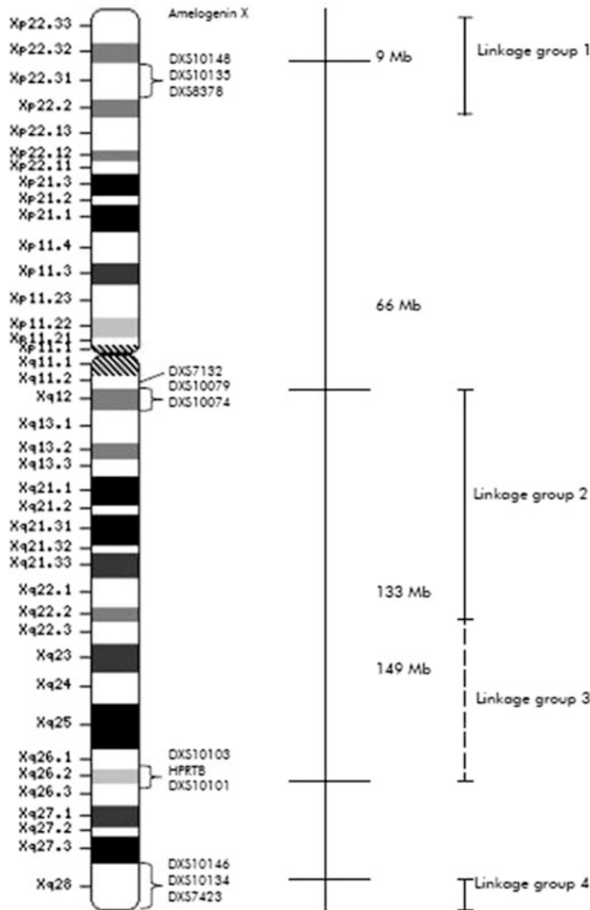


Fig. 15.4 Chromosome X ideogram, positioning of X-STR markers used in forensic casework. The order and positioning of the markers are based on the work published by Marshfield, NCBI. Rest of the information regarding the linkage group and genetic distances were taken from www.chrx-str.org

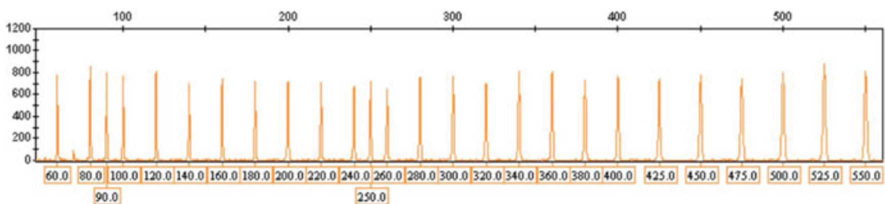


Fig. 15.5 DNA size standard, BTO 550 electropherogram showing fragments length in basepair

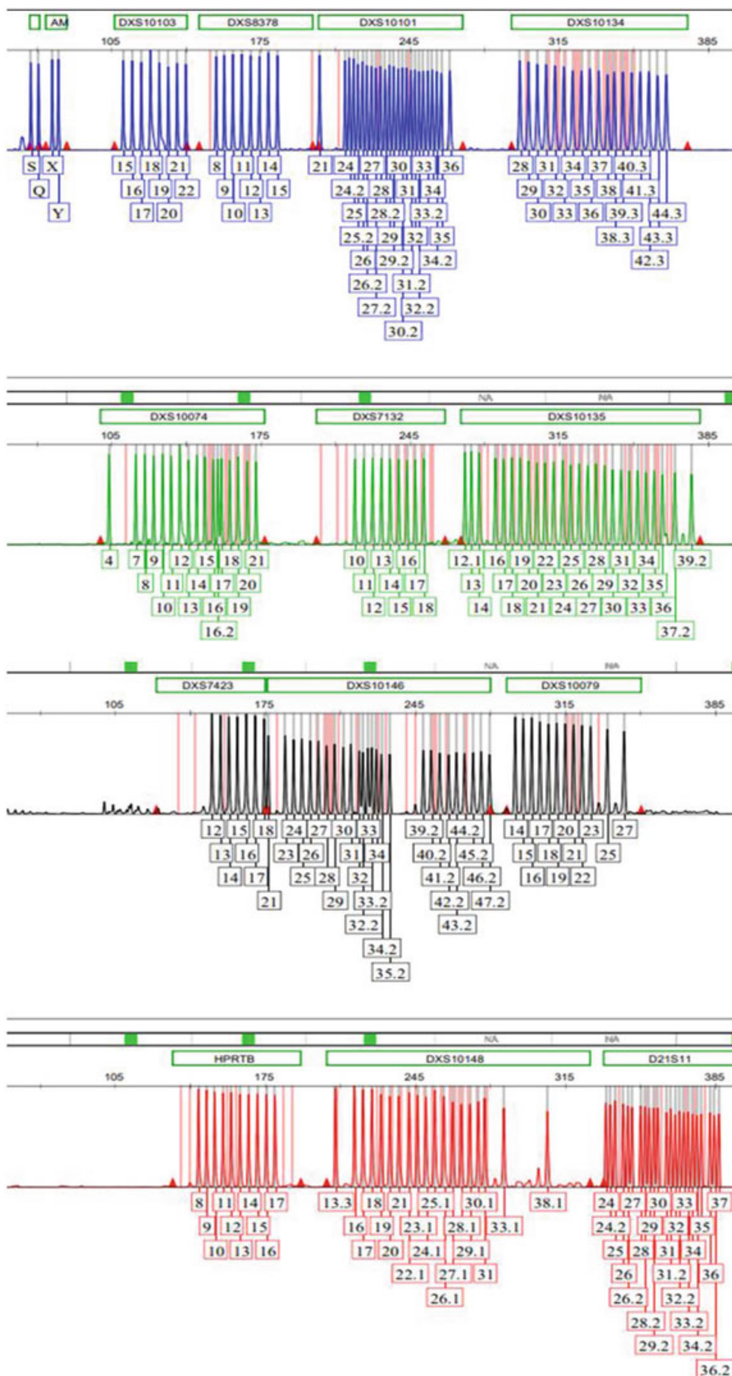


Fig. 15.6 Electropherogram of allelic ladder of Investigator® Argus X-12 kit (with quality sensor)

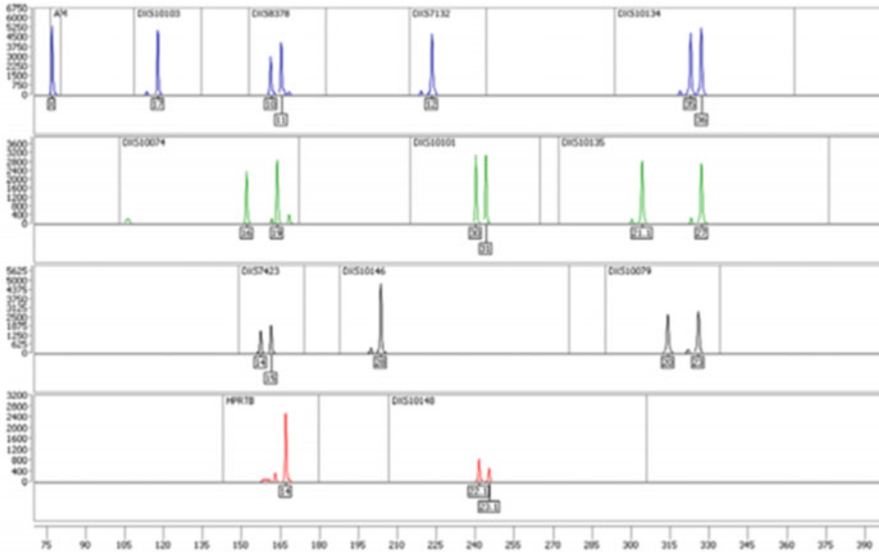


Fig. 15.7 Electropherogram of of a sample amplified by Investigator Argus X-12 kit

After thorough research, it was demonstrated that microsatellite mutation rate on autosomal STR markers is influenced by the structure and length of the repeat pattern and mutation in male germ line was reported to be five to six times higher than the female germ line. The more polymorphic and complex STRs exhibit more mutations (Brinkmann et al. 1998). The mutations in STRs have been explained by the stepwise mutation model by increase or decrease in repeat unit (Xu and Fu 2004). Insertions and deletions are other types of mutations that give rise to microvariant alleles and such alleles are one or few nucleotides lesser than the full repeat unit. The main mechanisms reported for STRs mutation are strand-slippage replication (Kornberg et al. 1964). In this mechanism, the slippage of Taq polymerase enzyme occurs during replication which leads to mispairing (by one or more repeat units) between the template and newly formed strand. The repeated DNA fragment is forced to "loop out" at the mismatched site. If DNA synthesis continued in this situation, the repeat number of the STR is altered on that site (Jobling et al. 2013). Reports have advocated that mutation occurs less frequently on the X chromosome than autosomes (Bohossian et al. 2000; Li et al. 2002; Makova and Li 2002; Schaffner 2004). The low mutation rate results in lower genetic diversity. Diversity is the function of effective population size. Lesser mutation rate and the smaller population size of X chromosome as compared with autosomes advocated that the diversity should be less on X chromosome. Though a direct comparison of diversity between X and autosomes is difficult. Only two out of three X chromosome recombines in a generation as recombination takes place only in females and not in the males. So the recombination rate is higher in females than in males. Also, the

Table 15.5 Comparison of population-genetic markers (Schaffner 2004)

	Marker type			
	mtDNA	Y chromosome	X chromosome	Autosomes
Size (Mb)	0.017	60	150	3000
Number of usable loci	1	1	Hundreds	Thousands
Mutation rate (mutations per Mb per generation)	Very high (1–300)	High (0.033)	Low (0.015)	Moderate (0.020)
Recombination rate (cM/Mb)	0	0	0.8	1.1
Diversity (fraction of discordant base pairs)	Very high (0.4%)	Low (0.02%)	Moderate (0.04%)	High (0.08%)
Accessible haplotypes[a]	Yes	Yes	Yes	No
Genetic drift[b,c]	High	High	Moderate	Low
Age[c]	100,000 years	100,000 years	750,000 years	1,000,000 years
Effective population see (relative to autosomes)	1/4	1/4	3/4	1

[a] A haplotype is a set of genetic markers that is present on one chromosome
[b] Genetic drift describes the random changes in allele frequency that occur because genes that appear in offspring are not a perfectly representative sample of the parental genes (e.g., as occurs in small populations)
[c] These entries are approximate population genetics inferences, based on the consensus estimate for the effective population size in humans
cM centiMorgan, *Mb* megabase, *mt* mitochondrial

mutation rate for maternal transfer was found to be 4.22×10^{-4} and for paternal transfer, it was 1.74×10^{-3} it has been reported in many studies that the paternal mutations are more frequent for both autosomal (Dauber et al. 2012; Fan and Chu 2007; Weber and Wong 1993) and X STRs (Tomas et al. 2012). Higher paternal mutation rates corroborate the idea that the mechanism of STR mutation may be independent of recombination phenomena (Ellegren 2004; Kayser et al. 2000). Recombination in males tends to be higher near the telomere, whereas in females it is higher near the centromere. The measured recombination rate of X chromosome is two-third of the genome average (Kong et al. 2002). The linkage disequilibrium is expected to be greater on X chromosome. Haplotype studies have a vital role in the determination of the phylogeny of a region. Since there is a single copy of X chromosome in males, it is easier to study the haplotype. Whereas haplotype studies cannot be performed on autosomal data. Y and mt DNA can also be used for haplotype studies for paternal and maternal inheritance, respectively. Since autosomes and X chromosome take part in recombination at every generation, this feature makes them a good pick for historical studies as different regions will show different histories. But recombination renders problematic phylogenetic studies, so while studying X-based phylogenies, the doer must be restricted to the areas with high linkage disequilibrium.

The mutation rates and other population genetic parameters of various forensic STR markers have been reported and research suggests that smaller population of the X chromosome makes it more prone to genetic drift than autosomes, which makes the populations differ more in their X chromosomes than in autosomes (Table 15.5).

15.3 X-STR in Forensic

X-STR markers have not been studied extensively in forensic casework so far. Unlike autosomal chromosomes, all the X STR markers are present on the same chromosome in linkage groups. Since the markers are linked, they are inherited in the same pattern as that of the parents. The markers located on X chromosome have a particular inheritance pattern: since women are dizygous and men are hemizygous, haplotype studies can be better performed with X-STR. And due to hemizygous nature of X chromosome in males, the propensity of any random person sharing the same allele in a population is reduced. Male individuals inherit their X-chromosome from their mother, while female individuals receive one X from mother and another from father (as a haplotype), so female individuals fathered by same man share the same paternal X chromosome. The unique inheritance pattern and linkage between the X-STR markers make them a better choice for kinship analysis.

15.3.1 X-STR in Cases of Sexual Assault and Incest

After sexual assault or incest, the microscopical dissection of chorionic villi of 6–8 weeks aborted fetus does not give satisfactory results as it gives a mixed profile

of the fetus and the mother. The aborted fetus also consists of fetal organs, maternal blood, and other tissues. Efficient paternity testing of these materials can be done by Y-STR analysis in case of a male fetus. But when the fetus is female, autosomal and X-STR are used and among them X-STR was found to be more efficient. While in the case of incest X-STR has a limitation as all the alleles of the aborted fetus will be the same as that of the daughter abused by the father. So no conclusive proof of paternity can be presented. It has been reported that X STR markers may help to expose a female profile in male background, such as vaginal cells on a penis or female cells from male fingernail scrapings (Szibor 2007). To identify female traces in male contamination X-STR markers are more efficient than autosomal markers because the merging of female alleles with male alleles will only happen if the female is homozygous at all loci (Szibor et al. 2003). Cases of a gang rape where brothers or paternal cousins are involved cannot be solved by Y-STR as all the brothers and paternal cousins share the same Y-haplotype, which makes it difficult to discern whether the case was of gang rape or not. In such doubtable situation, X-STR can solve the purpose as the offenders will be having a different X STR profile since the X chromosome is inherited from the mother to the son.

15.3.2 Mixture Analysis by X-STR

Mixed DNA profiles may be obtained in forensic casework which makes the individualization a difficult task. In a male–female mixture, autosomal STRs can detect the sex of the contributors when both of the contributors are of the same gender or the male contributor is the minor in the mixture (Budowle et al. 2009). Mixed profiles are usually obtained in cases like mass disasters and in sexual assault, where the sample from the victim is mixed with that of the accused. Various methods have been described earlier to cope with such situations of mixed profile analysis (Buckleton et al. 2016). A multiplex system, MIXplex, consisting of sex chromosome X and Y markers have been reported to aid in mixture interpretation in cases of sexual assaults and is even helpful in deciphering the number of contributors (Diegoli et al. 2013) But the mixture analysis methods given so far can resolve the mixture of unrelated individuals where the linkage is not taken into consideration. Mixed profiles of relatives have been reported to be resolved by some scientists (Fung and Hu 2004, 2008; Hu and Fung 2005). Dørum et al. (2016) have worked on the calculation of the likelihood ratios for mixed profiles involving close relatives (Dørum et al. 2016). Also, the MIXplex can clarify situations where the male allele is not amplified at the amelogenin locus due to a deletion on the Y chromosome without Y chromosome typing (Kashyap et al. 2006; Steinlechner et al. 2002).

15.3.3 Parentage, Deficiency Paternity, and Complex Paternity Analysis by X-STR

As the mutation rates in autosomes are higher, the propensity of false-positive results increases, but it has been reported that false-positive results can be reduced by studying extra loci to confirm the kinship. The problem with studying extra loci is that the rate of mutation also tends to be 1.6 times more (Tamura et al. 2015). Chen et al. (2009) reported that X-STR analysis is helpful in confirming mutations in parentage casework analysis. Any mismatch seen in the routine parentage, analysis by autosomal STRs can be confirmed by X-STR analysis (Chen et al. 2009). Better likelihood ratios (LRs) have been reported by the use of X chromosomal markers in paternity cases (Szibor et al. 2005a, b). Bayesian networks (BNs) describe rules for the inheritance of alleles in various blood relationships. These rules can include inheritance of various markers like autosomal, X chromosomal, mitochondrial, and Y chromosomal. BN is a graphical tool that can represent independent and probabilistic variables. BNs have been designed to derive LRs for half-sisters/half-sisters and mother/daughter/paternal grandmother relationships. These networks are validated against known formulae and are useful in other pedigree-related studies as well. These networks are applicable to motherless cases also (Hatsch et al. 2007). Loci DXS10079-DXS10074-DXS10075 and DXS6801-DXS6809-DXS6789 at X chromosome have been reported to provide stable haplotype blocks that can help to solve complex kinship cases (Castaneda et al. 2012). The usefulness of X-STRs has also been reported in deficiency paternity cases (Aquino et al. 2009). Increased Paternity Index (PI) and more reliable results have been reported in cases of deficiency paternity with the use of X-STRs rather than autosomal STRs (Trindade-Filho et al. 2013). In case of deficiency paternity, the mother of the putative father (Grandmother) can be used to establish the identity of the grandchild (Szibor et al. 2003). Genotypes of Chromosome X markers of the grandmother can also be reconstructed. If the grandmother has several daughters, the paternal origin of the most of the X alleles can be determined and so the X Chr genotype of the grandmother can be deciphered. If the brothers of the alleged father are present the data become even more informative, if the brothers carry different alleles it represents the grandmother is heterozygous on the contrary same alleles in brothers represent that the grandmother is homozygous (Szibor et al. 2003).

X-STR markers are advantageous in deficiency cases where samples from the paternal relatives are available for matching. Even in cases when none of the parents is present X STR can be of great importance, like in the case of two sisters or half-sisters. Though the proof of paternity in this case is less reliable as the sisters usually inherit partially matching haplotypes from the mother. Tillmar et al. (2011) used three different models to check the utility of X-STR markers in various sorts of relationship testing and found that if the mother is available, pedigrees like two maternal half-sisters or two females with common father can be better identified by the use of X-STRs and better likelihood ratios can be achieved (Tillmar et al. 2011). Investigator® Argus X-12 kit by Qiagen has been reported to be useful for sibship analysis as its likelihood ratios (LRs) were found to be relatively similar to the LRs

of autosomal STRs (Zidkova et al. 2014). Autosomal STR-based studies have been conducted on calculation likelihood ratios in sibship analysis (Musanovic et al. 2012; Tzeng et al. 2000; Wenk et al. 1996). But no studies pertaining to the LR in case of X-STR analysis have been published so far. Some complex paternity pedigrees like grandparent grandchildren, avuncular, and half-siblings cases are indistinguishable by unlinked autosomal STRs but have been reported to be distinguished by the use of linked X-STR markers (Pinto et al. 2011, 2012). X-STRs have been reported to complement the information given by the autosomal STRs in certain complex paternity cases (Krawczak 2007).

Maternity analysis by X-STR: Some situations may demand the mother/child testing, e.g., due to increased rates of illegitimate paternity it is better to check maternity while identifying unknown corpses or skeletonized remains. Maternity in mother–son duo can be best assessed by X-STR as the exclusion chances in such cases are similar to that of X-STRs in father–daughter cases, on the other side, in case of mother–daughter testing, the X-STRs are similar to autosomal STRs and do not provide any extra advantage. Maternity can also be assessed by mitochondrial DNA but the technique is comparatively expensive, requires more time and standardization, and doesn't always produce the level of certainty required in forensic (Parson et al. 1998). When population data for the individuals involved are not available, X-STR can be a better alternative for studying maternity.

15.4 Genetic Markers for Sex Identification

15.4.1 Amelogenin Marker

Amelogenin locus has the genes that code for the tooth enamel protein (von Wurmb-Schwark et al. 2007). AMELlocus has two homologous genes AMEL X and AMEL Y. AMEL X is located on the distal short arm of the chromosome X in the p22.1–p22.3 region. AMEL Y is located near the centromere of the Y chromosome at p11.2 (Nakahori et al. 1991).

Mutations: many cases of mutations in amelogenin gene have been reported (Kashyap et al. 2006; Santos et al. 1998; Steinlechner et al. 2002). Mutation in the primer binding site of AMEL X leading to amplification failure has been published lately (Maciejewska and Pawłowski 2009; Ou et al. 2012). Translocation in the SRY (sex-determining region Y) to X has led to the development of male reproductive organs and secondary sex characters in XX individuals (Ma et al. 2012; Steinlechner et al. 2002) in such cases amelogenin has failed to detect the sex of an individual.

15.4.2 DXYS156 Marker

It is a multialleleic pentanucleotide with the consensus sequence of motif (TAAAA). It is present on both X and Y chromosome and can be used for sex determination. DXYS156 is located on the long arm of the X chromosome and shows a sequence

homology with the short arm of the Y chromosome (Mukerjee et al. 2013). Allele of DXYS156 on Y chromosome can be distinguished from the alleles on X chromosome by detecting the presence of a Y-specific adenine insertion in the repeat unit (Calì et al. 2002). The polymorphism in the DXYS156 marker can be detected by using single primer pair amplifying both X and Y alleles due to high sequence similarity in the flanking region of the repeat unit of DXYS156 (Chen et al. 1994). There have been reports of DXYS156 polymorphism in almost all the studied populations. Y counterpart of the same shows more polymorphism than the X counterpart (Mukerjee et al. 2013). In case of Amelogenin test failure, DXYS156 locus has been used successfully for sex determination.

15.4.3 STS Marker

The STS gene codes for a membrane-bound enzyme named steroid sulfates, which hydrolyze steroid sulfates, a metabolic precursor for estrogen, androgen, and cholesterol (Reed et al. 2005). STS gene is located on the X chromosome at p22.31. Size of the gene is 146 kb. STS gene consists of 10 exons and 9 introns.

Mutations in the STS gene are associated with X-linked ichthyosis, a skin disease which affects the XY individuals (Alperin and Shapiro 1997; Shapiro et al. 1989), additional copies of this gene are reported in both male and female (Furrow et al. 2011).

A pseudogene of the same is reported to exist on Yq and named as STSP1 (Yen et al. 1988). The sequence homology between STS and STSP1 is 84.5%. Though STSP1 doesn't code for a functional protein. Primer sets targeting a 35 kb sequence in the first intron of STS and its homologs sequence in STSP1 produces two different kinds of amplicons that allow the identification of STS and STSP1 and determination of sex.

15.4.4 Population Studies on X STR Markers

Before STR markers are used for forensic casework, its allelic distribution in a particular population has to be known so that the allelic distribution in the population can be considered as standard against which casework analysis is done and to compute the value of match between the genetic profiles of two individuals (Diegoli 2015). The minimum number of unrelated individuals recommended to create a population database is 100–250 ("1991 report concerning recommendations of the DNA Commission of the International Society for Forensic Haemogenetics relating to the use of DNA polymorphisms" 1992; Schneider 2007). Population studies are also done to see the interrelatedness of various populations. Studies based on genetic relatedness among the individuals play an important role in many areas of genetics and forensics. Knowledge of relatedness is used to estimate genetic parameters such as heritability and genetic correlations in genetic analyses (Falconer and Mackay 1996). Estimation of breeding values in artificial selection relies on knowledge of

relatedness of individuals (Henderson 1984; Lynch and Walsh 1998). Knowledge of relatedness between interacting individuals is required to reconstruct the evolutionary trees and predict evolutionary consequences of social interaction in evolutionary biology (Hamilton 1964). In the context of anthropological research, relatedness between individuals is an important element of evolutionary investigations.

Around the world, studies have been performed on various populations to study the diversity in allelic distribution on X-STR markers to be used for forensic purpose (Afonso Costa et al. 2014; Li et al. 2014; Lim et al. 2009; Lin et al. 2011; Liu et al. 2013; Uchigasaki et al. 2013; Wu et al. 2014; Zeng et al. 2011; Zhang et al. 2011). A study on 16 and 12 X-STR markers has enlisted almost all the populations that have been studied for the frequency distribution of X-STR alleles (Gao et al. 2017; Guo 2017). From India, only one study has been published so far by Shrivastava et al. (2015) on the Bhil tribe of Madhya Pradesh (Shrivastava et al. 2015).

15.4.5 X Chromosome in Disease Diagnosis

Microsatellite instability (MSI), which occurs due to defective mismatch repair (MMR) pathway and loss of heterozygosity (LOH), has been corelated with various types of cancers and tumors (Bertagnolli et al. 2011; Buekers et al. 2000; Peris et al. 1995; Rizki and Lundblad 2001). Males being hemizygous for X chromosome are more prone to X chromosome-related disorders, but X linked disorders are not transmitted from male to male. Even the recessive phenotype gets exposed in

Table 15.6 Formulae for evaluating the forensic efficiency of genetic markers [Reprinted from Szibor et al. (2003)]

No.	Formula	References
1.	$\sum_i f_i^3(1-f_i)^2 + \sum_i f_i(1-f_i)^3 + \sum_{i<j} f_i f_j(f_i+f_j)(1-f_i-f_j)^2$	Ge et al. (1968)
2.	$\sum_i f_i^3(1-f_i)^2 + \sum_i f_i(1-f_i)^2 + \sum_{i<j} f_i f_j(f_i+f_j)(1-f_i-f_j)$	Kishida et al. (1997)
3.	$1-\sum_i f_i^2 + \sum_i f_i^4 - \left(\sum_{i<j} f_i^2\right)^2$	Desmarais et al. (1998)
4.	$1-2\sum_i f_i^2 + \sum_i f_i^3$	Desmarais et al. (1998)
5.	$1-2\left(\sum_i f_i^2\right)^2 + \sum_i f_i^4$	Desmarais et al. (1998)
6.	$1-\sum_i f_i^2$	Desmarais et al. (1998)

Formula 1: MEC for AS markers in trios
Formula 2: MEC FOR Chr X markers in trios involving daughter
Formula 3: MEC FOR Chr X markers in trios involving daughter (Desmarais version)
Formula 4: MEC FOR Chr X markers in father/daughter duos
Formula 5: Power of discrimination (PD) in females
Formula 6: PD for Chr X in males
$f_i(f_j)$: population frequency of the *i*th and *j*th marker allele

males because of hemizygosity. Around 100 genes on X chromosome have been correlated with mental retardation, which affects males more than females. Along with mental retardation, X-linked genes have been associated with epilepsy as well (Ross et al. 2005). Amir et al. (1999) reported that Rett syndrome in females was caused by mutation in MECP2 gene present on X chromosome (Amir et al. 1999), but later on this gene was found to be mutated in patients with mental retardation also (Orrico et al. 2000). Intrachromosomal rearrangements in X chromosome have been linked with multiorgan autoimmunity (Haltrich et al. 2015).

None of the STR markers including 13 CODIS loci and other autosomal STRs, X-STRs, Y-STRs, and mtDNA markers used for forensic identification are located in the exonic (coding) region of the DNA (Szibor et al. 2005a, b). Being present in the noncoding region of the DNA, forensic markers do not reveal any genetic abnormality of an individual but the numerical abnormalities have been reported to be revealed by these markers (Gornik et al. 2011). The coding loci like HumARA were not considered for forensic purposes because they revealed the private information of an individual though revelation of ethnic origin that was supposedly done by mt DNA and Y-STRs were not considered to be the private information of an individual (Szibor et al. 2005a, b).

15.4.6 Calculation of Forensic Parameters in X-STR Analysis

Polymorphic information content (PIC) and the expected heterozygosity (H_{exp}) are the two parameters that are valid for both AS and Chr X markers (Botstein et al. 1980). Formulas were made for MEC (mean exclusion chance) calculation in kinship testing (Table 15.6).

Formula 1 can also be used for Chr X STR markers while working on a deficiency paternity case in which paternal grandmother is investigated instead of alleged father. While MEC (calculated by formula 1) is compared with MEC (formula 2), the latter one is always higher showing the better efficiency of X-STR markers in paternity cases involving daughter (Szibor et al. 2003). Formula for the mean exclusion chance of Chr X markers in father–daughter duo (formula 4) was given by Desmarais et al. (1998); this formula can also be used in testing of mother/son duo. While formula 2 for MEC calculation is equal to formula 3 (Desmarais et al. 1998). The product rule is applied to calculate the forensic parameters in case of unlinked markers like autosomal STRs but since the markers present on X chromosome are linked the product rule cannot be applied but it has been reported that where linkage equilibrium is assumed the product rule can be applied for the calculations involving linked markers, e.g., in pedigrees like paternal half-sisters, parenthood, and paternal grandmother–granddaughter (Gill et al. 2012; Pinto et al. 2013).

15.4.7 MiniX-STRs

Forensic exhibits in cases of mass disasters are exposed to harsh environmental conditions like heat and humidity and hence their DNA undergoes severe degradation and the full profiles could not be generated. Mini-STRs are short-length amplicons that can be analyzed in very degraded DNA samples (Graham 2005; Israr et al. 2014). So far in forensic casework, autosomal mini-STRs have been used but since X STR analysis offers advantage over autosomal STRs in certain complex paternity cases and extended family analyses, miniX-STRs have also been developed to solve the complex kinship cases in situations like mass disasters (Israr et al. 2014).

15.4.8 Inclusion of X-STRs in the Latest Trending Technique, NGS

The utility of X-STRs in forensic analysis has rendered them a place in the latest technology, next-generation sequencing (NGS) also called the massive parallel sequencing (MPS). Though this technology has not made its place in regular forensic analysis but because of its alluring usages many forensic scientists are taking interest and it is being explored. Along with the use of STR markers, NGS technology supports the use of different kind of markers like single nucleotide polymorphisms (SNPs). The SNPs are comparatively less challenging to analyze and may also help in predicting the biogeographical ancestry of an unknown crime sample and external character like eye color, hair color of the sample donor (Borsuk 2016). Also, it has been reported that sample in very trace amounts and degraded states can be analyzed by NGS analysis, which is the main problem faced by every forensic analyst as the evidence samples are usually in degraded state and are present in very minute quantities (Alvarez-Cubero et al. 2017). Two biotech giants, Illumina (MiSeq) and Thermo Fisher Scientific (Ion Torrent Inc.), have developed panels for the forensic use of the NGS technology.

Platform for forensic use of NGS technology (Alvarez-Cubero et al. 2017)

Panel available for forensic analysis	Ion torrent (PGM)	Illumina (MiSeq)
Application	• Analysis of highly degraded or trace DNA • Missing person identification • Kinship analysis • More efficient mixture profiling	• Forensic profiling for DNA database • Missing person identification • Disaster victim identification
Kit	• HID-Ion AmpliSeq identity panel and • HID-ion AmpliSeq ancestry panel	• ForenSeq DNA signature prep kit
Markers		• *200 genetic markers*: 27 global autosomal STRs, 24 Y-STRs, *7*

(continued)

	• *124 autosomal markers*: 34 upper Y-clade SNPs and 90 autosomal SNPs, 165 autosomal markers	*X-STRs*, 95 identity SNPs, 22 phenotypic SNPs, and 56 biogeographical ancestry SNPs

Database: www.chrx-str.org website provides a database for chromosome X-STRs and chromosome X-STR haplotype. This database is comprised of published population data of X-STRs from various countries. The data contained in the database are published in peer-reviewed scientific journals. However, data of population of interest can also be submitted if the quality requirements are met. The genetic localization of the established X-STRs is displayed on the chromosome X linkage table and the ideogram of chromosome X ("www.chrx-str.org," n.d.). Another database available for X-STRs is the Brazilian Genetic Database of Chromosome X (BGBX—BancoGene'tico Brasileiro do Cromossomo X) website is freely available in Portuguese and English versions and was developed with the main purpose of compiling all Brazilian population genetic data for X chromosome short tandem repeat (XSTR) markers published in scientific journals searchable via PubMed. Furthermore, this database presents other relevant information concerning X-STRs, such as genetic and physical locations, allele structure, nomenclature, mutation rates, primers described in the literature, and likelihood ratio calculation. The entire scientific community is now encouraged to submit their X-STR population genetic data to this website, available at http://www.bgbx.com.br (Martins et al. 2014).

15.4.9 Softwares Available for STR Analysis

Numerous freely available software can be found online to calculate the likelihood ratios in various sorts of pedigrees. Softwares like Familias, FamLink, and FamLink X. These softwares use different computer programs and various sorts of pedigree analyses and likelihood ratio calculations for autosomal and X STR markers.

Familias: This software is commonly used for forensic cases to establish paternity, identify victims or analyze genetic evidence at crime scenes when kinship is involved and also for the difficult situations including inbred families, mutations, and missing data from degraded DNA. The software can be used for the purpose of identification in mass disasters, familial searching, non-autosomal marker analysis, and relationship inference using linked markers. Analysis by Familias focuses on more statistical issues such as estimation and uncertainty of model parameters. Although designed for use with human DNA, the principles can be applied to nonhuman genetics for animal pedigrees and/or analysis of plants for agriculture purposes (Egeland et al. 2015).

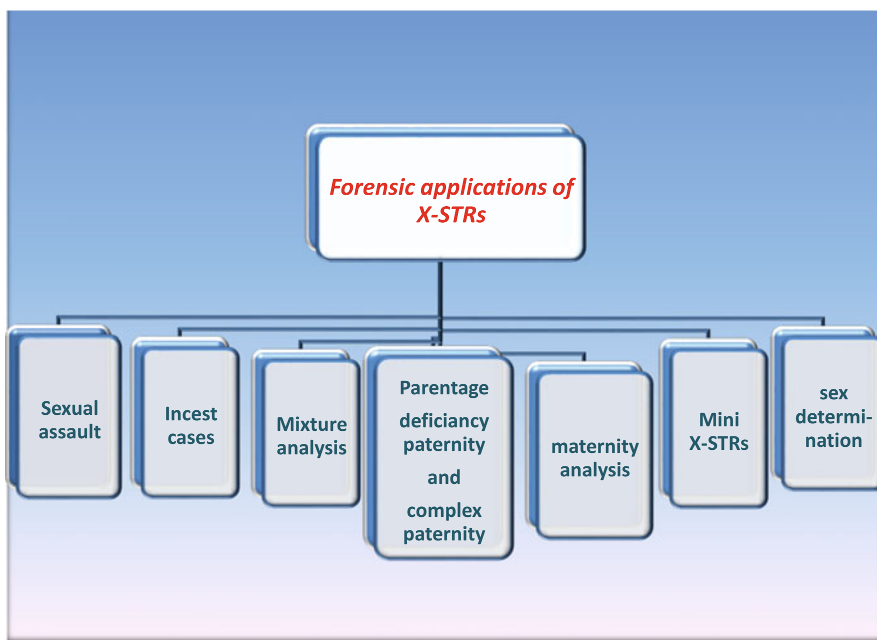
FamLink: this statistical software available freely online has a user-friendly interphase allowing the linkage calculations. It uses the Merlin computing program for pedigree analysis. Merlin performs fast haplotyping, detection of genotype errors, and pair linkage analyses. FamLink calculates the likelihood ratios from the

linked autosomal STR markers (Phillips et al. 2012; Tamura et al. 2015) data by overlooking the impact of linkage in pedigree analysis. FamLink software has the potential to analyze complex pedigrees as well (Kling et al. 2012).

FamLink X: it is a freeware available online for the calculation of likelihood ratios for X-STR data for different pedigrees like parentage, deficiency paternity, complex paternity, etc. While calculating the likelihood ratios, the software takes into account, linkage, linkage disequilibrium, and mutations as well. This software is based on Markov chain model to study the inheritance pattern, which allows to account for LD (Kling et al. 2015a, b). One step Markov chain is used to deal with linkage between the neighboring markers, i.e., linkage equilibrium and multistep Markov chain are used to manage linkage disequilibrium (Kling et al. 2015b).

15.5 Summary

X-STRs have been used to resolve a variety of forensic cases. The potential of X-STRs is summarized in the figure.



References

1991 Report concerning recommendations of the DNA Commission of the International Society for Forensic Haemogenetics relating to the use of DNA polymorphisms (1992) Forensic Science International, Ireland

Afonso Costa H, Morais P, da Silva C, Matos S, Marques Santos R, Espinheira R et al (2014) X-chromosome STR markers data in a Cabo Verde immigrant population of Lisboa. Mol Biol Rep 41(4):2559–2569. https://doi.org/10.1007/s11033-014-3114-9

Alperin ES, Shapiro LJ (1997) Characterization of point mutations in patients with X-linked ichthyosis effects on the structure and function of the steroid sulfatase protein. J Biol Chem 272(33):20756–20763

Alvarez-Cubero MJ, Saiz M, Martínez-García B, Sayalero SM, Entrala C, Lorente JA, Martinez-Gonzalez LJ (2017) Next generation sequencing: an application in forensic sciences? Ann Hum Biol 44(7):581–592

Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY (1999) Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. Nat Genet 23(2):185

Aquino J, Peixe C, Silva D, Tavares C, de Carvalho EF (2009) A X-chromosome STR hexaplex as a powerful tool in deficiency paternity cases. Forensic Sci Int Genet Suppl Ser 2(1):45–46

Bertagnolli MM, Redston M, Compton CC, Niedzwiecki D, Mayer RJ, Goldberg RM et al (2011) Microsatellite instability and loss of heterozygosity at chromosomal location 18q: prospective evaluation of biomarkers for stages II and III colon cancer—a study of CALGB 9581 and 89803. J Clin Oncol 29(23):3153–3162

Bohossian HB, Skaletsky H, Page DC (2000) Unexpectedly similar rates of nucleotide substitution found in male and female hominids. Nature 406(6796):622–625. https://doi.org/10.1038/35020557

Borsuk L (2016) Forensic applications of next generation sequencing. NIST, Gaithersburg

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32(3):314–331. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1686077&tool=pmcentrez&rendertype=abstract

Brinkmann B, Klintschar M, Neuhuber F, Hu J, Rolf B (1998) Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. Am J Hum Genet 62(6):1408–1415

Buckleton JS, Bright J-A, Taylor D (2016) Forensic DNA evidence interpretation. CRC Press, Boca Raton

Budowle B, Onorato AJ, Callaghan TF, Manna AD, Gross AM, Guerrieri RA et al (2009) Mixture interpretation: defining the relevant features for guidelines for the assessment of mixed DNA profiles in forensic casework. J Forensic Sci 54(4):810–821

Buekers TE, Lallas TA, Buller RE (2000) Xp22. 2-3 loss of heterozygosity is associated with germline BRCA1 mutation in ovarian cancer. Gynecol Oncol 76(3):418–422

Calì F, Forster P, Kersting C, Mirisola MG, D'Anna R, De Leo G, Romano V (2002) DXYS156: a multi-purpose short tandem repeat locus for determination of sex, paternal and maternal geographic origins and DNA fingerprinting. Int J Legal Med 116(3):133–138

Cann RL, Stoneking M, Wilson AC (1987) Mitochondrial DNA and human evolution. Nature 325:31. https://doi.org/10.1038/325031a0

Castaneda M, Mijares V, Riancho JA, Zarrabeitia MT (2012) Haplotypic blocks of X-linked STRs for forensic cases: study of recombination and mutation rates. J Forensic Sci 57(1):192–195. https://doi.org/10.1111/j.1556-4029.2011.01916.x

Chen H, Lowther W, Avramopoulos D, Antonarakis SE (1994) Homologous loci DXYS156X and DXYS156Y contain a polymorphic pentanucleotide repeat (TAAAA) n and map to human X and Y chromosomes. Hum Mutat 4(3):208–211

Chen D-P, Tseng C-P, Tsai S-H, Wang M-C, Lu S-C, Wu T-L et al (2009) Use of X-linked short tandem repeats loci to confirm mutations in parentage caseworks. Clin Chim Acta 408 (1–2):29–33

Dauber E-M, Kratzer A, Neuhuber F, Parson W, Klintschar M, Bär W, Mayr WR (2012) Germline mutations of STR-alleles include multi-step mutations as defined by sequencing of repeat and flanking regions. Forensic Sci Int Genet 6(3):381–386

Desmarais D, Zhong Y, Chakraborty R, Perreault C, Busque L (1998) Development of a highly polymorphic STR marker for identity testing purposes at the human androgen receptor gene (HUMARA). J Forensic Sci 43(5):1046–1049

Diegoli TM (2015) Forensic typing of short tandem repeat markers on the X and Y chromosomes. Forensic Sci Int Genet 18:140–151

Diegoli TM, Linacre A, Coble MD (2013) Forensic science international: genetics supplement series a gonosomal marker multiplex to aid in mixture interpretation. Forensic Sci Int Genet Suppl Ser 4(1):e184–e185. https://doi.org/10.1016/j.fsigss.2013.10.095

Dørum G, Kling D, Tillmar A, Vigeland MD, Egeland T (2016) Mixtures with relatives and linked markers. Int J Legal Med 130(3):621–634

Egeland T, Kling D, Mostad P (2015) Relationship inference with familias and R: statistical methods in forensic genetics. Academic Press, San Diego

El-Alfy SH, El-Hafez AFA (2012) Paternity testing and forensic DNA typing by multiplex STR analysis using ABI PRISM 310 genetic analyzer. J Genet Eng Biotechnol 10(1):101–112

Ellegren H (2000) Heterogeneous mutation processes in human microsatellite DNA sequences. Nat Genet 24(4):400

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. Nat Rev Genet 5 (6):435

Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics, 4th edn. Longman, Harlow

Fan H, Chu J-Y (2007) A brief review of short tandem repeat mutation. Genomics Proteomics Bioinformatics 5(1):7–14

Fukuta M, Gaballah MH, Kato H, Aoki Y (2018) A simple method for calculating the likelihood ratio in a kinship test using X-chromosomal markers incorporating linkage, linkage disequilibrium, and mutation. Legal Med 32:9. https://doi.org/10.1016/j.legalmed.2018.02.002

Fung WK, Hu Y (2004) Interpreting DNA mixtures with related contributors in subdivided populations. Scand J Stat 31(1):115–130

Fung WK, Hu Y-Q (2008) Statistical DNA forensics: theory, methods and computation. Wiley, Chichester

Furrow A, Theisen A, Velsher L, Bawle EV, Sastry S, Mendelsohn NJ et al (2011) Duplication of the STS region in males is a benign copy-number variant. Am J Med Genet A 155 (8):1972–1975

Gao HM, Wang C, Han SY, Sun SH, Xiao DJ, Wang YS et al (2017) Analysis of the 19 Y-STR and 16 X-STR loci system in the Han population of Shandong province, China. Genet Mol Res 16 (1)

Ge JKR, Ann WF, Lichte K, Steffens C (1968) Zur Verwendung des Polymorphismus der sauren Erythrocytenphosphatase bei der Vaterschaftsbegutachtung 146

Gill P, Phillips C, McGovern C, Bright J-A, Buckleton J (2012) An evaluation of potential allelic association between the STRs vWA and D12S391: implications in criminal casework and applications to short pedigrees. Forensic Sci Int Genet 6(4):477–486

Gornik KC, Grubić Z, Štingl K, Đurišević IT, Begović D (2011) Application of microsatellite loci on the chromosome X for rapid prenatal detection of the chromosome X numerical abnormalities. Croat Med J 52(3):392

Graham EAM (2005) Mini-STRs. Forensic Sci Med Pathol 1(1):65–67

Guo F (2017) Population genetic data for 12 X-STR loci in the northern Han Chinese and StatsX package as tools for population statistics on X-STR. Forensic Sci Int Genet 26:e1–e8. https://doi.org/10.1016/j.fsigen.2016.10.012

Gyllensten U, Wharton D, Wilson AC (1985) Maternal inheritance of mitochondrial DNA during backcrossing of two species of mice. J Hered 76(5):321–324

Haltrich I, Pikó H, Pamjav H, Somogyi A, Völgyi A, David D et al (2015) Complex X chromosome rearrangement associated with multiorgan autoimmunity. Mol Cytogenet 8(1):51

Hamilton WD (1964) The genetical evolution of social behaviour. J Theor Biol 7(1):17–52

Hatsch D, Keyser C, Hienne R, Bertrand L (2007) Resolving paternity relationships using X-chromosome STRs and Bayesian networks. J Forensic Sci 52(4):895–897. https://doi.org/10.1111/j.1556-4029.2007.00483.x

Henderson CR (1984) Applications of linear models in animal breeding. University of Guelph, Guelph

Hill CR, Butler JM, Vallone PM (2009) A 26plex autosomal STR assay to aid human identity testing. J Forensic Sci 54(5):1008–1015

Hu Y-Q, Fung WK (2005) Evaluation of DNA mixtures involving two pairs of relatives. Int J Legal Med 119(5):251–259

Israr M, Rafiq S, Shahid AA, Hussain H, Rakha A (2014) Scope of X-chromosomal MiniSTRs: current developments. J Forensic Res 5(5):1

Jaruzelska J, Zietkiewicz E, Batzer M, Cole DEC, Moisan JP, Scozzari R et al (1999) Spatial and temporal distribution of the neutral polymorphisms in the last ZFX intron: analysis of the haplotype structure and genealogy. Genetics 152(3):1091–1101

Jobling MA, Pandya A, Tyler-Smith C (1997) The Y chromosome in forensic analysis and paternity testing. Int J Legal Med 110(3):118–124

Jobling M, Hurles M, Tyler-Smith C (2013) Human evolutionary genetics: origins, peoples & disease. Garland Science, New York

Kaessmann H, Heißig F, Von Haeseler A, Pääbo S (1999) DNA sequence variation in a non-coding region of low recombination on the human X chromosome. Nat Genet 22(1):78–81

Kashyap VK, Sahoo S, Sitalaximi T, Trivedi R (2006) Deletions in the Y-derived amelogenin gene fragment in the Indian population. BMC Med Genet 7(1):37. https://doi.org/10.1186/1471-2350-7-37

Kayser M, Sajantila A (2001) Mutations at Y-STR loci: implications for paternity testing and forensic analysis. Forensic Sci Int 118:116–121

Kayser M, Roewer L, Hedman M, Henke L, Henke J, Brauer S et al (2000) Characteristics and frequency of germline mutations at microsatellite loci from the human Y chromosome, as revealed by direct observation in father/son pairs. Am J Hum Genet 66(5):1580–1588

Kishida T, Wang W, Fukuda M, Tamaki Y (1997) Duplex PCR of the Y-27H39 and HPRT loci with reference to Japanese population data on the HPRT locus. Jpn J Legal Med 51(2):67–69

Kling D, Egeland T, Tillmar AO (2012) FamLink–a user friendly software for linkage calculations in family genetics. Forensic Sci Int Genet 6(5):616–620

Kling D, Dell'Amico B, Tillmar AO (2015a) FamLinkX–implementation of a general model for likelihood computations for X-chromosomal marker data. Forensic Sci Int Genet 17:1–7

Kling D, Tillmar A, Egeland T, Mostad P (2015b) A general model for likelihood computations of genetic marker data accounting for linkage, linkage disequilibrium, and mutations. Int J Legal Med 129(5):943–954

Kong A, Gudbjartsson DF, Sainz J, Jonsdottir GM, Gudjonsson SA, Richardsson B et al (2002) A high-resolution recombination map of the human genome. Nat Genet 31(3):241–247. https://doi.org/10.1038/ng917

Koreth J, O'Leary JJ, O'D Mcgee J (1996) Microsatellites and PCR genomic analysis. J Pathol 178 (3):239–248

Kornberg A, Bertsch LL, Jackson JF, Khorana HG (1964) Enzymatic synthesis of deoxyribonucleic acid, XVI. Oligonucleotides as templates and the mechanism of their replication. Proc Natl Acad Sci U S A 51(2):315–323

Krawczak M (2007) Kinship testing with X-chromosomal markers: mathematical and statistical issues. Forensic Sci Int Genet 1(2):111–114

Lahn BT, Page DC (1999) Chromosome four evolutionary strata on the human X chromosome. Science 286(5441):964–967. https://doi.org/10.1126/science.286.5441.964
Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. Nature 409(6822):860–921. https://doi.org/10.1038/35057062
Li WH, Yi S, Makova K (2002) Male-driven evolution. Curr Opin Genet Dev 12(6):650–656. https://doi.org/10.1016/S0959-437X(02)00354-4
Li X, Zheng J, Yao J, Ding M, Pang H, Wang B (2014) Population data of 12 X-chromosome STR loci in Yanbian Korean samples from China. Forensic Sci Int Genet 13:143. https://doi.org/10.1016/j.fsigen.2014.07.012
Lim EJ, Lee HY, Sim JE, Yang WI, Shin K-J (2009) Genetic polymorphism and haplotype analysis of 4 tightly linked X-STR duos in Koreans. Croat Med J 50(3):305–312
Lin M, Liu QS, Huang YL, Che M, Li CQ, Liu YB et al (2011) Population genetics investigation of 12 X chromosome short tandem repeats in the She ethic minority in Fujian. China J Fam Plann 19:628–631
Liu YJ, Su WM, Yue JT, Shi SX, Guo LH (2013) Genetic polymorphisms of twelve X-chromosomal STR loci in Han population from Henan Province. J Forensic Med 29:297–299
Lucotte G, Guérin P, Hallé L, Loirat F, Hazout S (1989) Y chromosome DNA polymorphisms in two African populations. Am J Hum Genet 45(1):16–20
Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits, vol 1. Sinauer, Sunderland, MA
Lyon MF (1961) Gene action in the X-chromosome of the mouse (Mus musculus L.). Nature 190 (4773):372–373. https://doi.org/10.1038/190372a0
Ma Y, Kuang J-Z, Zhang J, Wang G-M, Wang Y-J, Jin W-M, Hou Y-P (2012) Y chromosome interstitial deletion induced Y-STR allele dropout in AMELY-negative individuals. Int J Legal Med 126(5):713–724
Maciejewska A, Pawłowski R (2009) A rare mutation in the primer binding region of the Amelogenin X homologue gene. Forensic Sci Int Genet 3(4):265–267
Makova KD, Li W-H (2002) Strong male-driven evolution of DNA sequences in humans and apes. Nature 416(6881):624–626. https://doi.org/10.1038/416624a
Martins JA, Kawamura B, Cardoso AE, Cicarelli RMB (2014) Brazilian genetic database of chromosome X. Mol Biol Rep 41(6):4077–4080. https://doi.org/10.1007/s11033-014-3277-4
Mukerjee S, Mukherjee M, Ghosh T, Kalpana D, Sharma AK (2013) Differential pattern of genetic variability at the DXYS156 locus on homologous regions of X and Y chromosomes in Indian population and its forensic implications. Int J Legal Med 127(1):1–6
Musanovic J, Filipovska-Musanovic M, Kovacevic L, Buljugic D, Dzehverovic M, Avdic J, Marjanovic D (2012) Determination of combined sibship indices “gray zone” using 15 STR loci for central Bosnian human population. Mol Biol Rep 39(5):5195–5200
Nadir E, Margalit H, Gallily T, Ben-Sasson SA (1996) Microsatellite spreading in the human genome: evolutionary mechanisms and structural implications. Proc Natl Acad Sci U S A 93 (13):6470–6475
Nakahori Y, Hamano K, Iwaya M, Nakagome Y (1991) Sex identification by polymerase chain reaction using X-Y homologous primer. Am J Med Genet 39(4):472–473. https://doi.org/10.1002/ajmg.1320390420
Nakamura Y, Samejima M, Tamaki K, Minaguchi K (2013) Multiplex PCR for 18 X-chromosomal STRs in Japanese population. Legal Med 15(3):164–170
Orrico A, Lam C-W, Galli L, Dotti MT, Hayek G, Tong S-F et al (2000) MECP2 mutation in male patients with non-specific X-linked mental retardation. FEBS Lett 481(3):285–288
Ou X, Chen W, Chen H, Zhao F, Zheng J, Tong D et al (2012) Null alleles of the X and Y chromosomal amelogenin gene in a Chinese population. Int J Legal Med 126(4):513–518
Parson W, Parsons TJ, Scheithauer R, Holland MM (1998) Population data for 101 Austrian Caucasian mitochondrial DNA d-loop sequences: application of mtDNA sequence analysis to a forensic case. Int J Legal Med 111(3):124–132

Peris K, Keller G, Chimenti S, Amantea A, Kerl H, Höfler H (1995) Microsatellite instability and loss of heterozygosity in melanoma. J Investig Dermatol 105(4):625–628

Phillips C, Ballard D, Gill P, Carracedo A, Lareu MV (2012) The recombination landscape around forensic STRs: accurate measurement of genetic distances between syntenic STR pairs using HapMap high density SNP data. Forensic Sci Int Genet 6(3):354–365

Pinto N, Gusmão L, Amorim A (2011) X-chromosome markers in kinship testing: a generalisation of the IBD approach identifying situations where their contribution is crucial. Forensic Sci Int Genet 5(1):27–32

Pinto N, Silva PV, Amorim A (2012) A general method to assess the utility of the X-chromosomal markers in kinship testing. Forensic Sci Int Genet 6(2):198–207

Pinto N, Gusmão L, Egeland T, Amorim A (2013) Paternity exclusion power: comparative behaviour of autosomal and X-chromosomal markers in standard and deficient cases with inbreeding. Forensic Sci Int Genet 7(2):290–295

Reed MJ, Purohit A, Woo LWL, Newman SP, Potter BVL (2005) Steroid sulfatase: molecular biology, regulation, and inhibition. Endocr Rev 26(2):171–202

Rizki A, Lundblad V (2001) Defects in mismatch repair promote telomerase-independent proliferation. Nature 411(6838):713

Ross MT, Grafham DV, Coffey AJ, Scherer S, McLay K, Muzny D et al (2005) The DNA sequence of the human X chromosome. Nature 434(7031):325

Santos FR, Pandya A, Tyler-Smith C (1998) Reliability of DNA-based sex tests. Nat Genet 18 (2):103

Schaffner SF (2004) The X chromosome in population genetics. Nat Rev Genet 5(1):43–51. https://doi.org/10.1038/nrg1247

Schneider PM (2007) Scientific standards for studies in forensic genetics. Forensic Sci Int 165 (2–3):238–243

Schwartz M, Vissing J (2002) Paternal inheritance of mitochondrial DNA. N Engl J Med 347 (8):576–580

Shapiro LJ, Yen P, Pomerantz D, Martin E, Rolewic L, Mohandas T (1989) Molecular studies of deletions at the human steroid sulfatase locus. Proc Natl Acad Sci U S A 86(21):8477–8481

Shrivastava P, Jain T, Gupta U, Trivedi VB (2015) Genetic polymorphism study on 12 X STR loci of investigator Argus X STR kit in Bhil tribal population of Madhya Pradesh, India. Legal Med 17(3):214–217. https://doi.org/10.1016/j.legalmed.2014.11.004

Steinlechner M, Berger B, Niederstätter H, Parson W (2002) Rare failures in the amelogenin sex test. Int J Legal Med 116(2):117–120

Subramanian S, Mishra RK, Singh L (2003) Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. Genome Biol 4(2):R13

Szibor R (2007) X-chromosomal markers: past, present and future. Forensic Sci Int Genet 1 (2):93–99

Szibor R, Krawczak M, Hering S, Edelmann J, Kuhlisch E, Krause D (2003) Use of X-linked markers for forensic purposes. Int J Legal Med 117(2):67–74. https://doi.org/10.1007/s00414-002-0352-5

Szibor R, Hering S, Edelmann J (2005a) The HumARA genotype is linked to spinal and bulbar muscular dystrophy and some further disease risks and should no longer be used as a DNA marker for forensic purposes. Int J Legal Med 119(3):179–180

Szibor R, Hering S, Kuhlisch E, Plate I, Demberger S, Krawczak M, Edelmann J (2005b) Haplotyping of STR cluster DXS6801-DXS6809-DXS6789 on Xq21 provides a powerful tool for kinship testing. Int J Legal Med 119(6):363–369. https://doi.org/10.1007/s00414-005-0550-z

Tamura T, Osawa M, Ochiai E, Suzuki T, Nakamura T (2015) Evaluation of advanced multiplex short tandem repeat systems in pairwise kinship analysis. Legal Med 17(5):320–325

Tillmar AO, Egeland T, Lindblom B, Holmlund G, Mostad P (2011) Using X-chromosomal markers in relationship testing: calculation of likelihood ratios taking both linkage and linkage disequilibrium into account. Forensic Sci Int Genet 5(5):506–511

Tomas C, Pereira V, Morling N (2012) Analysis of 12 X-STRs in Greenlanders, Danes and Somalis using Argus X-12. Int J Legal Med 126(1):121–128

Trindade-Filho A, Ferreira S, Oliveira SF (2013) Impact of a chromosome X STR Decaplex in deficiency paternity cases. Genet Mol Biol 36(4):507–510

Tzeng CH, Lyou JY, Chen YR, Hu HY, Lin JS, Wang SY, Lee JCI (2000) Determination of sibship by PCR-amplified short tandem repeat analysis in Taiwan. Transfusion 40(7):840–845

Uchigasaki S, Tie J, Takahashi D (2013) Genetic analysis of twelve X-chromosomal STRs in Japanese and Chinese populations. Mol Biol Rep 40(4):3193–3196

von Wurmb-Schwark N, Bosinski H, Ritz-Timme S (2007) What do the X and Y chromosomes tell us about sex and gender in forensic case analysis? J Forensic Legal Med 14(1):27–30. https://doi.org/10.1016/j.jcfm.2005.09.003

Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P et al (2002) Initial sequencing and comparative analysis of the mouse genome. Nature 420(6915):520–562. https://doi.org/10.1038/nature01262

Weber JL, Wong C (1993) Mutation of human short tandem repeats. Hum Mol Genet 2 (8):1123–1128

Wenk RE, Traver M, Chiafari FA (1996) Determination of sibship in any two persons. Transfusion 36(3):259–262

Wu XJ, Chen CQ, Jia X, Sun HM, Luo SM, Wei WX et al (2014) Genetic polymorphisms of 12 X-STR loci in Guangdong Zhuang population. Forensic Sci Technol 4:23–26

www.chrx-str.org (n.d.). http://www.chrx-str.org/

Xu H, Fu Y-X (2004) Estimating effective population size or mutation rate with microsatellites. Genetics 166(1):555–563

Yen PH, Marsh B, Allen E, Tsai SP, Ellison J, Connolly L et al (1988) The human X-linked steroid sulfatase gene and a Y-encoded pseudogene: evidence for an inversion of the Y chromosome during primate evolution. Cell 55(6):1123–1135

Zeng X, Ren Z, Chen J, Lv D, Tong D, Chen H, Sun H (2011) Genetic polymorphisms of twelve X-chromosomal STR loci in Chinese Han population from Guangdong Province. Forensic Sci Int Genet 5(4):e114–e116. https://doi.org/10.1016/j.fsigen.2011.03.005

Zhang S-H, Zhu R-X, Li L, Li C-T (2011) [Forensic application of investigator Argus X-12 kit in Han population from East China]. Fa Yi Xue Za Zhi 27(5):365–368. http://europepmc.org/abstract/MED/22259866

Zidkova A, Capek P, Horinek A, Coufalova P (2014) Investigator® Argus X-12 study on the population of Czech Republic: comparison of linked and unlinked X-STRs for kinship analysis. Electrophoresis 35(14):1989–1992

Ziętkiewicz E, Yotova V, Jarnik M, Korab-Laskowska M, Kidd KK, Modiano D et al (1997) Nuclear DNA diversity in worldwide distributed human populations. Gene 205(1–2):161–171. https://doi.org/10.1016/S0378-1119(97)00408-3

Ziętkiewicz E, Yotova V, Jarnik M, Korab-Laskowska M, Kidd KK, Modiano D et al (1998) Genetic structure of the ancestral population of modern humans. J Mol Evol 47(2):146–155. https://doi.org/10.1007/PL00006371

Applications of Mitochondrial DNA in Forensic Science

16

Moumita Sinha, Manisha Rana, and Pushpesh Kushwaha

Abstract

Mitochondria is an intracellular double-membraned organelle that exists inside each nucleated cell of mammals. Mitochondrial DNA (mtDNA) is a tiny 16.6 kilobase (kb) dsDNA circle coding 13 critical respirational chain components. The displacement loop (D-Loop) of mtDNA or control region is a noncoding region with 1.1 kb size that controls molecule transcription and replication. The D-Loop includes three short regions with a highly variable population sequence relative to the rest of the genome: a hypervariable HVS-I, HVS-II, and a HVS-III. mtDNA is strongly preserved, and it is recombined with what would be similar copies of itself when it is recombined. However, the mutation rate of mitochondrial DNA is tenfold greater than the DNA of the nucleus. This property makes the creation of species extremely useful in the matrilineal line for several generations and utilized in several forensic investigations. Therefore, mitochondrial DNA profiling is one of the most important tools in maternal dispute and child swapping cases.

Keywords

mtDNA · D-Loop · Profiling · Forensic · Recombination

M. Sinha (✉) · M. Rana
Department of Forensic Science, Guru Ghasidas University, Bilaspur, Chhattisgarh, India

P. Kushwaha
Dr. APJ Abdul Kalam Institute of Forensic Science and Criminology, Bundelkhand University, Jhansi, Uttar Pradesh, India

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_16

16.1 Introduction

16.1.1 Mitochondria

Mitochondria's are intracellular double-membrane organelles that exist inside each nucleated cell of mammals. On an organizational level, the established picture that mitochondria are a bean-shaped or cigar-shaped small structure is possibly ingenuous, and it is more precise to consider mitochondria as an endoplasmic reticulum-like budding and fusing network (Iborra et al. 2004). Mitochondria are closely enmeshed in homeostasis on the cellular. They play a role among other functions in intracellular signaling and apoptosis, intermediary metabolism, and in the metabolism of lipids, amino acids, steroids, cholesterol, and nucleotides. Likewise, mitochondria also perform a vital role in the metabolism of cellular energy. This process involves the development of carboxylic acid β-oxidation, the urea process, and the definitive important nucleotide pathway (ATP)—the respirational chain. The mitochondrial respirational chain is a combination of complexes of five enzymes located on the membrane inside mitochondria. Each of the complexes comprises of several subunits, the most important being complex I consisting of over 40 components of polypeptides. NADH and FADH2 are reduced cofactors produced by the intermediate break down of carbohydrates, fats, and proteins giving complex I and II electrons. Such electrons migrate downwards and are electrochemical gradients between complexes III and IV and two mobile electron carriers, ubiquinone and cytochrome. The e-transfer purpose of complexes I–IV is achieved through subunits harboring prosthetic groups (e.g. iron-sulfur groups in complexes I, II, and III and heme iron in cytochrome c and in complex IV). Mitochondrial subunits containing prosthetic groups, i.e., heme iron in cytochrome (c) and complex IV), sulfur iron group in complexes I, II, and III and the electron transference function of complexes I–IV is carried out. Complexes I, III, and IV use the released energy to transmission protons from the mitochondrial matrix into the vacuum between the intermembrane. This proton gradient is connected by complex V to synthesize ATP from adenosine diphosphate (ADP) and inorganic phosphate, which produces the bulk of the mitochondrial membrane potentials. Oxidative phosphorylation (OXPHOS) is the basic method. ATP is a high-energy source used by virtually all-important metabolic processes in the cell. In exchange for cytosolic ADP, it must be released from the mitochondrion. Adenine nucleotide (ANT), which comprises a

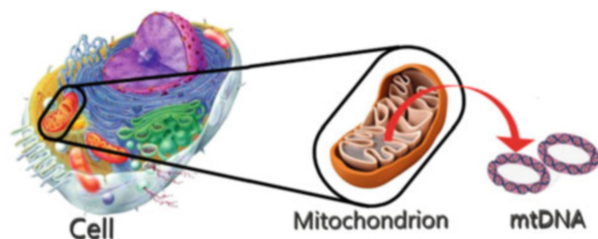


Fig. 16.1 Location of mtDNA within the cell

variety of tissue exclusive isoforms, is used for this purpose (Bandelt et al. 2006) (Fig. 16.1).

16.1.2 Mt Genome (Location and Structure)

Mitochondrial DNA is a tiny 16.6 kb dsDNA circle coding 13 critical respirational chain components. Seven subunits of complex 1 are encoded by ND1-ND6 and ND4 L. The cytb-compound (ubiquinol cytochrome c reductase) is a single mitochondrial DNA-encrypted III complex subunit. CO (I)-CO (III) encrypt3 complex IV subunits (cytochrome c oxidase) with two complex V-subunits (ATP synthesis) encoding for ATP6 and ATP8 genes. Besides, mtDNA codes for two ribosomal RNAs (rRNAs) genes and 22 genes of transfer RNAs (tRNAs), which provide RNA components for intra-mitochondrial protein synthesis.

The displacement loop (D-Loop) of mtDNA or control region is a noncoding region with 1.1 kb size that controls molecule transcription and replication. D-loop stretches from 576 nt to16,024 nt in mtDNA and is the major region not influenced by the production of polypeptides in the chain. There are a variety of small and single-base genome segments that don't engage in RNA or protein coding directly. The D-Loop includes three short regions with a highly variable population sequence relative to the rest of the genome: a hypervariable HVS-I, a HVS-II, and a HVS-III, corresponding in other references (Brandstätter et al. 2004a). The precise description of the various hypervariable sequences differs between contexts. The forensic group historically took HVS-I in 16,024–16,365, HVS-II in 73–340, and HVS-III in 438–576 (Brandstätter et al. 2004b). New population genetic studies, on the other hand (Brandstätter et al. 2004b). To depict the phylogenetically significant positions 16,390, 16,391, and 16,399 (HVS-I 16,024–16,400, HVS-II 44–340, and HVS-III 438–576), take extensive ranges, in particular in HVS-I. The full purpose of these three hypervariable regions is not understood, but they tend to be essential for replication of genome and transcriptions and either one contain or close to the source of mtDNA replication by heavy strand and light strand: O_H and O_L. The genetic code of mtDNA varies in many ways from the worldwide genetic code (Anderson et al. 1981). In human mtDNA, AUG codon codes for methionine, but not for isoleucine, and the termination codons, i.e., AGA and AGG codes for Arginine, and tryptophan is coded with UGA unlike nuclear DNA. Ultimately, the codons of initiation are AUA and AUG. Early trials established the molecular asymmetric stranding with a "weighty " strand rich in guanosine or "H-Strand" and "light" Strandor" L-strand rich in cytosine. In the custom, mtDNA is tote up according to the original CRS—Cambridge Reference Sequence of mtDNA in reference to the light strand. The revised Cambridge Reference Sequence as the novel light strand system of numbering should be preserved for reserach purposes (Andrews et al. 1999) (Fig. 16.2).

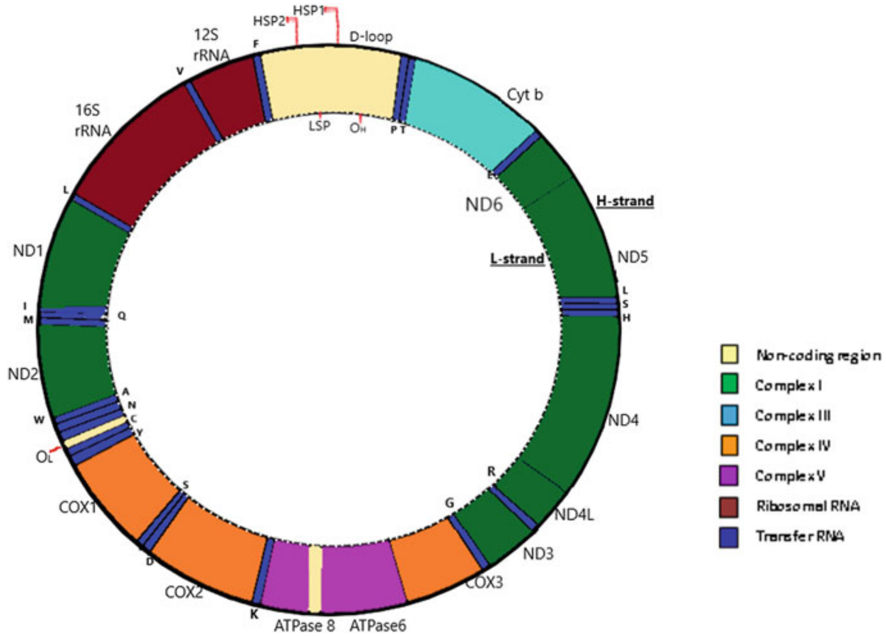


Fig. 16.2 Mitochondrial genome

16.2 Transcription, Translation, and Replications in mtDNA

The human mitochondrial DNA was revealed in the 1980s as the fundamental mechanism for transcription (Clayton 1982). Two main sites are launched for transcript: ITH1 and ITL, both in the noncoding region at 150 bp. Strong beach transcription begins at position 561 in a strong beam promoter; light beach transcriptions start at position 407 inside a light beach, along with elements of enhancer right upstream binding the mitochondrial transcript factor A (TFAM or mtTFA). TFAM allows a two-way version, isolated from the DNA template.

In stoichiometry, theBM1 and BM2 transcription factors join 1:1 to the heart and perform a critical part when the transcription begins (Falkenberg et al. 2002). The beam is transliterated as a solo polycistronic precursor that is later sort out and improved until the proteins encoded by mtDNA are synthesized on mitochondrial ribosomes in the mitochondrial matrix. While the fundamental processes following this phase are familiar, the monitoring processes are merely explained (Gagliardi et al. 2004).

In comparison to nuclear DNA that only repeats one time in the course of each cell cycle, mt DNA is processed constantly in non-tissues, for instance, the muscle of the skeleton and the brain (Bogenhagen and Clayton 1977; Birky 2001). Consequently, mitochondrial DNA is self-contained of cell cycle replication. The correct

replication process of mtDNA is asymmetrical to the strand. The heavy strand leads to a replication process starting at OH according to this model (Clayton 1982) with rift of an initial transcript synthesized with the light beam promoter. As the heavy strand reproduction reaches the origin of replication for the light strand, the light strand is synthesized in the opposite direction. Current experimental data does support a substitutional symmetric or rolling-circulation strand model, which starts to replicate MtDNA in a crucial 5.5 kb region flanked by the D-Loop and the ND4 gene at various points (Bowmaker et al. 2003). Such duplication bubbles then continue bidirectionally, ending at O_H and stopping succinctly in the O_L region, until the replication process is completed, and the rear strand rates the binding of Okazaki fragments.

16.3 Pathogenic Mutations of mtDNA-Heteroplasmy

Earlier studies in 1980s, reported the pathogenic mutations in mitochondrial DNA. The cases with long-lasting, developing, outward ophthalmoplegia and (KSS) Kearns Sayre Syndrome (Holt et al. 1988; Zeviani et al. 1988) large-scale deletions of mitochondrial DNA, together with deleting tRNAs and functional genes, occurred in the muscles of skeleton. In a case suffering from mitochondrial encephalomyopathy along with stroke (Goto et al. 1990) as well as Leber Inherited Optic Neuropathy (LION), point mutations of mitochondrial DNA were identified (Wallace 1997).

The prevalence of initial pathogenic mitochondrial DNA mutations in affected individuals remain heteroplasmic by means of different quantities of mutated mitochondrial DNA in the similar organism. While the majority of humans hold two copies of nuclear DNA, they have much extra copies of mtDNA (depending on the cell type from 1000 to 100,000). All of these exist frequently same in salubrious birth persons (homoplasmas), but a certain mixture (heteroplasmas) can occur, particularly involving polycytosine stretch hypervariable sites and polymorphs in duration. In reality, the altered and the wild type mitochondrial DNA often interact in patients who have pathogenic mtDNA defects (Holt et al. 1988; Wallace 1997). There is a major difference between different cases and in tissues of various organs and also within the cells of the same person in the percentage of mutated mtDNA. The bulk of mutated mitochondrial DNA mutations remain strongly recessive in in vitro experiments using the "trans-mitochondrial cytoplasmic hybrid" cells (King and Attardi 1988). In other words, before a biochemical fault within the respirational chain was created, cells could tolerate high level of mutilated mtDNA (typically 70–90%). The emergence of Heteroplasmy and sequencing of D-Loop suggests a recent transformation at the same site from population genetics perspective (Man et al. 2003).

16.4 Inheritance of mtDNA

It is common to say, but recently this has been called into question, that mitochondrial DNA is transmitted only along the matrilineal lines (Giles et al. 1980). The problem of paternal transmission was posed by a surplus of homoplasms and the apparent inversion of the intermolecular mtDNA rebound rate with the genetic gap (Awadalla et al. 1999), and by the fact that paternalistic mtDNA is associated with the analyses of individual development (Awadalla et al. 1999).

Post Studies after above findings created questions on the explanations given. The excessive homoplasmy identified in earlier reports is clarified in different recombinant measures in various data sets (Mcaulay et al. 1999; Elson et al. 2001) along with sequencing artifacts (Hagelberg et al. 1999). The latest study of the deletion of paternal small pathogenic mtDNA found that paternal mtDNA leakage can occur (Schwartz and Vissing 2002).

Indeed, in the past decade several substantial number of families with mtDNA disease were researched around the world and other cases of paternal transmission have not been reported. Therefore, there is little evidence that mtDNAs are recovered in vivo (Howell 1997). Latest research of Kraytsberg and his colleagues has provided an exciting proof of recombination in vivo (Kraytsberg et al. 2004), which, however, is also used to create entities that may perhaps yield confusing outcomes (Bandelt et al. 2006).

In addition, the available evidence suggests that paternal mtDNA leaking is extremely rare, and while it does occur, a major recombination of paternal and maternal mtDNAs is highly unlikely. So there seems to be no reason to question, nevertheless as of the population genetics perspective, the conventional doctrine of maternal transmission.

The findings are still vital on the way to explain the exact molecular mechanisms in the rear of strict motherly transmission. Although paternal mitochondria were originally thought to not have entered the oocyte, this is not the case. In a study, the recently implanted embryo in uterus have found paternal mtDNA molecules in them (St John et al. 2000). The father's mitochondrial DNA is likely to be killed by an effective ubiquitis process (Sutovsky et al. 1999).

16.5 Forensic Applications of mtDNA

Mitochondria exists as a key organelle intended for the manufacture of ATP in the form of energy for cells. Human beings contain DNA in mitochondria as mtDNA, which is around 16,500 base pairs as well as individual-specific. This makes mtDNA useful when DNA is compromised or degraded in forensic science. Human mtDNA is highly preserved, and it is recombined with similar copies of itself when recombine. However, the mutation rate of mitochondrial DNA is tenfold greater than DNA of nucleus. This property makes the creation of species extremely useful in the matrilineal line for several generations. Some of the most important applications of mtDNA in forensic science are as follows:

1. Identification of largely decayed bodies where only the teeth, bones, or hair is available.
2. Maternal dispute cases.
3. Baby swapping.
4. Identification of ancient DNA.
5. Phylogenetic studies.

16.6 Mitochondrial DNA Sequencing Approaches

The first DNA sequence research (Sanger et al. 1977), also regarded as the first-generation sequences, was published by Sanger in 1977. The incorporation of ddNTPs into the freshly synthesized DNA molecules contributes to the completion of the stretching cycle and subsequent awareness of the individual nucleotide in each location in the chain. The method of Sanger sequencing can generate readings of 25–1200 nucleotides, which allow reading of up to 96 kb of nucleotides within 2 h. Since 2005, modern methods of DNA sequencing were developed (Bruijns et al. 2018), often referred to as next-generation sequencing (NGS).

The synthesis-sequence approaches, for instance Roches "454" pyrosequencing approach and Illuminas "HiSeq," enable sequencing within 1–2 week, capable of analyzing 80 million bp in 2 h, and able to analyze 6 billion base pairs. Sequencing through hybridization and ligation, for instance-ABI SOLiD 3plus, results in 60 gb per exercise of functional deoxyribose nucleic acid details. The sequenced DNA fragments will vary from 35 to 75 nt while in the analysis with SOLiD with 453 pyrosequencing and major parallel sequencing (MPS) technologies (Bruijns et al. 2018; Ondov et al. 2010; Shendure et al. 2005) the fragments are 100–1000 nts long. (Bruijns et al. 2018; Dames et al. 2010). Personal Genome Machine™ (PGM) via Ion Torrent utilizes a pH-based identification technique by introducing a radio-active material to a series (Rothberg et al. 2011). Once this occurs, protons are emancipated and a proportional electrical signal is produced. Data processing is performed via the CMOS (Complementary Metal Oxide Semiconductor) capable of calculating millions to billions instantaneous sensing responses, with a sensor surface on the bottom of the well layer. Finally, a portable real-time sequence device (MION) (Oxford Nanopore Technologies) enables extremely long reading lengths in hundered kb. In forensics the NGS tools were easily used (Bruijns et al. 2018). In the form of forensic sampling (Parson et al. 2013) and heteroplasm experiments (Magalhães et al. 2015), for example, Ion Torrent's PGM device was used. Though PGM has proven resilient and accurate to detect and measure mixture and heteroplasm, the reporting of the mtDNA in certain regions with significant feature biases and fake positive results mainly triggered by problems of alignment in the research algorithms has been troublesome. In recent years, the accurate identification of mitochondrial DNA Complete Genome Panel (Woerner et al. 2018) was used with S5 Ion Technique (Thermo Fisher Scientific) and Desktop MiSeqFGx Sequencer (Ilumina). Both structures provide consistent mtDNA haplotypes esti-mate. Some additional findings on the usage of forensic molecular genetics

technology and mitochondrial DNA examination have been circulated (Templeton et al. 2013; Just et al. 2014; Chaitanya et al. 2015; Lopopolo et al. 2016; Ovchinnikov et al. 2016; Hollard et al. 2017; Marshall et al. 2017; Park et al. 2017; Young et al. 2017; Churchill et al. 2018; Ma et al. 2018).

Nonetheless, more feasibility studies and advanced forensic-based software capabilities need to be created to allow NGS analysis to be integrated into traditional legal purposes (Amorim and Pinto 2018; Peck et al. 2016). Although, Sanger sequence appears as an effective tool for mtDNA research for the forensic detection of individuals, and is used in most cases work laboratories around the world, in keeping with existing universal recommendations (Parson et al. 2014; Prinz et al. 2007; Ballard 2016). A number of forensic science laboratories conduct the sequencing using Sanger method of the whole control zone as a standard technique (Chaitanya et al. 2016; Poletto et al. 2019; Turchi et al. 2016; Yasmin et al. 2017), although others have also expanded the analysis to HVIII and, in modern years, nearly all forensic science laboratories applies amplifications of the whole control region in standard technique. Over the last decade, several experiments have been based on applying research of the whole mtDNA genomes in an effort of enhance the role of mtDNA in human identifying (Duan et al. 2019; Strobl et al. 2018; Woerner et al. 2018).

Amorim and Pinto (2018), contribute details as of the complete genome of mtDNA to enhance the identity obtained by the HVI, HVII and HVII analysis or from the complete control region without modifying the previous findings on separate haplogroups with an entirely different regional ancestry.

16.7 Current Trends of mtDNA Analysis: Hybridization Enrichment Coupled with Massively Parallel Sequencing (MPS)

Control region sequencing using Sanger method possibly will be a concern in forensic examination of mitochondrial DNA of disintegrated human remains and with PCR-based Sanger sequencing the obtained DNA evidences are incredibly small. Enrichment of hybridization coupled with next-generation sequencing –massively parallel sequencing (MPS) provides an efficient way to retrieve fragments from poorly preserved samples of DNA as minute as 30 base pairs (bp). Cases where human identifications are required from extremely disintegrated and fragmented DNAs (< 100 bp) are mostly solved by analysis of mtDNA, and also in cases where persons missing from longer time, immigrant problems, Disaster Victim Identification, etc. Also, with poor amplification of PCR has been observed due to high rates of fragmented DNA, i.e. which are below required size of amplicon, requires hybridization for better analysis (Gilbert et al. 2003; Maciejewska et al. 2013; Chaitanya et al. 2015).

The performance of mtDNAPCR is linked to the input of DNA, which indicates that low DNA quantity PCR-based approaches have failed (Just et al. 2015). Targeted hybridization enhancement (Templeton et al. 2013; Hofreiter et al. 2015)

coupled with MPS provides a range of advantages in contrast to the PCR-based analytical approaches for decomposed DNA. They are capable of obtaining sequence facts from pieces of DNA as petite as 30 base pair, fairly beneath the minimum PCR level and identifying single fragments of DNA, thereby reducing PCR-related problems.

Enrichment of hybridizing involves conversion by bar-coded adjusters of disintegrated genomic DNA to the DNA lending library. The PCR primers are used to immortalize the DNA, complementing the adapters, until enrichment is hybridized. Biotinylated ssRNA or DNA hybridization dressings are consequently utilized so as to separate from the library the relevant sequences (in this case human mtDNA).

Various forms of DNA like well preserved DNA, decomposed DNA, chemically treated DNA, telogen-rested blood, skeletal remain traces, and archaeological specimens have been examined and used in forensic analysis (Templeton et al. 2013; Marshall et al. 2017; Shih et al. 2018). (MDNA) To enhance and sequence of the mitochondrial DNA (mtDNA) control regions, Eduardoff et al. (2017) used primary extension capture (PEC) from superior quality DNA, human hairs samples and ancient human skeletal remains.

Hybridization enrichment can therefore be said to have a new ability and capacity to distinguish traces and highly damaged DNA based on expert DNA. In this latest strategy, trace sources will be enhanced, decomposed human remains are more accurately detected, and uncertain sample detection costs and delay are reduced, thereby improving criminal and coronal investigative findings (Fig. 16.3).

16.8 Mitochondrial DNA Population Data and Databases

If two mitochondrial DNA sequences, one commencing from a sample proof and one as of a reference lending library, cannot be ruled out as they come from precisely the similar resource, some detail on the rarities of the mtDNA profile should be given. The traditional method is to calculate how many times in a population sample a given order is found (Budowle et al. 1999). In general, the population databanks utilized in forensics include numerous convenient specimens which reflect, in terms of facts, the main populations of possible contributors. The EDNAP Mitochondrial DNA Population Database is the most important mtDNA Haplotype database (EMPOP, http://www.empop.org).

EMPOP was developed and planned in its early stages to function like a population reference databank exclusively for assessment of mitochondrial DNA proof throughout the world, with the purpose to endow mtDNA data with the superior quality. The design and research capabilities of this online database have developed over the past few years, while mtDNA data consistency remains the key objective of the EMPOP database. EMPOP is not just a population reference resource but also a direct result intended for forensic genetics researchers as well as other disciplines. It also acts as a quality control device. EMPOP is the most robust dataset, especially from the point of view of community covered by this database, although there is a

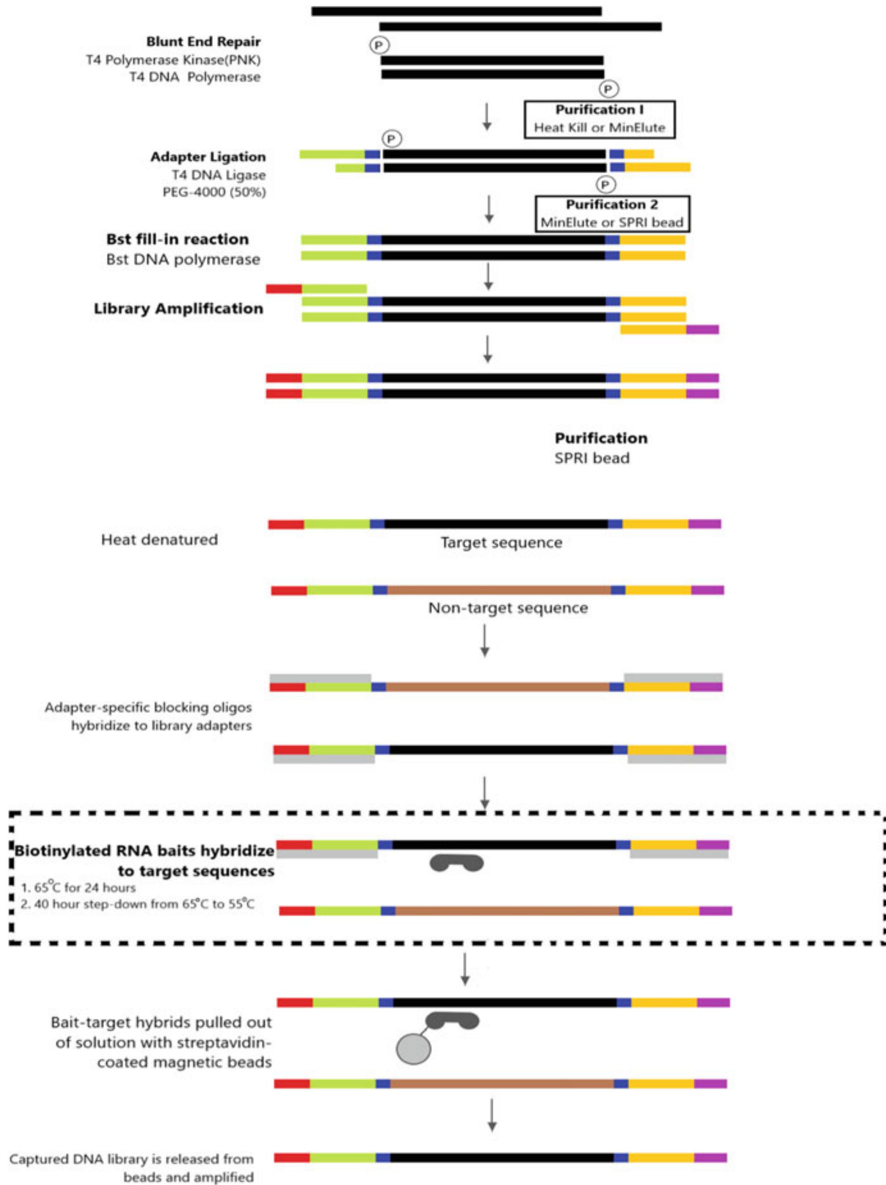


Fig. 16.3 Steps of massively parallel sequencing (MPS) with coupling of hybridization enrichment

significant amount of standard reference populations meant for the purpose of forensic comparisons (Parson et al. 2014). EMPOP utilizes a sequence-based SAM search algorithm that transforms enquiry and database arrangements on nuclear sequences that are alignment safe, thus ensuring that a haplotype is identified in an

unchecked database application. SAM-E is already in operation. It is an improved SAM variant and uses InDels block as phylogenetic cases. In EMPOP the tool haplogroup explorer represents in convenient searchable formats all known Phylotree Haplogroups and offers the number of the EMPOP sequence allocated toward the respective category by means of the maximum probability method EMMA to estimate mitochondrial DNA Haplogroups (Röck et al. 2013). As described earlier, more current common ancestor (MRCA) haplogroups are valid for many potential haplogroups. PhyloTree offers a revised full phylogenetics of a global human variant of mtDNA centered on mutations in the coding regions and control regions (Van Oven and Kayser 2008).

The entire mitochondrial DNA phylogenetic tree is accessible at http://www.phylotree.org, which contains both recently documented populations and newly found haplogroups. EMPOP offers spatial haplogroup variations to imagine and appreciate the regional distribution of the graphs. Mitomap (Ruiz-Pesini et al. 2007) is another significant mtDNA human database.

This database was created in 1996 as an electronic database (http://www.mitomap.org) comprising reported differences of human mtDNA, along with different variants of geography and disease. Today, Mitomap is manually designed with the aid of superior quality human mitochondrial DNA used for medical professionals, researchers, and molecular biologists (Ruiz-Pesini et al. 2007) and modified to technically rich tools. Mitomap has three primary use groups. This provides certain relevant details on human mtDNA, including the overall distribution of mitochondrial DNA haplogroups and their frequency.

In addition, consumers can include details on certain repositories, applications, and helpful services unique to mtDNA. Mitomap shops all healthy people and doctors to annotate mtDNA variations. The mutation frequency is determined from the GenBank's human mitogenomes. Thus, users can access the details and save the most significant information on loci, the nucleotide shift, codon location, and number in various file sizes. Mitomap includes the computational method Mitomaster, which actually offers the computer programming interface.

The key purpose of this method is to classify polymorphic locations, to quantify variant statistics and to assign complete or partial mitogenomes to haplogroups. Such a question may be achieved by repeating to mtDNA, GenBank, or single nucleotides (Brandon et al. 2009). Such a question will not be feasible. Throughout the application of mtDNA databases, ethical and legal concerns can emerge from another viewpoint. Study of mtDNA may produce knowledge regarding privacy concerns (Guillen et al. 2000; Wallace et al. 2014).

Main mitochondrial disorder is a mutation that can be transmitted in most men. Consequently, knowledge regarding the structure of the mitochondrial genome will classify a person's current or potential safety. The study of MtDNA can only be conducted in noncoding regions, where no illness or phenotypical details has been identified with it.

16.9 Conclusion

For the past 25 years, mitochondrial DNA typing is used internationally to address various human identity problems linked to organized crime, small offences, criminal attacks, major accidents, and incidents of disappearances. Advances in mtDNA typing are incredible, varying from analyzing tiny fragments within days to decoding many entire mtDNA genomes in a few hours. The mtDNA genome may include details on ancestors, including knowledge on health/disease, as a genetic marker of lineage. While there are many widely recognized explanations for researchers to collect knowledge on the possible history of a suspected perpetrator, others nevertheless consider it unethical to establish the genetic structure behind such diseases. Therefore, emerging technologies that enable the sequencing of mitochondrial DNA should be wisely utilized in the field of forensic identification and for the purposes expected. These remain issues with respect, in particular, to the topic of heteroplasmy and in recent times with the flexibility of biparental heirlooms, the allowability of mtDNA research in court. Clear explanation of biparental ancestry molecular pathways for mtDNA, the capacity to evaluate the conditions when this happens, and the potential to extremely reliably diagnose and classify the heteroplasmy are big challenges that need to be dealt with in order to render mtDNA a viable alternate resource in forensic identification of humans.

References

Amorim A, Pinto N (2018) Bigdata in forensic genetics. Forensic Sci Int Genet 37:102–105

Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, Young IG (1981) Sequence and organization of the human mitochondrial genome. Nature 290:457–465

Andrews RM, Kubacka I, Chinnery PF, Turnbull DM, Lightowlers RN, Howell N (1999) Reanalysis and revision of the Cambridge reference sequence. Nat Genet 23:147

Awadalla P, Eyre-Walker A, Maynard Smith J (1999) Linkage disequilibrium and recombination in hominid mitochondrial DNA. Nature 286(5449):2524–2525

Ballard D (2016) Analysis of mitochondrial control region using sanger sequencing. Human Press, New York, pp 143–115

Bandelt HJ, Richards M, Macaulay V (eds) (2006) Human mitochondrial DNA and the evolution of Homo sapiens, vol 18. Springer-Verlag, Berlin Heidelberg

Birky CW (2001) The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. Annu Rev Genet 35:125–148

Bogenhagen DF, Clayton DA (1977) Mouse L cell mitochondrial DNA molecules are selected randomly for replication throughout the cell cycle. Cell 11:719–727

Bowmaker M, Yang MY, Yasukawa T, Reyes A, Jacobs HT, Huberman JA, Holt IJ (2003) Mammalian mitochondrial DNA replicates bidirectionally from an initiation zone. J Biol Chem 278:50961–50969

Brandon MC, Ruiz-pesini E, Mishmar D, Procaccio V, Marie T, Nguyen KC, Douglas C (2009) MITOMASTER—abio informatics tool for the analysis of mitochondrial DNA sequences. Hum Mutat 30(1):1–6

Brandstätter A, Niederstätter H, Parson W (2004a) Monitoring the inheritance of heteroplasmy by computer-assisted detection of mixed basecalls in the entire human mitochondrial DNA control region. Int J Legal Med 118:47–54

Brandstätter A, Peterson CT, Irwin JA et al (2004b) Mitochondrial DNA control region sequences from Nairobi (Kenya): inferring phylogenetic parameters for the establishment of a forensic database. Int J Legal Med 118:294–306

Bruijns B, Tiggelaar RM, Gardeniers HJGE (2018) Massively parallel sequencing techniques for forensics: a review. Electrophoresis 39(29):2642–2654

Budowle B, Wilson MR, DiZinno JA, Stauffer C, Fasano MA, Holland MM, Monson KL (1999) Mitochondrial DNA regions HVI and HVII population data. Forensic Sci Int 103(1):23–35

Chaitanya L, Ralf A, Van Oven M, Kupiec T, Chang J, Lagacé R, Kayser M (2015) Simultaneous whole mitochondrial genome sequencing with short overlapping amplicons suitable for degraded DNA using the ion torrent personal genome machine. Hum Mutat 36(12):1236–1247

Chaitanya L, Van Oven M, Brauer S, Zimmermann B, Huber G, Xavier C, Kayser M (2016) High-quality mtDNA control region sequences from 680 individuals sampled across the Netherlands to establish a national forensic mtDNA reference database. Forensic Sci Int Genet 21:158–167

Churchill JD, Stoljarova M, King JL, Budowle B (2018) Massively parallel sequencing enabled mixture analysis of mitochondrial DNA samples. Int J Legal Med 132(5):1263–1272

Clayton DA (1982) Replication of animal mitochondrial DNA. Cell 28:693–705

Dames S, Durtschi J, Geiersbach K, Stephens J, Voelkerding KV (2010) Comparison of the illumin a genome analyzer and Roche 454 GS FLX for resequencing of hypertrophic cardiomyopathy-associated genes. J Biomol Tech 21(2):73–80

Duan M, Chen L, Ge Q, Lu N, Li J, Pan X, Qiao Y, Tu J, Lu Z (2019) Evaluating heteroplasmic variations of the mitochondrial genome from whole genome sequencing data. Gene 699:145–154

Eduardoff M, Xavier C, Strobl C, Casas-Vargas A, Parson W (2017) Optimized mt DNA control region primer extension capture analysis for forensically relevant samples and highly compromised mt DNA of different age and origin. Genes 8(10):237

Elson JL, Andrews RM, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N (2001) Analysis of European mt DNAs for recombination. Am J Hum Genet 68(1):145–153

EMPOP-EDNAP (n.d.) Mitochondrial DNA population database. http://www.empop.org

Falkenberg M, Gaspari M, Rantanen A, Trifunovic A, Larsson NG, Gustafsson CM (2002) Mitochondrial transcription factors B1and B2 activate transcription of human mt DNA. Nat Genet 31:289–294

Gagliardi D, Stepien PP, Temperley RJ, Lightowlers RN, Chrzanowska-Lightowlers ZM (2004) Messenger RNA stability in mitochondria: different means to an end. Trends Genet 20:260–267

Gilbert MT, Hansen AJ, Willerslev E, Rudbeck L, Barnes I, Lynnerup N, Cooper A (2003) Characterization of genetic miscoding lesions caused by postmortem damage. Am J Hum Genet 72:48–61

Giles RE, Blanc H, Cann HM, Wallace DC (1980) Maternal inheritance of human mitochondrial DNA. Proc Natl Acad Sci U S A 77:6715–6719

Goto Y, Nonaka I, Horai S (1990) A mutation in the tRNALeu (UUR) gene associated with the MELAS subgroup of mitochondrial encephalomyopathies. Nature 348:651–653

Guillen M, Lareu MV, Pestoni C, Salas A, Carracedo A (2000) Ethical-legal problems of DNA databases in criminal investigation. J Med Ethics 26(4):266–271

Hagelberg E, Goldman N, Lió P et al (1999) Evidence for mitochondrial DNA recombination in a human population of island Melanesia [published correction appears in Proc R Soc Lond B Biol Sci 2000; 267(1452):1595–1596]. Proc Biol Sci 266(1418):485–492

Hofreiter M, Paijmans JL, Goodchild H, Speller CF, Barlow A, Fortes GG et al (2015) The future of ancient DNA: technical advances and conceptual shifts. Bio Essays 37:284–293

Hollard C, Keyser C, Delabarde T, Gonzalez A, Vilela Lamego C, Zvénigorosky V, Ludes B (2017) Case report: on the use of the HID-ion AmpliSeqTM ancestry panel in a real forensic case. Int J Legal Med 131(2):351–358

Holt IJ, Harding AE, Morgan-Hughes JA (1988) Deletions of muscle mitochondrial DNA in patients with mitochondrial myopathies. Nature 331:717–719

Howell N (1997) mt DNA recombination: what do in vitro data mean? Am J Human Genet 61 (1):19–22
Iborra FJ, Kimura H, Cook PR (2004) The functional organization of mitochondrial genomes in human cells. BMC Biol 2:9. https://doi.org/10.1186/1741-7007-2-9
Just RS, Scheible MK, Fast SA, Sturk-Andreaggi K, Higginbotham JL, Lyons EA, Irwin JA (2014) Development of forensic-quality full mtGenome haplotypes: success rates with low template specimens. Forensic Sci Int Genet 10(1):73–79
Just RS, Irwin JA, Parson W (2015) Mitochondrial DNA heteroplasmy in the emerging field of massively parallel sequencing. Forensic Sci Int Genet 18:131–139
King MP, Attardi G (1988) Injection of mitochondria into human cells leads to a rapid replacement of the endogenous mitochondrial DNA. Cell 52:811–819
Kraytsberg Y, Schwartz M, Brown TA, Ebralidse K, Kunz WS, Clayton DA, Vissing J, Khrapko K (2004) Recombination of human mitochondrial DNA. Science 304(5673):981
Lopopolo M, Børsting C, Pereira V, Morling N (2016) A study of the peopling of Greenland using next generation sequencing of complete mitochondrial genomes. Am J Phys Anthropol 161 (4):698–704
Ma K, Zhao X, Li H, Cao Y, Li W, Ouyang J, Liu W (2018) Massive parallel sequencing of mitochondrial DNA genomes from mother-child pairs using the ion torrent personal genome machine (PGM). Forensic Sci Int Genet 32(2017):88–93
Maciejewska A, Jakubowska J, Pawłowski R (2013) Whole genome amplification of degraded and nondegraded DNA for forensic purposes. Int J Legal Med 127(2):309–319
Magalhães S, Marques SL, Alves C, Amorim A, Alvarez L, Goios A (2015) Evaluation of heteroplasmy detection in the ion torrent PGM. Forensic Sci Int Genet Suppl Ser 5:e13–e15
Man YW, Griffiths PG, Brown DT, Howell N, Turnbull DM, Chinnery PF (2003) The epidemiology of Leber hereditary optic neuropathy in the North East of England. Am J Hum Genet 72 (2):333–339
Marshall C, Sturk-Andreaggi K, Daniels-Higginbotham J, Oliver RS, Barritt-Ross S, McMahon TP (2017) Performance evaluation of a mitogenome capture and Illumina sequencing protocol using non-probative, case-type skeletal samples: implications for the use of a positive control in a next-generation sequencing procedure. Forensic Sci Int Genet 31:198–206
Mcaulay V, Richards M, Sykes BC (1999) Mitochondrial DNA recombination—no need to panic. Proc R Soc B 266:2037–2039. https://doi.org/10.1098/rspb.1999.0883
MITOMAP (n.d.) A human mitochondrial genome database. https://www.mitomap.org
Ondov BD, Cochran C, Landers M, Meredith GD, Dudas M, Bergman NH (2010) An alignment algorithm for bisulfite sequencing using the applied biosystems SOLiD system. Bioinformatics 26(15):1901–1902
Ovchinnikov IV, Malek MJ, Kjelland K, Drees K (2016) Whole human mitochondrial DNA sequencing. Humana Press, New York, pp 157–171
Park S, Cho S, Seo HJ, Lee JH, Kim MY, Lee SD (2017) Entire mitochondrial DNA sequencing on massively parallel sequencing for the Korean population. J Korean Med Sci 32(4):587–592
Parson W, Strobl C, Huber G, Zimmermann B, Gomes SM, Souto L, Lagace R (2013) Forensic science international: genetics evaluation of next generation mtGenome sequencing using the ion Torrent Personal Genome Machine (PGM). Forensic Sci Int Genet 7:543–549
Parson W, Gusmão L, Hares DRR, Irwin JAA, Mayr WRR, Morling N, Parsons TJJ (2014) DNA Commission of the International Society for Forensic Genetics: revised and extended guidelines for mitochondrial DNA typing. Forensic Sci Int Genet 13:134–142
Peck MA, Brandhagen MD, Marshall C, Diegoli TM, Irwin JA, Sturk-Andreaggi K (2016) Concordance and reproducibility of an Ext generation mtGenome sequencing method for high quality samples using the Illumina MiSeq. Forensic Sci Int Genet 24:103–111
Poletto MM, Malaghini M, Silva JS, Bicalho MG, Braun Prado K (2019) Mitochondrial DNA control region diversity in a population from Parana state—increasing the Brazilian forensic database. Int J Legal Med 133(2):347–351

Prinz M, Carracedo A, Mayr WR, Morling N, Parsons TJ, Sajantila A, Schneider PM (2007) DNA Commission of the International Society for Forensic Genetics (ISFG): recommendations regarding the role of forensic genetics for disaster victim identification (DVI). Forensic Sci Int Genet 1(1):3–12

Röck AW, Dür A, Van Oven M, Parson W (2013) Concept for estimating mitochondrial DNA haplogroups using a maximum likelihood approach (EMMA). Forensic Sci Int Genet 7 (6):601–609

Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, Bustillo J (2011) An integrated semiconductor device enabling non-optical genome sequencing. Nature 475 (7356):348–352

Ruiz-Pesini E, Lott MT, Procaccio V, Poole JC, Brandon MC, Mishmar D, Wallace DC (2007) An enhanced MITOMAP with a global mt DNA mutational phylogeny. Nucleic Acids Res 35 (Suppl 1):823–828

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 74(12):5463–5467

Schwartz M, Vissing J (2002) Paternal inheritance of mitochondrial DNA. N Engl J Med 347:576–580

Shendure J, Porreca GJ, Reppas NB, Lin X, McCutcheon JP, Rosenbaum AM, Church GM (2005) Molecular biology: accurate multiplex polony sequencing of an evolved bacterial genome. Science 309(5741):1728–1732

Shih YS, Bose N, Gonçalves BA, Erlich AH, Calloway DC (2018) Applications of probe capture enrichment next generation sequencing for whole mitochondrial genome and 426 nuclear SNPs for forensically challenging samples. Genes 9(1):49

St John J, Sakkas D, Dimitriadi K, Barnes A, Maclin V, Ramey J, Barratt C, De Jonge C (2000) Failure of elimination of paternal mitochondrial DNA in abnormal embryos. Lancet 356 (9225):257–258

Strobl C, Eduardoff M, Bus MM, Allen M, Parson W (2018) Evaluation of the precision ID whole MtDNA genome panel for forensic analyses. Forensic Sci Int Genet 35:21–25

Sutovsky P, Moreno R, Ramalho-Santos J et al (1999) Ubiquitin tag for sperm mitochondria. Nature 402:371–372

Templeton JEL, Brotherton PM, Llamas B, Soubrier J, Haak W, Cooper A, Austin JJ (2013) DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification. Investig Genet 4(1):1–13

Turchi C, Stanciu F, Paselli G, Buscemi L, Parson W, Tagliabracci A (2016) The mitochondrial DNA make up of Romanians: a forensic mtDNA control region database and phylogenetic characterization. Forensic Sci Int Genet 24:136–142

Van Oven M, Kayser M (2008) Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum Mutat 30:386–394

Wallace DC (1997) Mitochondrial DNA in aging and disease. Sci Am 277:40–47

Wallace HM, Jackson AR, Gruber J, Thibedeau AD (2014) Forensic DNA databases ethical and legal standards: a global review. Egypt J Forensic Sci 4(3):57–63

Woerner AE, Ambers A, Wendt FR, King JL, Moura-Neto RS, Silva R, Budowle B (2018) Evaluation of the precision ID mtDNA whole genome panel on two massively parallel sequencing systems. Forensic Sci Int Genet 36:213–224

Yasmin M, Rakha A, Noreen S, Salahuddin Z (2017) Mitochondrial control region diversity in Sindhi ethnic group of Pakistan. Legal Med 26:11–13

Young B, King JL, Budowle B, Armogida L (2017) A technique for setting analytical thresholds in massively parallel sequencing-based forensic DNA analysis. PLoS One 12(5):e0178005

Zeviani M, Moraes CT, DiMauro S (1988) Deletions of mitochondrial DNA in Kearns-Sayre syndrome. Neurology 38:1339–1346

SNP in Forensic DNA Testing

17

Rohit Saluja, Pankaj Shrivastava, and Swapnil Sinha

Abstract

Single nucleotide polymorphisms (SNPs) are the most common mutations found in humans and are special molecular signatures providing intricate information pertaining to human migration patterns, ancestral history, and predisposition to diseases. In present modern forensic era, SNPs in the human populations are being used by investigative agencies across the world as productive and sometimes definitive evidence in solving cases especially pertaining to suspect/victim identification or establishing relatedness. In the present book chapter, we have provided a comprehensive account of the use of SNPs in forensic sciences. We have explained the various conventional and most modern cutting edge strategies and approaches of typing SNPs. Most importantly we have detailed the numerous applications of SNP genotyping in DNA forensics including disaster victim identification, suspect identification at the crime scene, and establishing precisely the disputes related to paternity. We have also included a few landmark case-studies which used SNP typing to reach conclusive verdict.

Keywords

SNP · Genotyping · NGS · Kinship-analysis

R. Saluja
Department of Biochemistry, All India Institute of Medical Sciences, Bibinagar, India

P. Shrivastava
DNA Fingerprinting Unit, State Forensic Science Laboratory, Department of Home (Police), Government of Madhya Pradesh, Sagar, Madhya Pradesh, India

S. Sinha (✉)
BioAptagen Laboratories Private Limited, Guwahati, India
e-mail: swapnilsinha@bioaptagen.com

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_17

17.1 Introduction

Forensic DNA testing and analysis, as compared to other molecular biology fields, is faced with scarce sample amount which is often highly degraded and contaminated. But in spite of these challenges, DNA testing is emerging as a very important tool in forensics. Single nucleotide polymorphism (**SNP**) is defined as a point mutation (base change, in/del) whose allele frequency is $>1\%$ in a population. SNPs are the most common genetic variation found in humans. It can be heterozygous (mutation on one allele only) or homozygous (mutation on both alleles). SNPs are of special interest in population genetics as they provide valuable information related to human migration and ancestral history including predisposition to diseases. SNPs can be detected in any given DNA/population with high accuracy and precision using cost-effective methods, and now in the post-genomic era, SNPs across genomes can be genotyped in massively parallel fashion using Next-Generation Sequencing technology. Since SNPs are signatures of populations, many investigative agencies across the world are using SNPs as productive and sometimes definitive evidence in solving cases especially pertaining to suspect/victim identification or establishing relatedness (kinship analysis).

This chapter is focused on the use of SNPs in forensic DNA testing and interpretation. The chapter also details the various methods and chemistries used for SNP genotyping. Also included in the chapter are few interesting case-studies where SNPs are used as a prime evidentiary support to identify suspects/victims so as to provide a glimpse how SNPs could be used real time in forensics.

17.2 SNP Typing Methodologies

Over 20 years, a number of new and sensitive technologies for SNP profiling have been developed. Modern technologies are high-throughput and very precise. Below is the description of landmark SNP profiling technologies and their principles. The aim of this section of the chapter is to understand the various methodologies used in SNP genotyping and to choose the most appropriate techniques according to the forensic need.

17.2.1 Allele-Specific Hybridization

Allele-specific hybridization or allele-specific oligonucleotide hybridization (ASO) is the most common SNP typing method used in forensics especially in suspect identification at crime scene. The ASO probe distinguishes SNPs on specific genetic loci. In this method, two probes are designed; each is specific to an SNP site (Fig. 17.1) and only the perfectly matching probe will hybridize to the target and is stable.

Various detection methods have been introduced since the inception of ASO method. Some are summarized below:

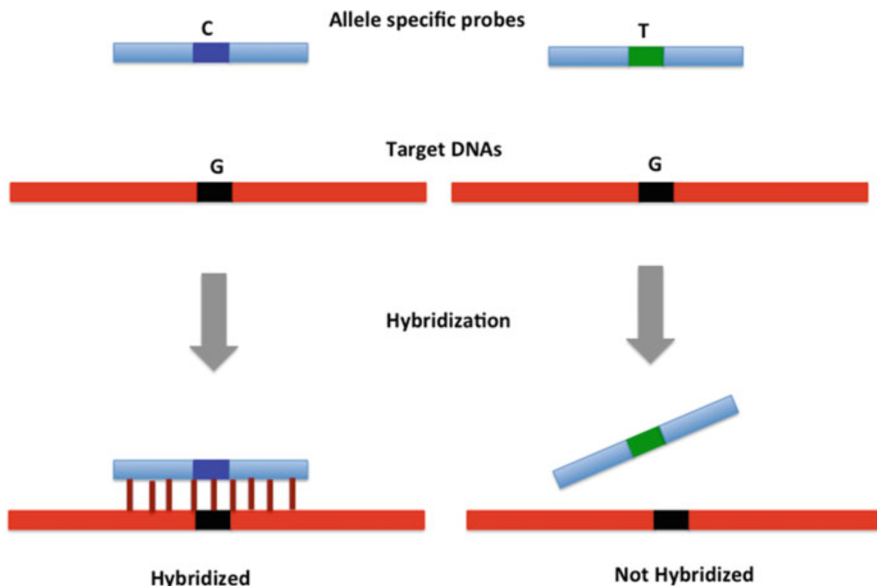


Fig. 17.1 Schematic representation of allele-specific hybridization method

17.2.1.1 FRET-Based Methods

Fluorescence resonance energy transfer (FRET) is a phenomenon in which the energy emitted by one flourophore is absorbed by another flourophore in the proximity and the second flourophore is excited by this energy transfer. In FRET-based methods of SNP genotyping, allele-specific probes are designed conjugated to two different flourophores. The detection is done in real time on qPCR platform. There are several variations of the FRET-based chemistry.

Commonly used TaqMan probe assay (ABI) is based on the 5′ nuclease activity of *Taq* polymerase during PCR (Livak et al. 1995). Two TaqMan probes are designed for each SNP locus, one probe is complementary to the wild-type allele and the other to the mutant allele. Each probe has different dyes conjugated to their 5′OH ends along with a quencher at the 3′OH end. When the probes are not binding to the complementary DNA, the quencher quenches the fluorescence of the flourophore. During the PCR, only specific complementary probe anneals to the template DNA and during extension step, Taq polymerase cleaves the 5′ fluorescent dye resulting in rise in its fluorescence (Fig. 17.2a). Mismatch probes which do not anneal to the template are not cleaved by *Taq* polymerase imparting no fluorescence. The genotype of the sample is detected by measuring the fluorescence intensity of dyes used for each probe.

Another method based on FRET chemistry was developed by Roche, the LightCyclar® method. In this process, two probes are designed in such a way that they are positioned adjacent to each other on target SNP site (Fig. 17.2b). Probe 1 has fluorescein label at its 3′-end and probe 2 is conjugated to LC Red at its 5′-end. When

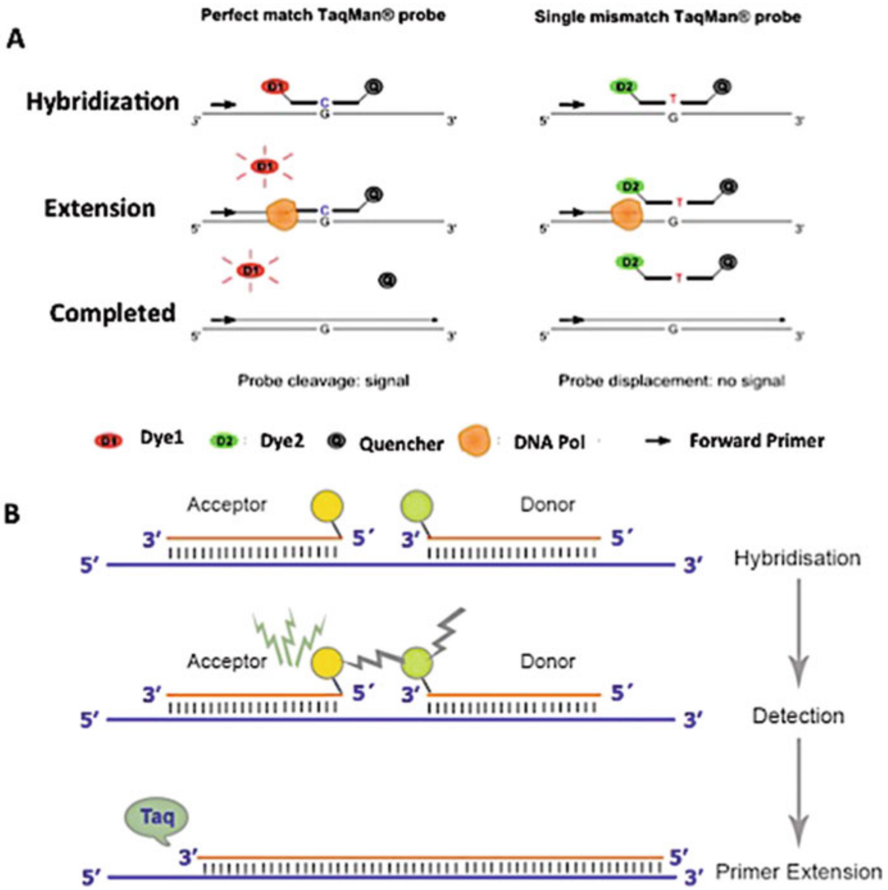


Fig. 17.2 (**a**) TaqMan assay technology (**b**) Schematic representation of LightCyclar chemistry

the probes find exact complementary match, they hybridize adjacent to each other in close proximity. First the fluorescein dye is excited and the wavelength emitted is captured by LC Red for its own excitation. This energy transfer is only possible when both probes are positioned close to each other. The fluorescence intensity of the light emitted by LC Red is measured. When the PCR cycles progress, the florescence intensity increases and is proportional to the amount of DNA amplified during PCR. One of the major advantage of FRET methods is their real-time SNP detection and therefore no post-PCR steps are required. This allows ease of SNP genotyping with precision but a major limitation of FRET-based methods is lack of multiplexing which makes this procedure expensive.

17.2.1.2 Array Hybridization

In this approach, primer sequences are immobilized on a matrix to form a microarray. These primers are SNP specific, and also the hybridization efficiency with the complementary PCR product depends upon the flanking sequence at the

SNP site. The target sample DNA is PCR amplified and a fluorescent label (e.g. SYBR Green) is added along with PCR amplification. The labeled PCR products are then hybridized to the oligo microarray. The hybridization is probe-specific and the fluorescence is scanned. This method provides parallel detection of many SNPs on a single array. One limitation of this process is difficulty in designing the optimal conditions for large multiplexing. But the GeneChip system (Affymetrix) has optimized using 6–10 ASO probes for each SNP. The test probes include all possible sequences at the polymorphic site penultimate nucleotides flanking the SNP, similar to tiling array (Fodor et al. 1991; Pease et al. 1994). This method can be used for typing large number of SNPs, fulfilling current forensic requirements.

17.2.2 Primer Extension

These methods are based on polymerase chain reactions. There is one type of reactions involving minisequencing where the polymorphic site is typed by incorporation of ddNTP by DNA polymerase complementary to the base of the sample DNA. The second type of reaction involves allele-specific extension of the polymorphic site only if the primers have a perfect match with the template. Below are the descriptions of signature primer extension methods used in SNP typing in forensics:

17.2.2.1 SNaPshot™ Method by Applied Biosystems

The SNaPshot method is the propriety of ABI and is the most common SNP detection method in Forensics. This is a single base extension PCR method in which the primer ends just penultimate to the SNP site (Fig. 17.3). Only ddNTPs are used in the reaction, and each ddNTP is fluorescently labeled with unique dyes. When DNA polymerase performs PCR, the fluorescently labeled complementary ddNTP is added and further extension is stalled. The single base extended fragments are separated and imaged in the capillary electrophoresis sequencer and the presence of homozygous/heterozygous allele is detected. SNaPshot can be multiplexed to 30-plex with ease providing flexibility and precision to the process. Recently Daniel et al. have combined SNaPshot with next-generation sequencing using Ion Torrent PGM (Life Technologies) system genotyping 136 SNPs in a single run.

17.2.2.2 MALDI-TOF Mass Spectroscopy

This matrix-assisted laser desorption ionization-time of flight method is a single base extension method in which the mass of the single based extension DNA fragment is measured (Fig. 17.4). The single based extension is done by DNA polymerase where a complementary ddNTP and the extended DNA fragment is immobilized onto a matrix over a chip or plate. The lasers are then fired on the matrix in a process called desorption which results in the vaporization (ionization) of the matrix particles including attached DNA. The DNA travels through the flight tube, where it travels to the detector under the influence of electric field. The time of flight and reaching

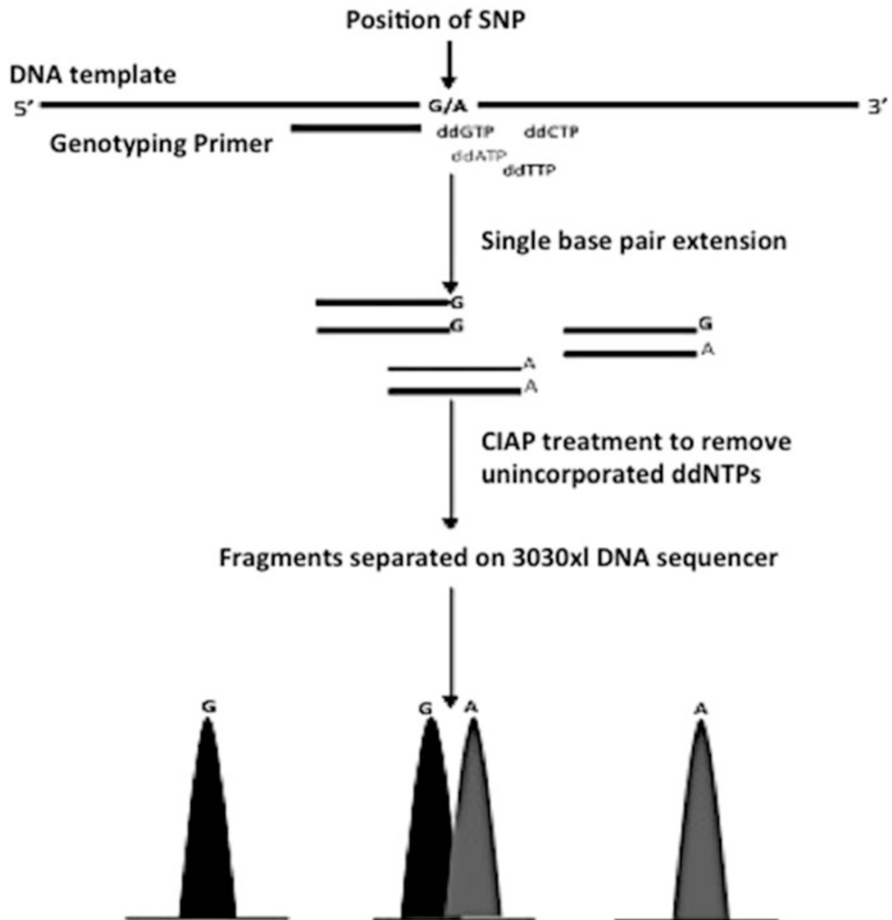


Fig. 17.3 Outline of the SNaPshot procedure for SNP detection

the detector is directly proportional to the mass (*e*/*m* ratio) of the DNA fragment. The mass of single base extended is measured in a very precise manner and the detection of nucleotide added is done. The resolution of this method is very high as it is totally based on mass of the individual nucleotide added at the end of the DNA target. There are many approaches SNP typing based on MALDI-TOF chemistry. The most popular is the MassEXTEND assay (Sequenom) (Braun et al. 1997) and PinPoint assay (ABI) (Haff and Smirnov 1997). One limitation of this technology is the requirement of extra pure DNA as a small contamination will result in noise mass measurement and spurious interpretation. The assay can also be multiplexed upto 20-plex (Kim et al. 2003) by adding a non-human tail sequence at the 5′-end of the primer.

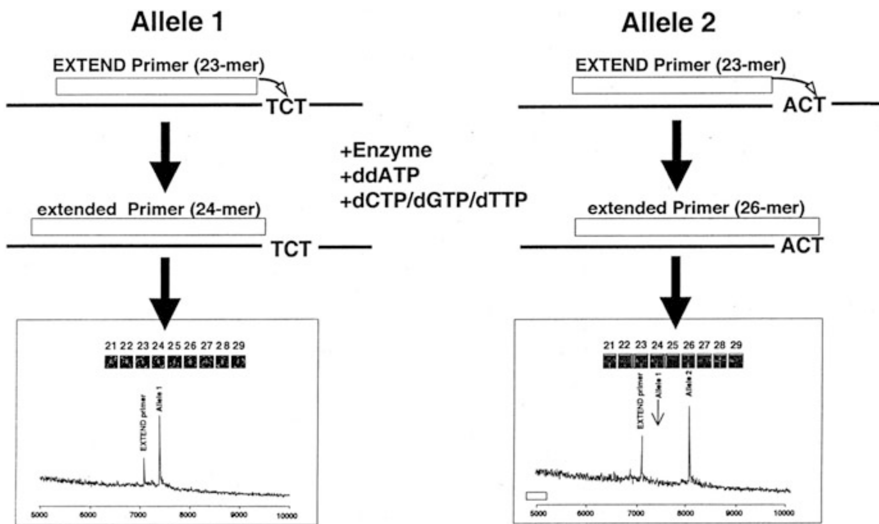


Fig. 17.4 Schematic representation of MassEXTEND MALDI-TOF method of SNP detection. (*Source: Cashman et al. Drug Met. Dispo. 2001, 12: 1629–1637*)

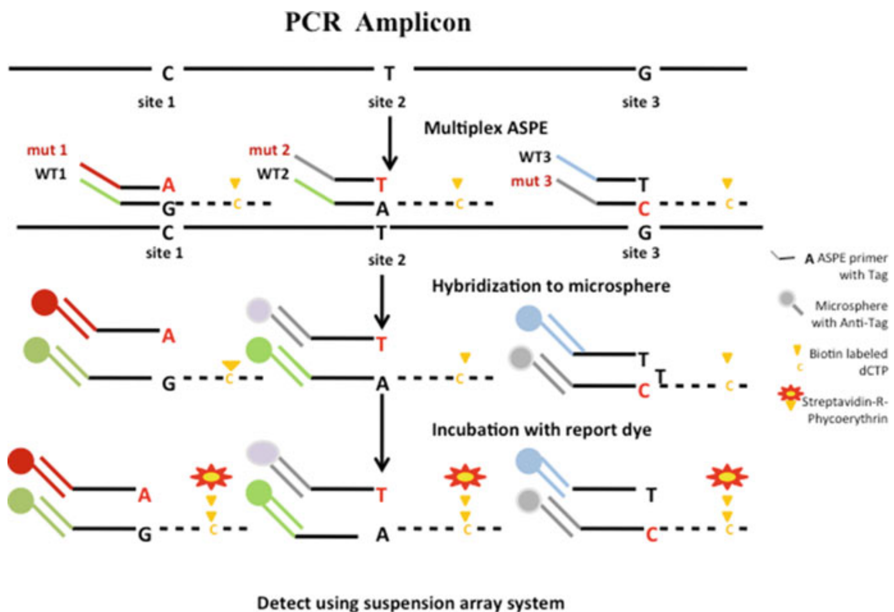


Fig. 17.5 Allele-specific primer extension done using microarray for multiplex-SNP detection

17.2.2.3 Allele Specific Extension

This process is based on the relative efficiency of DNA polymerase to extend the 3′OH ends of the matched and mismatched primers used (Fig. 17.5). Two primers are used each specific for each polymorphic allele. DNA polymerase extends a primer only when the 3′-end is perfectly complementary to the DNA sample template. The primer extending the SNP site and forming product is detected and the genotype is detected. The detection can be done on microarray using fluorescently labeled nucleotides. The primers are immobilized on the microarray chip and the target DNA is provided along with labeled NTPs and DNA polymerase.

17.2.3 Next-Generation Sequencing (NGS) Methods

Over the last decade, NGS technology has completely changed the way genomics was approached. With its high-throughput capacity of massively parallel DNA sequencing, low cost as compared to Sanger sequencing and rapid results, NGS methods have become an important toolkit for genomic research including forensics. Since the advent of Roche 454 Flex sequencing technology in 2005 (Margulies et al. 2005), NGS methods have undergone rapid modifications and development for optimal output, which are very sensitive and cost-effective. The utility of next-generation sequence reads is highlighted by the facts that the length of the sequence is much shorter than that from a capillary sequencers and that each NGS platform has a unique error model. Both of these features are instrumental in forensic DNA analysis where sample DNA is mostly fragmented and low in quantity and the error check is crucial when these methods are used for SNP typing in forensics. Explained below are very recent NGS platforms that are used for SNP typing which are or could be used in forensic DNA testing:

17.2.3.1 Roche 454 FLX Platform

The 454 FLX sequencer was the first NGS platform introduced by Roche in 2005. This method is based on pyrosequencing chemistry. It is a PCR-based method where the sample DNA is used as template. The incorporation of base by DNA polymerase results in the release of pyrophosphate (PPi) which is utilized by the APS (Adenosine Phospho Sulfate) to be converted in ATP, this ATP is used by firefly Luciferase enzyme to convert Luciferin into Oxyluciferin which generates light. Quantity of light produced is directly proportional to the number of nucleotides incorporated. In 454 FLX system, DNA samples are fragmented and attached to Roche 454-specific adaptor sequences (Fig. 17.6a). These sequences are next hybridized to complementary adaptor sequences attached on agarose beads (Fig. 17.6b). Next, these fragment: bead complexes is isolated into individual oil: water micelles that also contain all the PCR reactants and PCR is carried out inside micelles. This emulsion PCR results in amplification of input DNA to million copies (Fig. 17.6c). These amplicons are then subjected to pyrosequencing after arraying the beads on picotiter plate which holds single bead in each well (Fig. 17.6d). The sequence/genotyping data is annotated by the in-built software. Currently, FLX system produces an average read length of

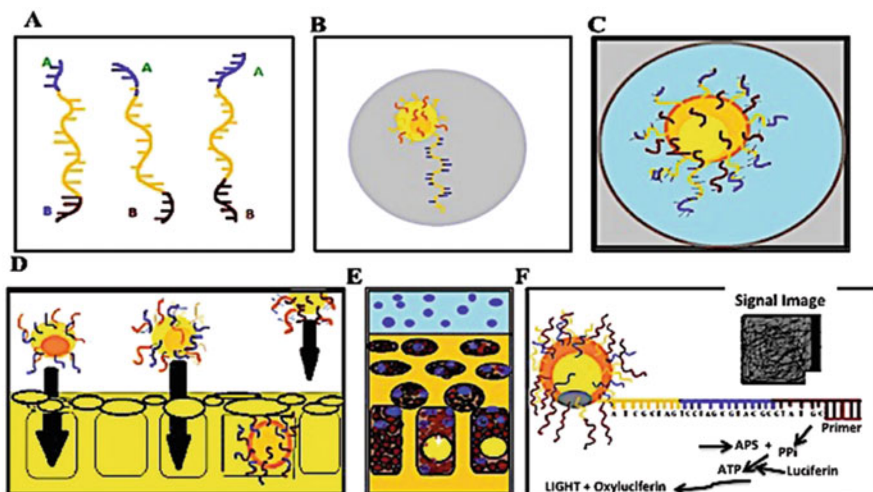


Fig. 17.6 Step-wise schematic representation of 454 FLX NGS procedure

250 nucleotides/cell in a 100 flow capacity. The resulting reads yield 100 Mb of quality data on average in about 8–10 h run.

17.2.3.2 ABI SOLiD Platform

The SOLiD NGS platform by Applied Biosystems uses DNA Ligase and a unique sequencing approach as outlined in Fig. 17.7. The initial steps are similar like other methods; DNA is fragmented and ligated to specific adaptor sequences. Template DNA amplification is done by emulsion PCR like in Roche method, but here the DNA fragments are hybridized to magnetic beads (Fig. 17.7). Next, a primer complementary to the adaptor sequence is added along with DNA ligase and specific fluorescent-labeled 8-mers, whose fourth and fifth bases are encoded by the attached fluorescent group. After each ligation step, fluorescence is detected followed by removal of bases from the ligated 8-mer along with the fluorescent group and next step is repeated with another round of ligation. Read lengths in SOLiD are generally between 25 and 35 bp, and each sequencing run yields between 2 and 4 Gb of DNA sequence data. The purpose behind two-base encoding is to differentiate true single base variants from base-calling errors.

17.2.3.3 Illumina Sequencing Platforms

Illumina offers flexible, innovative, and cost-effective NGS platform for analysis of genetic variation. The initial steps are similar to all NGS procedures, DNA is fragmented and attached to adaptors followed by annealing on the complementary oligo lawn immobilized on a “flow cell.” This step is performed by an automated device called a Cluster Station. The flow cell is a sealed glass slide-like structure with 8 channels harboring the oligo lawn. Each DNA fragment is now amplified by PCR to generate million copies of input DNA by a process called “Bridge Amplification”

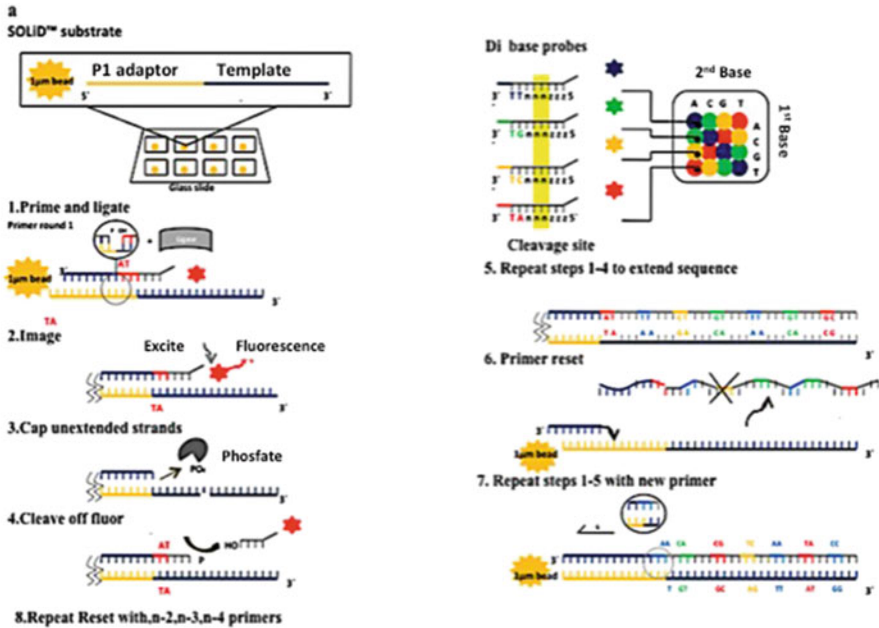


Fig. 17.7 Schematic representation of the strategy used in ABI SOLiD method

(Fig. 17.8). Massively parallel DNA sequencing is next carried out by Illumina's patented Sequence By Synthesis (SBS) process. In SBS technology, all four nucleotides are added to the flow cell channels and PCR is carried out. The nucleotides are labeled uniquely and the 3′-OH group is chemically blocked to stop the incorporation of the next complementary base. After each incorporation of nucleotide, imaging step follows during which each flow cell lane is imaged. Following imaging, the 3′OH blocking group is chemically removed for the next incorporation of base by DNA polymerase. Data is analyzed by alignment of sequence reads to a reference genome to look for SNPs and other genetic variations. Illumina have introduced a number of platforms based on the above described chemistry. Choice of platforms depends on experimental needs and type of data to be acquired. Like the NextraXT offers efficient library preparation methods in record time. The benchtop MiniSeq™ System has an output ranging from 1.8 to 7.5 Gb sequencing while the NovaSeq™ 6000 System can generate 6 Tb and 20 B reads in approximately 2 days. The HiSeq 3000/4000 Series offers cost-effective production-scale sequencing.

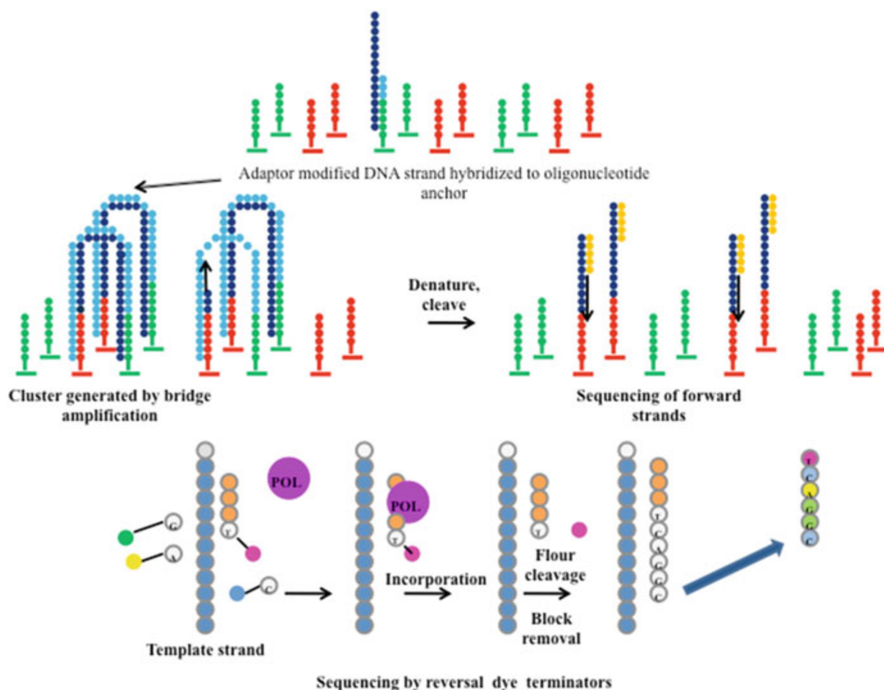


Fig. 17.8 Schematic representation of Illumina SOLEXA NGS technology

17.3 Applications of SNP Typing in Forensics

In recent years, SNPs have emerged as novel genetic markers getting sufficient spotlight in forensic field. As SNPs are the most common polymorphisms in human genome having low mutation rate and ease of their profiling using relevant automated technology, their use in forensics is becoming increasingly crucial. Given below are some major applications of SNPs in forensic DNA testing:

17.3.1 Disaster Victim Identification

The World Health Organization (WHO) defines disaster as "serious disruption of the functioning of society causing widespread human, material, economic or environmental losses." Disasters can be natural like floods, earthquake, fire, epidemics or it could be manmade like wars, any major terrorist attack, plane crash, and road mishaps. Whatever the cause, disasters result in huge loss of human life. One of the major issues during any disaster is the identification of victims. Traditionally, Disaster Victim Identification (DVI) is done by identification of personal effects like documents, IDs, and other belongings or more reliably by collection and analysis of

fingerprints, hair, or dental samples. DNA-based identification of disaster victims is generally used as a last resort when all other identification methods have failed. This is mainly because most of the time the tissue samples are degraded and heavily mixed with other victims. Earlier, molecular procedures like blood group polymorphisms or HLA typing were used, but these methods have their own limitations. They provide insufficient genetic information and these procedures are not always possible due to the degraded quality of the samples. Recently after the introduction of modern PCR-based methods, genetic markers have been used for DVI. One major reason behind the increasing popularity of using genetic markers for DVI is the availability of optimized and robust methods of DNA extraction methods and highly sensitive genotyping NGS procedures which can be carried out with as little as 10 ng of input sample DNA. SNPs provide useful and sometimes unique genetic information which makes them increasingly popular in DVI. Also, with the availability of SNP databases like NCBI dbSNP, and IGVdb (Indian Genome Variation database), the SNP-based methods performed on disaster victims can give very crucial information on their identification most importantly about their ethnicity and origin. There are many SNPs, LD (Linkage Density) blocks which are specific to certain population(s) which becomes instrumental and narrowing down the identification of victims from a mass destruction site. Also, ease of using as small as 60–100 bp of DNA fragments in NGS methods, the highly degraded DNA samples of victims can be used. SNPs are characterized by low mutation rates as compared to STR markers which, makes them stable genetic marker to be used to identify the ethnicity of the victims. SNPs have differential frequency distribution across populations making them highly useful to provide valuable information on demographic histories and geographical origin (Ziętkiewicz et al. 2003) both of which can be very instrumental in DVI. SNP genotyping results of the victims can be scanned against the available databases belonging to specific populations across the world and in this way information like ethnic origin, population identification can be a major asset in DVI. Presently, the SNP markers used in forensics are those related to ancestry-informative markers (AIMs), autosomal SNPs which have large inter-population allele frequency differences and some mtDNA related SNPs. Few, currently available panels of AIM SNPs are 73 AIMs DNAWitness Bio-Geographical Ancestry Kit, 320 AIMs Eurasian panel and 1476 AIMs European panel (Budowle and van Daal 2008). These SNP marker panels are a huge help in answering missing links about victims' continental origin and relative admixture across populations (Decorte 2010; Nassir et al. 2009).

The choice of SNPs for DVI needs to be combined with statistical skill of interpretation, construction of more population specific databases and markers to be analyzed depending on purpose of investigation. For example, if kinship analysis is performed using SNPs, the mtDNA or Y-chromosome are the markers of choice. Often unlinked SNPs (rather than LD blocks) having high population-specific heterozygosity are ideal for identification of victims by narrowing down the search periphery to a specific ethnicity (Dixon et al. 2005, 2006; Kidd et al. 2006). In cases where no match is found, personalized databases like AIMs could be used to

establish the ethnicity of the victim and further investigations could be carried out for precise identification (Phillips et al. 2009).

17.3.2 Kinship Testing

Investigating relatedness between suspects/individuals, missing persons, victims are the areas where DNA testing plays very crucial and definitive role in forensics. DNA-based Kinship testing helps to establish precise relationship between two potentially related individuals. Very recently, Morimoto et al. (2017) have computationally constructed second-degree kinships (i.e., uncle-nephew and grandfather-grandson) and third-degree kinships (i.e., first cousins and great-grand-father-great-grandson) for 174,254 SNP markers based on linkage disequilibrium and recombination events. They envisage that this method can be a useful tool for distinguishing relationships between two alleged individuals having same degree of kinship. The various kinship testing that can be achieved using SNP information of human genome are the following:

- ***Paternity testing***: DNA paternity testing helps to determine the true biological father/mother of a child with almost 99.9% accuracy. Generally mouth swab samples of are collected followed by DNA extraction and PCR amplification of specific loci containing informative SNPs/other polymorphisms and are analyzed by any chosen method of genotyping.
- ***Sibling testing***: In this kinship testing the biological relationship between two alleged individuals is established through its DNA polymorphism information. This testing is generally performed when the biological parents of the people involved are not present or alive. The test procedures and SNP markers used are similar to that done in paternity testing.
- ***Lineage/Ancestry profiling***: In forensic studies, where there is little or no information of possible suspect/victim and database, information obtained by DNA ancestry studies can be very crucial in narrowing the search criteria and identifying the potential suspect. Use of DNA ancestry is relatively new to forensics, but in future it could be pivotal in crime scene investigation or missing person links including victims of disasters. Walsh et al. (2011) have devised a method, called IrisPlex for the accurate prediction (precision value ~90%) of individual's ancestral information based on blue and brown eye color. They have used 6 SNPs which are associated with blue and brown eye color. The allele frequencies of these SNPs vary and are phenotypically associated with eye color. Han et al. (2008) conducted a genome-wide association study on population of European ancestry and found 173 SNPs on gene loci IRF4 and SLC24A4 associated to hair color and skin pigmentation.
- ***Ethnicity testing***: One of the most accurate information provided by SNP genetic markers is that of ethnicity. Humans in the course of evolution, harbor mutations throughout their genome. Migration, disease pressure, environmental changes manipulate our genome in such a way so as to make survival fittest in our

surrounding habitat: the classic Darwin's theory of evolution. There are many SNPs which are specific to populations and some are even confined to one population group. These individual SNPs or most of the time haplotypes (a group of closely placed SNPs) provide valuable information of person's ethnic group. With the availability of human SNP (dbSNP) and haplotype (HapMap project) databases on many populations across the world, ethnicity testing is proving to be an important information provider especially in cases of victim identification during disasters. Klimentidis and Shriver (2009) have shown association of facial features of various ethnic groups in Americas with various SNP markers and validated their findings using facial reconstruction (molecular photo editing).

17.3.3 Suspect Identification

Suspect identification using SNP genotyping as search tool typically means matching of the allele found in the sample collected at the crime scene to the alleles of the various DNA samples identified as suspects. Traditionally suspect identification was done by battery of physical and circumstantial evidences but after the advent of modern DNA fingerprinting and other sophisticated molecular techniques, suspect identification using SNPs as genetic signatures can be done with high precision even when other physical evidences are inconclusive. There are a number of genetic loci on which the associated SNPs are used for suspect identification. Samples are collected from the crime scene in the form of blood, hairs, semen, etc. DNA is extracted and SNP markers are genotyped. Search for the convicts can be done directly, by matching the SNP profile obtained from samples at the crime scene to few individuals who are potential suspects. Alternatively, the SNP profile of the crime scene DNA sample can be scanned with available databases of individuals who are identified as serial offenders. The FBI (Federal Bureau of investigation), USA maintains a database called as "CODIS" (Combined DNA Index System) containing the DNA profiles contributed by federal, state, and local participating forensic laboratories. This database is widely used by the department to narrow down or for definitive identification of suspected perpetrator by scanning his/her SNP profile against their state database of convicted offender and arrestee profiles which are contained in CODIS.

17.3.4 Forensics Microbiological Analysis (FMA)

Following the Anthrax attack in USA in 2001, the Federal Bureau of Investigation has developed a new field of Microbial Forensics which detects and identifies microorganisms found in biological crime scene ultimately tracing the trail to its source (McEwen et al. 2006). Considering the serious consequences of potential biological warfare, FMA has emerged as a very important tool. SNP diagnostic PCRs done on human samples found at the crime scene for profiling the presence of

microorganisms can provide sufficient information about the source of biological attack. In an investigative study, Fierer et al. (2010) have examined bacterial traces on human skin thereby inferring that bacteria left by human skin can provide valuable DNA information for forensics analysis. In another study by Lilje et al. (2013) the workers collected 11 samples from different environments like grassfields, forests, and urban parks with different microbial flora and genotyped the 18S RNA gene markers demonstrating that these SNP markers could be used for constructing database and running filtered matching tests.

With the availability of high-speed NGS methods, forensic microbial polymorphism database can be constructed which will become instrumental in fast and accurate identification of terrorists in biological warfare.

17.4 Examples of Case-Studies

SNPs have found important and definitive application in forensic field and have been employed regularly to gather valuable information on suspects/victims. There are voluminous case-studies where SNP genotyping is employed as a part of investigation. Few of these interesting case-studies are mentioned below:

- ***Madrid Train Bombing Case***:
  On the 11th of March 2004 ten improvised explosive devices (IED) were detonated on four commuter trains in Madrid in a coordinated terrorist attack that took the lives of 191 people. Samples were collected primarily from contact traces, or the items used by suspects e.g. toothbrush, blanket, hat, and scarfs. Samples were obtained after cutting or swabbing the tissue traces from the items used by the suspects. The forensic experts perform the ancestry studies and phenotypic inference derived from a DNA sample. SNP profiling was carried out using 34-plex autosomal ancestry-informative-marker single nucleotide polymorphism (AIM-SNP) assay to extract the precise ancestry on the DNA from seven unmatched case samples. To estimate the likely ancestry assignment error from each marker set used, a panel of 96 individuals was constructed using two varied ethnic groups living in Madrid; Spanish and Moroccan and these samples were genotyped before the test samples. One of the suspects was predicted to be of North African origin by AIM-SNP analysis and one suspect was traced to Algeria. The results from genotyping aided hugely in narrowing down the suspect search and ultimately took the authorities to the attackers along with other evidential support (Phillips et al. 2009).
- ***Suicide Bombing of the Australian Embassy in Jakarta***:
  On September 4, 2009, a car bomb exploded in a delivery van outside Australian embassy in South Jakarta, Indonesia, killing nine people including the perpetrator. The explosion was so massive that very small traces of tissue of the perpetrator was recovered from the site making it almost impossible to identify the bomber by traditional means. Crime scene investigation revealed the serial locations of the suspects before the blast and that helped in the collection of

tissue samples. First, MtDNA analysis was carried on 17 tissue samples, profiling HV1 gene markers which revealed the common source of the perpetrator. Further analysis connected the perpetrator maternally to the family member of one of the four suspects. Standard genetic marker procedures confirmed the identification of the prime suspect. This case is of special interest as application of DNA genotyping played the central role in identification of the perpetrator.

- ***Melanie Road Murder Case***:
  In 1984, Melanie Road, 17, was sexually assaulted and murdered in Bath, Somerset. Police collected blood and semen samples from Melanie's clothing and crime scene. The blood group analysis revealed nothing and due to lack of conclusive evidences and non-availability of DNA tests at that time, failed to grab the real culprit. The police kept and preserved all the samples. As scientific techniques advanced, various tests were performed with the evidentiary samples on police persistence. In 1990s the DNA of the sample was run against the available police databases but there was not match found then as no kin or the culprit was ever arrested. In 2014, the culprit's daughter was arrested following a domestic violence case in which, according to the police protocol, she was cautioned and released but her DNA sample was collected and uploaded on national database. In 2015, familial DNA testing for Melanie's murder was rerun and a match was found on culprit's daughter. The police then arrested Christopher Hampton on June 1, 2015, whose mouth swab samples were a perfect match to the semen and blood samples collected from Melanie's murder site. After 32 years of murder, Melanie Road finally got justice with the help of DNA forensics.

17.5 SNP Databases Used in DNA Forensics

Given below is a list of most commonly used SNP databases for forensic purposes.

Database name	Curator	Populations included
dbSNP (Database of Single Nucleotide Polymorphisms)	NCBI, NIH Bethesda, USA	4 major populations along with many ethnic populations across world
ALFRED (ALlele FREquency Database)	US National Science Foundation	744 populations across Africa, Asia, Americas, and Europe
FROG-kb (Forensic Resource Reference On Genetics-knowledge base)	US National Science Foundation	744 populations across Africa, Asia, Americas, and Europe
Forensic SNP Information	NIST, USA	All dbSNP and TSC populations
IGVdb (Indian Genome Variation database)	Indian Genome Variation Consortium	55 ethnic populations across 4 linguistic groups of India
International HapMap project	NCBI, NIH Bethesda, USA	11 populations spanning African, Asian and European populations

17.6 Limitations of SNP Typing in Forensics

17.6.1 Sample Quality

The most common problem with SNP typing in forensics is the scarcity of DNA obtained from crime scene or disaster site. Although, modern SNP typing techniques for example various NGS methods require relatively low amount of DNA, still many times DNA methods cannot be applied as the starting quantity is too low. Sometimes low-quantity DNA makes it almost impossible to rerun the test if first run fails or is inconclusive. Another problem pertaining to sample quality is high fragmentation and degradation. This especially happens in DVI where samples are hugely degraded as a result of disaster.

17.6.2 Lack of SNP Databases

Even though DNA forensics came into existence in 1990s but still many countries cannot use this useful and sometimes definitive tool in their police investigation as the main procedure. This is because many countries including India does not have a detailed genetic variation database. FBI has CODIS, which is a huge DNA resource of covering large ethnic groups across USA. The public SNP databases like dbSNP and ALFRED have populations covering many continents but still more databases are required which cover many populations who are never been genotyped.

17.6.3 Sample Mixing

When any crime scene or disaster site is scanned for DNA evidence, samples obtained contain mixing of two or more individuals. Sample mixing is a serious problem in DVI where there is huge intermixing of samples. Sometimes during disasters, e.g. airplane crash, human samples are mixed with microflora of soil or environment which makes downstream analysis and interpretation difficult. The sample intermixing becomes a serious problem when the test SNP markers are present in heterozygous form. Mixed SNP profiles of many individuals have a direct association with choice of markers to be analyzed and also on the statistical method for the interpretation of results.

References

Braun A, Little DP, Koster H (1997) Detecting CFTR gene mutations by using primer oligo base extension and mass spectrometry. Clin Chem 43:1151–1158

Budowle B, van Daal A (2008) Forensically relevant SNP classes. BioTechniques 44:603–610

Decorte R (2010) Genetic identification in the 21st century—current status and future developments. Forensic Sci Int 201:160–164
Dixon LA, Murray CM, Archer EJ, Dobbins AE, Koumi P, Gill P (2005) Validation of a 21-locus autosomal SNP multiplex for forensic identification purposes. Forensic Sci Int 154:62–77
Dixon LA, Dobbins AE, Pulker HK, Butler JM, Vallone PM, Coble MD, Parson W, Berger B, Grubwieser P, Mogensen HS, Morling N, Nielsen K, Sanchez JJ, Petkovski E, Carracedo A, Sanchez-Diz P, Ramos-Luis E, Briõn M, Irwin JA, Just RS, Loreille O, Parsons TJ, Syndercombe-Court D, Schmitter H, Stradmann-Bellinghausen B, Bender K, Gill P (2006) Analysis of artificially degraded DNA using STRs and SNPs—results of a collaborative European (EDNAP) exercise. Forensic Sci Int 164:33–44
Fierer N, Lauber CL, Zhou N, McDonald D, Costello EK, Knight R (2010) Forensic identification using skin bacterial communities. Proc Natl Acad Sci U S A 107:6477–6481
Fodor SP, Read JL, Pirrung MC, Stryer L, Lu AT, Solas D (1991) Light-directed, spatially addressable parallel chemical synthesis. Science 251:767–773
Haff LA, Smirnov IP (1997) Single-nucleotide polymorphism identification assays using a thermostable DNA polymerase and delayed extraction MALDI-TOF mass spectrometry. Genome Res 7:378–388
Han JL, Kraft P, Nan H, Guo Q, Chen C, Qureshi A (2008) A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. PLoS Genet 4: e1000074
Kidd KK, Pakstis AJ, Speed WC, Grigorenko EL, Kajuna SL, Karoma NJ, Kungulilo S, Kim JJ, Lu RB, Odunsi A, Okonofua F, Parnas J, Schulz LO, Zhukova OV, Kidd JR (2006) Developing a SNP panel for forensic identification of individuals. Forensic Sci Int 164:20–32
Kim S, Shi S, Bonome T, Ulz ME, Edwards JR, Fodstad H, Russo JJ, Ju J (2003) Multiplex genotyping of the human beta2-adrenergic receptor gene using solid-phase capturable dideoxynucleotides and mass spectrometry. Anal Biochem 316:251–258
Klimentidis YC, Shriver MD (2009) Estimating genetic ancestry proportions from faces. PLoS One 4:e4460
Lilje L, Lillsaar T, Rätsep R, Simm J, Aaspõllu A (2013) Soil sample metagenome NGS data management for forensic investigation. Forensic Sci Int Genet 4:e35–e36
Livak KJ, Flood SJ, Marmaro J, Giusti W, Deetz K (1995) Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. PCR Methods Appl 4:357–362
Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA (2005) Genome sequencing in microfabricated high-density picolitre reactors. Nature 437:376–380
McEwen SA, Wilson TM, Ashford DA, Heegaard ED, Kournikakis B (2006) Microbial forensics for natural and intentional incidents of infectious disease involving animals. Rev Sci Tech 25:329–339
Morimoto C, Manabe S, Fujimoto S, Hamano Y, Tamaki K (2017) Discrimination of relationships with the same degree of kinship using chromosomal sharing patterns estimated from high-density SNPs. Forensic Sci Int Genet 33:10–16
Nassir R, Kosoy R, Tian C, White PA, Butler LM, Silva G, Kittles R, Alarcon-Riquelme ME, Gregersen PK, Belmont JW, De La Vega FM, Seldin MF (2009) An ancestry informative marker set for determining continental origin: validation and extension using human genome diversity panels. BMC Genet 10:39
Pease AC, Solas D, Sullivan EJ, Cronin MT, Holmes CP, Fodor SP (1994) Light-generated oligonucleotide arrays for rapid DNA sequence analysis. Proc Natl Acad Sci U S A 91:5022–5026
Phillips C, Prieto L, Fondevila M, Salas A, Gómez-Tato A, Alvarez-Dios J, Alonso A, Blanco-Verea A, Brión M, Montesino M, Carracedo A, Lareu MV (2009) Ancestry analysis in the 11-M Madrid bomb attack investigation. PLoS One 4:e6583

Walsh S, Liu F, Ballantyne KN, van Oven M, Lao O, Kayser M (2011) IrisPlex: a sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. Forensic Sci Int Genet 5:170–180

Ziętkiewicz E, Yotova V, Gehl D, Wambach T, Arrieta I, Batzer M, Cole DE, Hechtman P, Kaplan F, Modiano D, Moisan JP, Michalski R, Labuda D (2003) Haplotypes in the dystrophin DNA segment point to a mosaic origin of modern human diversity. Am J Hum Genet 73:994–1015

SNP Testing in Forensic Science 18

Moumita Sinha and I. Arjun Rao

Abstract

SNP is the abbreviation for single nucleotide polymorphism (SNP) which occurs at particular site in the genome as variant of single nucleotide. These variations are represented in different population and individuals to some considerable degree. The presence of SNPs is ubiquitous not only in human genome but also among some plants and microorganisms like bacteria. The popularity of SNPs among forensic researchers has been fascinated due to their capable advantages in parentage testing. SNPs have low mutation rates and are very capable in the identification and examination of degraded samples with small amplicons. SNPs provide valuable information on geographical origin and individual identification of unknown humans, plants and microorganisms samples.

Keywords

SNP · Forensic · Testing · DNA · Detection

18.1 Introduction

A single change of sequence in base between individuals and population of different geographical origin as well as ancestry at a specific location in the genome is termed as single nucleotide polymorphisms or SNPs. SNPs as genetic markers are abundant in human genome and readily available for individual identification and geographical identity. SNPs are assessed by forensic scientists for a variety of reasons: (a) SNPs can be amplified in small amplicons of 100 bp by using PCR, (b) degraded DNA samples are easily analysed for SNPs, (c) as size-based separation for SNP is not required in a fully automated processing of sample and analysis of data can be

M. Sinha (✉) · I. A. Rao
Department of Forensic Science, Guru Ghasidas University, Bilaspur, Chhattisgarh, India

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_18

carried out, (d) presence of no stutter peaks are associated with SNP analysis which can make interpretation difficult and (e) above all, SNP markers ease the prediction of ethnic origin and some physical traits of the unknown samples.

SNP markers are potential genetic markers for identification in forensic investigations with higher degree of challenges while typing SNP markers. The foremost challenge is SNP markers are least informative in case of low amount of DNA analysis, as they are mostly bi-allelic. Therefore, their discriminating power is lower than STR markers. But with the advancement of molecular genetics techniques like multiplex PCR, this problem is resolved so far with 35plex for Y chromosome SNPs (Sanchez et al. 2003).

18.2 SNP Typing Methods

Typing of SNP mainly includes two constituents: Allele discrimination method: a technique of determining the base present at SNP locus under study and signal detection method: a technique for reporting the allelic presence. The first method includes three general steps: annealing/hybridization method, primer extension and cleaving through enzymes. The signal detection methods can be enzyme labelled or enzyme unlabelled suitable for end pont studies. To identify the allele-specific product of a discriminant assay without DNA databases by using mass spectrometry is used (Twyman 2005).

18.2.1 Allele Discrimination Methods

18.2.1.1 Hybridization Based (Allele-Specific)

A small portion of synthetic DNA which is complement to the variable target DNA sequence is an allele-specific hybridization. It may play a role of a probe to identify the presence of target region in southern blot assay or other methods. For analysis of allelic variants, two forward primers were designed with the 3′ base of each primer matching only one of the biallelic SNP bases to be evaluated. Incorporation of a primer mismatch at the third base from the 3′ end of the primer has been shown to enhance the specificity of the PCR by further destabilizing the extension of the doubly mismatched primer. To additionally distinguish the allelic primers during amplification, different mismatches in the third 3′ base of both forward primers were employed (Okimoto and Dodgson 1996). Routinely a reverse primer would be designed downstream to the polymorphic site and which have a Tm of 55–60 °C were selected.

18.2.1.2 Primer Extension Method

Primer extension method is absolutely based on a specific sequence enzymatic reaction which is used to asses multiple SNPs in a single reaction. The allele-specific primer extension (ASPE) method includes two different stages which areas the following: To determine the targeted genotype by using an enzymatic

reaction by capturing on a solid microsphere surface for detection. The solution kinetics phase technique allows the microspheres which are sequence labeled to be used in detecting the new templates, which can be done by the help of a suitable capture sequence attached to the allele-specific oligonucleotides. With this technique, the targeted molecules of nucleic acid react with individual base specified primer in individual conditions where the primer hybridization and extension of primer may occur. The allele specific primers are designed to the opposite strand of DNA and the 3′ ends of the primers correspond to the alleles that are interrogated. The 5′ ends of the allele specific primers are hybridization sites for universal PCR primers. The said primer extension reactions are performed using labeled nucleotides, whereas the nucleotide degraded enzymes will be present during the primer extension reactions.

18.2.1.3 Cleaving Through Enzymes

This type of enzyme reactions offers a precise and simple process for SNPs genotyping without doing any targeted genomic DNA PCR amplification (Twyman 2005), whereas this method is known for analyzing individual SNPs. In these methods the surrounded genome sequence of the SNP was identified and amplification done by using specified PCR primers. In this process the SNPs which can modify the site of restriction enzyme can be genotyped by doing the restriction and digestion of the following PCR product and through agarose gel electrophoresis. In the following process the PCR products which hold the site of restriction enzyme will break in to two products, which results in two bands of different sizes on the agarose gel plate. Similarly, in the recognition site of the samples, if an SNP is present then it will no longer be cut by the restriction enzyme and forms a single gel band on gel electrophoresis. A PCR product of heterozygous samples for the SNP identification will be result in three different bands, one band results for the uncut product and two other bands of PCR product for the cut. The present day laboratories use most of these techniques due to their robust usage in molecular biology laboratory. The universe of enymes growing rapidly and various commercial forms of restriction enzymes are available for researchers which increases the genotyping of SNPs widely (Gad 2007).

18.2.1.4 Microarray-Based SNP Genotyping

This type of SNP genotyping is based on the progress of DNA chips technology, SNP microarray for similar type of genotype at different SNP loci. A polymorphic SNP loci subset is selected mostly on the base of their genome position, polymorphism level and assay suitability to be used to build up a microarray or DNA chip. A DNA chip or microarray is a minute panel of glass, silicon or type of metal in which one end point of a number of single strand DNA molecule is linked covalently and organized as spots. The individual spot has a number of copies of a single DNA particle of 20–25 nucleotides, which represents the SNP locus and contains the nucleotides surrounding the SNP. To make sure the high consistency, each SNP is designated by different number of oligonucleotides; in every oligonucleotide the SNP is situated at a different location which ranges from two bases on one side and

the other two bases on the other side. Whereas the set of oligonucleotides is traced at two or three various locations, this may assist in the replication and help in eliminating the false negative results (probably due to improper hybridization). It is also noted down that the SNP locus and the DNA sequences neighboring this locus affects the hybridization efficacy. Therefore, it may be very challenging to adjust the different conditions for detecting the panel of SNPs using an array, whereas the new methods are getting prepared to overcome this difficulty (Sobrinoa et al. 2005).

Each individual genomic DNA to be genotyped for SNPs is utilized for a chain of PCR reactions for the amplification of the short genome region which have different SNPs. For a reliable amplification a set of primers are first defined which converts the SNP to STS. In this process the amplified PCR products are fluorescently labeled, whereas the rest of the PCR products from a single individual are collected and utilized for hybridization with DNA chip technology. During this the PCR products which are not hybridized during the reaction are removed by cleaning and only purely base pair PCR products that attached with the oligonucleotides spots on the DNA chip are allowed. A new way to measure the fluorescence is through the fluorescence scanner at specific spot on the chip, and with the help of image analysis software the output data are analyzed. As each oligonucleotide position on the chip is known, the presence of alleles at different SNP loci is construed. This technique observes and identifies all the SNP loci of the individuals to be tested (Sobrinoa et al. 2005).

18.2.1.5 Biosensors-Based Signal Detection Method

Conventional SNP recognition techniques have a few impediments: they have moderately poor affectability and particularity; they expect enhancement to get various duplicates for identification; they need the utilization of massive instruments; and they can't work remotely.

The chip catches a strand of DNA that has a particular SNP transformation. At that point it makes an electrical sign that is sent remotely to a cell phone. The chip is produced using a graphene field impact transistor with a built bit of twofold stranded DNA that is explicitly intended for the chip and appended to the surface. The DNA is twisted in the center with the goal that it is molded like a couple of tweezers and one side of the DNA can code for explicit SNP. On the off chance that a DNA strand with that SNP comes into contact with the built DNA strand, the DNA with SNP will connect to the side that codes for the change. This opens up the tweezer state of the DNA and makes an adjustment in the electrical flow that the graphene field impact transistor recognizes. The DNA is twisted in the center with the goal that it is molded like a couple of tweezers and one side of the DNA can code for explicit SNP. On the off chance that a DNA strand with that SNP comes into contact with the built DNA strand, the DNA with SNP will connect to the side that codes for the change. This opens up the tweezer state of the DNA and makes an adjustment in the electrical flow that the graphene field impact transistor recognizes. DNA strand removal, the sub-atomic procedure behind the innovation, happens when a DNA twofold helix trades one of its strands for another correlative strand. In the designed variant of the DNA, the tweezers trade a typical strand with a specific SNP. The 'normal' strand is

joined to a graphene transistor and has a reciprocal arrangement for a particular SNP. The other is a 'weak' strand in which some of the nucleotides are replaced with a different molecule to weaken its bonds to the normal strand. SNP strands would then be able to tie all the more firmly to the typical strand to dislodge the frail one. The DNA tweezers at that point have a net electric charge that is effortlessly distinguished by the graphene transistor (Hwang et al. 2018).

18.3 Utility of SNP in Forensic DNA Testing

DNA Typing is a potential technique in forensic Investigations like paternity and maternity disputes, immigration matters, Disaster Victim Identification and identification of missing persons from skeletal remains. In most of the investigations investigation is mainly concerned on human biological materials to find a relation between evidences and to trace the sources of the human biological materials (Budowle and van Daal 2008). In forensic investigations the sources of DNA evidences varies from human to non-human DNA types; which can be a microbial DNA, a plant DNA or an animal DNA. Wild Life Crime involves hunting, poaching, killing, gathering, selling and illegal trafficking of endangered plants and animal species and their products also. In many aspects of forensic investigations species identification is required if due to any reasons a suspicion arises that a national or international border has been broken (Linacre 2006). At present, microbial forensics evolved as one of the important branch of forensic science investigations in which microbes itself used as biological weapon for biological wars or bio terrorism (Keim 2003).

Genetic evidence can be obtained from any material of biological origin, like blood, saliva, hair, bone, teeth, hair and muscle tissue, which can be utilized to trace the species origin of that plants, animal and microorganisms. The collection of DNA markers utilized for identification of biological materials has developed considerably inside the forensic science community with every marker set and currently developed innovation age, enlarging determination or potential sensitivity of the discoveries. In this chapter we quickly depict the uses of particular SNPs utilized in forensic identification of human individuals, endangered animal and plant species with their geographical identity and furthermore its efficacy in identification of microbial species and biogeography in cases of bioterrorism attacks. Therefore, SNiPs provide future platform for Scientic Working Team of Forensic DNA to determine genetical identity of questioned samples or to identify unknown DNA Samples (Budowle and van Daal 2008).

18.3.1 SNPs in Human Identification

Indeed, in forensic cases there are biological specimens which are not acquiescent to examination even after screening with STR; usually, these forensic samples contain too little amount of DNA or are excessively destroyed. Advanced techniques are

available and tried for typing disintegrated samples by sequencing hypervariable control regions and to some extent the coding regions of human mtDNA. Complete Typing of mitochondrial DNA (mtDNA) by sequening is an inevitable, tiresome, and very expensive for the working research laboratories. Because of this only few Forensic DNA Research laboratories executes complete sequencing facilities of mtDNA for criminal investigations. Furthermore, the power of discrimination accomplished by utilizing mtDNA typing isn't as high as that given by the panel of STR loci. Alternative sets of DNA markers that may demonstrate helpful, especially to type degraded biological specimens and for expanding the amount of genetic data gathered from encountered forensic specimens, are SNPs (Budowle 2004; Budowle et al. 2004).

Insertions or deletions, base substitutions in any organism, at single locations in the entire genome creates SNPs. Major of these SNPs are bi-allelic in nature and consequently are not as explanatory on the basis of per-locus as facilitated with forensically chosen STR loci. Since all forensic DNA files, mainly criminal DNA databanks, are well-accepted and established on STR loci, it is improbable that SNPs will be established into major forensic markers. It would take a substantial decrease in expenditure, improved output, and enhanced abilities to regulate various samples for the forensic society to displace the STR loci. With this facts and aspects, it is unfortunate that SNPs will switch STR loci as the dominant genetic markers for the identification of humans and human remains in forensic scenario (Gill et al. 2004). Although, there are plenty of genetic evidences that can be selected; subsequently nearly 85% of human difference is resulting from SNPs (Cooper et al. 1985; Wang et al. 1998; Holden 2002). SNPs, on the other hand, suggest benefits for forensic investigations in certain circumstances, such as utility in disasters victim identification and missing person's identification where the DNA may be significantly fragmented. The output results delivered will be no use as if the remains of the consequent fragments are lesser than the required length for STR typing. Certainly it is possible to develop the PCR for amplification of SNP to reduce the essential fragments length nearly to 60–80 base pairs (Budowle 2004; Budowle et al. 2004; Divne and Allen 2005). In the major outputs of damaged biological samples, SNP assay offers reliable outcomes than STR typing. There are major advantages that majority of the research and progress is continuing to enlarge research abilities, possibly designing large multiplex reactions and whole genome analysis. New developed programs, rapid testing and typing of numerous samples with high impact features can be achieved; SNP foxed tools may progress current capabilities in this respect. As a result SNPs result in moderately low mutation rate and subsequently remained as ancestry-based genetic markers for ancestral studies, missing person's identity, legacy issues and in situations where matching reference sample is poorly available (Budowle and van Daal 2008).

18.3.2 SNPs in Plant DNA Identification

Wild Life DNA forensics is a juvenile branch and is most importantly concerned with the identification and recognization of suspected DNA samples for Species identification. Wildlife cases involves illegal trafficking of wildlife products like processed woods of trees, animal materials present in processed traditional medicine, etc. as they lost their morphological features their identification demands information regarding their species origin (Deguilloux et al. 2002; Hsieh et al. 2003; Wetton et al. 2005; Peppin et al. 2008). Correct identification and individual characterization of plant materials is necessary for the effective conservation of endangered plant and animal species and to confirm their justifiable usages. In the past few years molecular tools established provides easy resources for allocating known and unknown plant species. These newly developed molecular techniques are comparably greater than earlier developed phenotypic tools or methods which may answer many evolutionary and taxonomic questions. Single nucleotide polymorphisms (SNP), microsatellites, amplified fragment length polymorphism (AFLP) and random amplified polymorphic DNA (RAPD) are some of the molecular methods which are recently been used for plant species identifications studies, whereas these techniques have their own limitations and fruitful advantages. These studied molecular techniques differ in their power of resolving to identify the genetic base differences, data generated and applicability towards specific taxonomic levels. SNPs are the most reliable genetic markers both in animal and plant genomes and appeared as one of the mew generation marker for wide applications (Coles et al. 2005; Gómez and Maloo 2009; Miller and Henry 2003).

A recent new technique in molecular identification for plant DNA polymorphism is developed based on micro array platform called as diversity array technology (DArT) (Jaccoud et al. 2001). A novel SNP genotyping technique was described by Van-Eijk et al. (2004) which is based on SNP wave, a chip-based SNP array which uses thousands of oligonucleotide probes attached to a glass, silicon wafer solid surface which allows many SNPs to be examined instantaneously (Rapley and Harbron 2001). The SNaPshot multiplex kit developed by ABI PRISM is designed to read up to 10 single nucleotide polymorphisms (SNPs) at known locations from nearly 10 DNA samples in a single tube. The process includes preparing the sample reactions using DNA template and primer, SNaPshot reaction performance by thermal cycler and treatment of final extension of the PCR products followed by automated gel electrophoresis of the samples and analyze the data.

Populace genetic structure identification of castor bean (*Ricinus communis*) using diverse SNPS from genome wide comparisons showed low genetic diversity level and mixed genotypes, which may lead to nominal geographic structuring of castor bean populations worldwide (Foster et al. 2010). A recent study on single nucleotide primer extension (SNuPE) assay of *gyrB* targeted gene, has reported the utility of gene in categorizing bacteria belongs to the family *Burkholderia cepacia,* which are difficult for identification using common molecular methods (Coenye et al. 2001). The novel SNP-based methods assures successful detection of specific genetic variations in samples within short duration, which indicates potential features of

such methods in medical treatments also detecting successfully and distinct the specific genetic variations which may successfully apply in medical treatments since its positive feature to analyze number of samples in short period (Fierer et al. 2010). Therefore, more number of scientific approaches are required for identification of plant DNA identification, because plants carries low genetic diversity among them.

18.3.3 SNPs in Non-human DNA Identification of Endangered Species

The major factors of non-human animal DNA identification are now genomic markers which indicate polymorphisms at DNA level. Furthermore, due to the potential nature as well as the varying biochemical consequences of different molecular genetics procedures, a range emerges upon which to make decisions as per the convenience of these techniques. On using molecular markers for non-human animal genetic studies, two key issues must be addressed. The genotyping technique must be straightforward and cost-effective to generate significantly large amount of genotype data which is often required from the scientific point of view of most of the molecular biologist.

Somewhat from the statistical point of view, certain features such as the use of relationships of dominance, relevant information quality, neutrality, map positions or genetic independence of markers are vitally important depending on the type of study to be carried out. Regardless of the chosen method, the source data must be as accurate as possible.

Perhaps from the molecular biology point of view, the three chief alternative forms at the DNA level are single nucleotide variations, currently termed SNPs used for single nucleotide polymorphisms; insertions or deletions (Indels) of different lengths vary between 1 to several hundred base pairs and VNTR, for alterations as the number of tandem repeats (Vignal et al. 2002).

The identification of genetic species is established on the separation and exploration of DNA markers which exhibit difference between species, but usually are conserved in species. Within animals, the gene-regions inside mitochondrial DNA, chiefly cytochrome b (Parson et al. 2000) and cytochrome oxidase subunit I (COI) (Hebert et al. 2003a, b), are the very often used markers, as their mutation rate links almost with species evolution.

In wildlife DNA profiling, the option of markers is the similar as for an extensive variety of genetic analyses of the nature and population. Disciplines for instance molecular ecology and genetics of preservation ascertain and employ SNP markers and microsatellite to an extensive variety of species, which can be extremely valuable for wildlife discipline (Linacre 2006).

The most widely used approaches are SNPs for the particular identification of organisms (São-Bento et al. 2009; Kitpipit et al. 2012; Kohnemann and Pfeiffer 2011). The primary species-specified are based on the inclusion, as a matter of principle design, of SNPs in the target animal species, on the restriction of the

fragment length of polymorphism (RFLP), where SNPs may alter the number of identifying sites in a sequence (Cooper and Cooper 2013). One benefit of the use of SNP loci is that mixtures of two or more species can be classified unlike DNA sequence information. Species mixtures, particularly where humankind is a species, regularly prevent the generation of interpretable results by traditional DNA series. The popularity of SNPs in species specific identification can increase the sensitivity of the test with no loss of specificity (Linacre and Tobe 2009).

18.3.4 SNPs in Microbial Species and Bio-geographical Identification

Microorganisms were used to conduct illegal activities, which were in recent times underlined by the terrorist attack on the USA in the fall of 2001, in which letters holding anthrax were used. The answers to these questions are: health and protection of the first respondents, inquiries into contaminated locations, sample collection, transmission of the samples, isolation and detection of the pathogen, and the source trace of the bioagent. Source traceability of a pathogen is important for the advancement of criminal investigations and conviction of the culprit (s) (Keim 2003).

A number of questions had to be answered in connection with anthrax attacks about the relationship and cause of the cases in the four nations, the link and origin of anthrax bacteria isolated in various states, and the identity of a pathogen releaser at home or abroad (Keim 2003). Microbial forensics have been based on the use of microbes as tools for bio-threats or their products (e.g. toxins).

In the last 15 years, the technical developments and application of the immensity and assets of the microbial environment have encouraged the meadow to spread to new extents where microbes and their commodities may help in time since death and even in human identification (Fierer et al. 2010; Pechal et al. 2014) and also support other forms of forensic investigations. In order to extend the use of microbes ahead of bioterrorism and biocrime investigation, a further thorough conception in the discipline of microbial forensics is essential.

The accountability for identifying microbial samples to generate forensic leads in judicial (and civil) cases should now be widely known as microbial forensics. Microbial forensics can use conventional analytical techniques, existing molecular technologies, and modern innovative techniques throughout the course of an investigation into the crime of bioterrorism which can still only be developed to identify and track the pathogen (Keim 2003).

Dr. Randal S. Murch already proposed in 1995 that forensic science should be used to investigate and solve bioterrorism circumstances (Murch 2003). For microbial forensics research laboratory research, molecular sequences, microbial biology, biochemical markers, electron microscopy, crystallography, and mass spectrometry may be involved (Yang and Keim 2012).

The categories of genetic markers, including repeated parts, insertions and deletions, mobile components, islands of virulence factors, genes of virus and tolerance, genes of homekeeping, functional generics, entire genome pieces, asexual and sexual reproduction, horizontal gene transfer, are among the forms of genetically

modifiable genes that may influence microbial forensic proof (Budowle et al. 2005). The extension of the discipline of microbial forensics widens to the field of forensic assay of human individuality. The studies can be pooled to comprise more data for enhanced correlations and stronger likelihoods of eliminating people wrongly identified with biological evidence by using short-tandem repeat (STRs) and SNPs testing (Schmedes et al. 2016). Owing to the related phylogenetic association between several microorganisms, certain molecular genetic markers could not distinguish them. The separation of closed pathogen species can in these cases be made possible by conservative SNPs within realistic coding sequences. Many SNPs have also been discovered for genetic typing by the use of entire genome sequence (Morelli et al. 2010).

18.4 Conclusion

In the molecular scenario, single nucleotide polymorphisms are earlier known to as base substitutions which were abbreviated in the molecular community as SNPs, these types of polymorphisms have always a newer importance in molecular biology too. Certainly, in particular cases the facts generated from the bi, tri or tetra-allelic character of SNPs is limited; there seem to be instances whereby they can specify important facts and information that associate relating exact genes or population genetic structures or genetic organization and phenotypes. The high region/segment of SNPs in genome structure allows expanding some of these bp in a particular locus of a few hundreds of base pairs. After restructuring the resulting haplotypes, multi-allelic system can finally be used for studies, to prevail over the drawbacks due to SNPs heterozygote limitations. With the progressive development in molecular biological techniques to create SNP data for genetic analysis, genetic softwares and other progressive algorithms together will always be helpful.

References

Budowle B (2004) SNP typing strategies. Forensic Sci Int 146:S139–S142

Budowle B, van Daal A (2008) Forensically relevant SNP classes. Biotechniques 44(5):603–608

Budowle B, Planz J, Campbell R, Eisenberg A (2004) SNPs and microarray technology in forensic genetics: development and application to mitochondrial DNA. Forensic Sci Rev 16:22–36

Budowle B, Johnson MD, Fraser CM, Leighton TJ, Murch RS et al (2005) Genetic analysis and attribution of microbial forensics evidence. Crit Rev Microbiol 31:233–254

Coenye T, Mahenthiralingam E, Henry D et al (2001) Burkholderia ambifaria sp. nov.: a novel member of the Burkholderia cepacia complex comprising biocontrol and cystic-fibrosis related isolates. Int J Syst Evol Microbiol 51:1481–1490

Coles ND, Coleman CE, Christensen SA, Jellen EN, Stevens MR, Bonifacio RA, Rojas-Beltran JA, Fairbanks DJ, Maughan PJ (2005) Development and use of an expressed sequenced tag library in quinoa (Chenopodium quinoa Willd.) for the discovery of single nucleotide polymorphisms. Plant Soc 168:439–447

Cooper JE, Cooper ME (2013) Wildlife forensic investigation: principles and practice. CRC Press, Boca Raton

Cooper DN, Smith BA, Cooke HJ, Niemann S, Schmidtke J (1985) An estimate of unique DNA sequence heterozygosity in the human genome. Hum Genet 69:201–205

Deguilloux M, Pemonge M, Petit R (2002) Novel perspectives in wood certification and forensics: dry wood as a source of DNA. Proc R Soc Lond B Biol Sci 269(1495):1039–1046

Divne AM, Allen M (2005) ADNA microarray system for forensic SNP analysis. Forensic Sci Int 154:111–121

Fierer N, Lauber CL, Zhou N, McDonald D, Costello EK, Knight R (2010) Forensic identification using skin bacterial communities. Proc Natl Acad Sci U S A 107:6477–6481

Foster JT, Allan GJ, Chan AP, Rabinowicz PD, Ravel J, Jackson PJ, Keim P (2010) Single nucleotide polymorphisms for assessing genetic diversity in castor bean (Ricinus communis). BMC Plant Biol 10:13–23

Gad SC (2007) Handbook of pharmaceutical biotechnology. Wiley, Hoboken

Gill P, Werrett DJ, Budowle B, Guerrieri R (2004) An assessment of whether SNPs will replace STRs in national DNA databases—joint considerations of the DNA working group of the European Network of Forensic Science Institutes (ENFSI) and the Scientific Working Group on DNA Analysis Methods (SWGDAM). Sci Justice 44:51–53

Gómez JMJ, Maloo JN (2009) Sequence diversity in three tomato species: SNPs, markers and molecular evolution. BMC Plant Biol 9:85

Hebert PDN, Ratnasingham S, de Waard JR (2003a) Barcoding animal life: cytochrome oxidase subunit 1 divergences among closely related species. Proc R Soc Lond Ser B Biol Sci 270 (suppl_1)

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003b) Biological identifications through DNA barcodes. Proc R Soc Lond B Biol Sci 270(1512):313–321

Holden AL (2002) The SNP consortium: summary of a private consortium effort to develop an applied map of the human genome. Biotechniques 32:S22–S26

Hsieh HM, Huang LH, Tsai LC, Kuo YC, Meng HH, Linacre A, Lee JC (2003) Species identification of rhinoceros horns using the cytochrome b gene. Forensic Sci Int 136(1–3):1–11

Hwang MT, Wang Z, Ping J, Ban DK, Shiah ZC, Antonschmidt L, Lee J, Liu Y, Karkisaval AG, Johnson ATC, Fan C, Glinsky G, Lal R (2018) DNA nanotweezers and graphene transistor enable label-free genotyping. Adv Mater 30(34):1802440

Jaccoud D, Peng K, Feinstein D, Kilian A (2001) Diversity arrays: a solid state technology for sequence information independent genotyping. Nucleic Acids Res 29:25e

Keim P (2003) Microbial forensics: a scientific assessment. A Report from American Academy of Microbiology

Kohnemann S, Pfeiffer H (2011) Application of mtDNA SNP analysis in forensic casework. Forensic Sci Int Genet 5:216–221. https://doi.org/10.1016/j.fsigen.2010.01.015

Kitpipit T, Tobe SS, Linacre A (2012) The complete mitochondrial genome analysis of the tiger (Panthera tigris). Mol Biol Rep 39:5745–5754

Linacre A (2006) Application of mitochondrial DNA technologies in wildlife investigations—species identification. Forensic Sci Rev 18(1):1–8

Linacre A, Tobe SS (2009) Species identification using DNA loci. Taylor & Francis Group, London

Miller PT, Henry R (2003) Single-nucleotide polymorphism detection in plants using a single-stranded pyrosequencing protocol with a universal biotinylated primer. Anal Biochem 317:166–170

Morelli G, Song Y, Mazzoni CJ, Eppinger M, Roumagnac P et al (2010) Yersinia pestis genome sequencing identifies patterns of global phylogenetic diversity. Nat Genet 42:1140–1143

Murch RS (2003) Microbial forensics: building a national capacity to investigate bioterrorism. Biosecur Bioterror 1:117–122

Okimoto R, Dodgson JB (1996) Improved PCR amplification of multiple specific alleles (PAMSA) using internally mismatched primers. Biotechniques 21(1):20. 24, 26

Parson W, Pegoraro K, Niederstatter H, Foger M, Steinlechner M (2000) Species identification by means of the cytochrome b gene. Int J Legal Med 114:23–28

Pechal JL, Crippen TL, Benbow ME, Tarone AM, Dowd S, Tomberlin JK (2014) The potential use of bacterial community succession in forensics as described by high throughput metagenomic sequencing. Int J Legal Med 128:193–205

Peppin L, McEwing R, Carvalho GR, Ogden R (2008) A DNA based approach for the forensic identification of Asiatic black bear (Ursus thibetanus) in a traditional asian medicine. J Forensic Sci 53:1358–1362

Rapley R, Harbron S (2001) Molecular analysis and genome discovery. Wiley, Chichester

Sanchez JJ et al (2003) Multiplex PCR and minisequencing of SNPs—a model with 35 Y chromosome SNPs. Forensic Sci Int 137(1):74–84

São-Bento M, Carvalho M, Bento AM, Andrade L, Lopes V, Serra A, Balsa F, Anjos MJ, Corte-Real F (2009) Y-chromosome SNP analysis in the Brazilian population of São Paulo state (Ribeirão Preto). Forensic Sci Int Genet Suppl Ser 2(1)):427–428

Schmedes SE, Sajantila A, Budowle B (2016) Expansion of microbial forensics. J Clin Microbiol 54(8):1964–1974

Sobrinoa B, Brion M, Carracedoa A (2005) SNPs in forensic genetics: a review on SNP typing methodologies. Forensic Sci Int 154:181–194

Twyman RM (2005) Single nucleotide polymorphism (SNP) genotyping techniques—an overview. Encyclopedia of Diagnostic Genomics and Proteomics

Van-Eijk MJT, Broekhof JLN, van der Poel HJA, Hogers RCJ, Schneiders H, Verstege KJE, van Aart JW, Geerlings H, Buntjer JB, van Oeveren AJ, Vos P (2004) SNPWave: a flexible multiplexed SNP genotyping technology. Nucleic Acids Res 32:e47

Vignal A, Milan D, Sancristobal M, Eggen A (2002) A review on SNP and other types of molecular markers and their use in animal genetics. Genet Sel Evol 34:275–305

Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolshy R, Chandona G, Perkins N et al (1998) Large scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science 280:1077–1082

Wetton JH, Tsang KW, Khan H (2005) Inferring the population of origin of DNA evidence within the UK by allele-specific hybridization of Y-SNPs. Forensic Sci Int 152:45–53

Yang R, Keim P (2012) Microbial forensics: a powerful tool for pursuing bioterrorism perpetrators and the need for an international database. J Bioterr Biodef S3:007. https://doi.org/10.4172/2157-2526.S3-007

Part IV

Non-human Forensic DNA Typing

19 DNA Analysis of Domestic Animals

Kamlesh Kaitholia, Pushpesh Kushwaha, Manisha Rana, Ila Gautam, and Ankit Srivastava

Abstract

The chapter describes utility and advancements in forensic DNA analysis of nonhuman evidences. Published markers and presently utilized molecular techniques for analyzing nonhuman DNA analysis are described in detail. The chapter also presents an extensive review of literature available on forensic DNA typing of domestic animals.

Keywords

STR · DNA · Forensic · Nonhuman · DNA markers

K. Kaitholia (✉) · I. Gautam
DNA Fingerprinting Unit, State Forensic Science Laboratory, Department of Home (Police), Govt. of MP, Sagar, Madhya Pradesh, India

P. Kushwaha
DNA Fingerprinting Unit, State Forensic Science Laboratory, Department of Home (Police), Govt. of MP, Sagar, Madhya Pradesh, India

Dr. APJ Abdul Kalam Institute of Forensic Science and Criminology, Bundelkhand University, Jhansi, Uttar Pradesh, India

M. Rana
DNA Fingerprinting Unit, State Forensic Science Laboratory, Department of Home (Police), Govt. of MP, Sagar, Madhya Pradesh, India

Department of Forensic Science, Guru Ghasidas University, Bilaspur, Chhattisgarh, India

A. Srivastava
Dr. APJ Abdul Kalam Institute of Forensic Science and Criminology, Bundelkhand University, Jhansi, Uttar Pradesh, India

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_19

19.1 Introduction

DNA typing has emerged as an imperative technique in forensic investigation for human identification since decades and is now augmenting its application toward nonhuman genetic analysis. Among various developments and landmarks in the field of forensic genetics, recent concept of animal DNA analysis is emerging as an advantageous, accessible, and decisive technique (Oldt et al. 2016). Examination of nonhuman (or animal)-based evidence present at crime scenes can be an influential source of information related with the crime, victim, or perpetrator (Cassidy and Gonzales 2005). Genetic evidences from domestic animals are substantial for identification and individualization purposes in cases where such animals are either the suspect (as is cases of animal attacks, road accidents) or victims (as in animal theft and animal abuse). Applications also expand to food contamination, illicit trading, and sexual crimes (Fig. 19.1). Genetic analysis of animal evidence can be useful in constructing links among victims, perpetrators, and/or crime scenes.

Duguay (1994) of Prince Edward Island, Canada was the first reported case in which genetic testing of animal biological materials served as the prime evidence. A leather jacket stained with victim's blood was found at the woods of the island. However, there were no evidences that could establish the link of the key suspect (victim's divorced husband) to the leather jacket. Numerous strands of nonhuman white hair embedded on the jacket were collected and their genetic analysis concluded that the white hairs belonged to the specific cat owned by the suspect. Based on the genetic evidence from cat's hair, link between the jacket, suspect, and victim was confirmed and the suspect was convicted of murder (Cassidy and Gonzales 2005).

With an increment in the presentation of animal evidence in court, standardization and generation of proper guidelines are expected for assurance of the admissibility of the DNA reports in the court. In Cases involving animals, adequacy of DNA to be accepted and applicable in the court of law requires precise affirmation of allele

Fig. 19.1 Varied applications of forensic DNA analysis in domestic animals

transmission and independence, as well as accuracy in evidence collection, handling, analysis, and generation of result. But, very few laboratories are capable of analyzing these samples (Cassidy and Gonzales 2005; Kanthaswamy 2015).

19.2 Techniques Used in Animal Forensics

The classical technique of species identification was based on the morphological, osteological, and microscopical analysis of the parts of animal. Detailed observation of the phenotypic markers such as thickness and color of skin coat and tail, color of eye, muzzle, dental pattern and dental arch, examination of skeletal remains, and microscopical examination of animal hair for morphological structure of hair and cuticular scale pattern emerged as suitable approach for species identification (Nishant et al. 2017) (Table 19.1). It requires a well-developed reference database and skilled expertise for accurate comparison and identification. Therefore, this method is quite unreliable in absence of proper database and knowledge. Also, individual identification is unattainable with morphological analysis (Panday et al. 2014). Species screening and identification of the animal product through serological techniques, such as ELISA, immunodiffusion assay, are based on the antigen–antibodies interaction. Such assays were helpful mostly for exclusion of samples due to unavailability of species-specific antibody for all the animals. Also, cross-reactions in nontarget species in certain cases provide unreliable result (Panday et al. 2014).

Genetical markers can provide a more accurate and reliable results. DNA-based assays like restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), short tandem repeat (STR), and single nucleotide polymorphism (SNP) provide a unique profile for individual animal. However, RFLP, AFLP, RAPD involves laborious, time-consuming procedure, and are not always reproducible (Table 19.2). Recent methods of STR and SNP provide comparatively satisfactory results in the case of mixed samples. Frequent occurrence and high degree of polymorphism within genome and discriminating power of STRs make it a suitable marker in the identification of the animals (Somanathan and Mathur 2017). SNPs are preferentially used in the case of degraded samples due to the requirement of smaller fragments of DNA (Sobrino et al. 2005).

19.2.1 Current Techniques Used for Forensic DNA Analysis in Animals

Proper collection, preservation, and transportation of biological specimens prior to laboratory analysis are demanded. Genetic analysis in the laboratory begins with isolation of DNA from the other cellular components and biomolecules. Purification of DNA from substances that may inhibit further processes into suitable form and proper concentration are required for processing the methods of DNA analysis

Table 19.1 Conventional tools and techniques for forensic analysis of domestic animals

Techniques	Applications
Morphological analysis	Examination of morphological features such as skin coat, eyes, tails, muzzles, dental patterns, etc. and bone examination for species identification
Footprint analysis	Useful in species identification, age estimation, and sometimes in individual identification through unique features in pugmark pattern
Bitemark analysis	Determination of dental features such as intercanine distance, dental arch shape, teeth morphology, malalignment may provide a lead in the species identification
Microscopic analysis	Morphological examination, scale pattern identification, and elemental analysis of hair which are the most common evidence found at the crime scene through scanning electron microscope
Serological analysis	Antigen–antibody interaction-dependent techniques such as ELISA, immunodiffusion assays, western blot assay, etc. are useful in species identification and diagnosing infectious diseases
Molecular analysis	DNA polymer-based techniques are considered to be reliable for species as well as individual identification. Determination of age, sex, and parental lineage is also possible through a variety of techniques such as RFLP-PCR, STR, SNP, mtDNA, and DNA methylation. Protein polymer-based markers are being focused on better discrimination in animals
Infrared thermography	Analysis of body temperature to diagnose physiological status of animal—detection of any disease, pregnancy, feeding efficiency, and stress response
Radioisotopic analysis	Analysis of radioisotopic signature and quantity of stable isotopes and trace elements within the skeletal remain is useful in determination place of origin, and geographic mobility as well as estimation of time since death
Inductively coupled plasma–atomic emission spectroscopy (ICP-AES)	Trace element signature in skeletal remains is examined to determine their origin and differentiation according to species, domestication status, and feeding type

discussed in this section. It is then followed by amplification, genotyping, and report generation (Fig. 19.2). The extensive list of genetic makers useful for DNA analysis of domestic animals has been given in Table 19.3.

Table 19.2 Comparative account of various molecular techniques for forensic analysis of domestic animals

Characteristics	RFLP	RAPD	AFLP	STR	SNP
Principle	Includes digestion with restriction enzyme and hybridization with probe	Random primers for PCR amplification	Analysis of heterosite restriction fragment in the genome by PCR amplification	Based on tandem repeat flanking sequence	Difference in DNA sequence due to single base–pair mutation
Classification	Hybridization based	PCR-based	PCR-based	PCR-based	Sequence based
Genomic abundance	Moderate	Moderate	Moderate	Moderate	Highest
Level of polymorphism	Moderate	Moderate	Moderate	Low	Moderate
Dominance	Codominance	Dominance	Dominance	Codominance	Codominance
Locus specificity	Yes	No	No	Yes	Yes
Loci per assay	1 to few	Many	Many	1–20	1–1000
Required DNA quantity	High	Low	Medium	Low	Low
Reproducibility	High	Low	Mid–high	High	High
Used restriction enzyme	Yes	No	No	No	No
Laborious	High	Low	Medium	Low	Low
Automation	No	Yes	Yes	Yes	Yes
Specialized equipment	Radioactive isotopes	Agarose gel	Polyacrylamide gel	Agarose/ polyacrylamide gel	Agarose gel

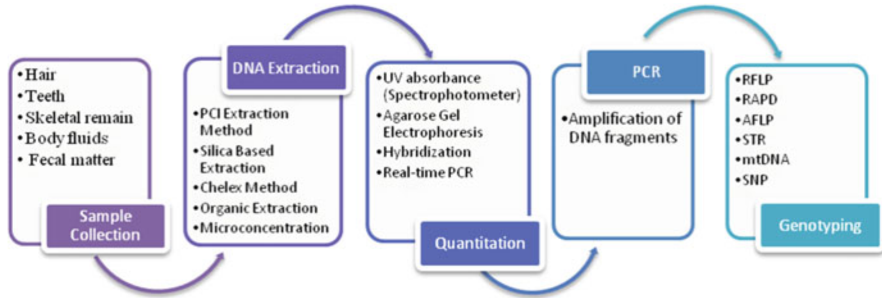


Fig. 19.2 Steps followed for DNA analysis in domestic animals

19.2.2 Sample Collection

Hair, skin, bones and teeth, body fluids such as blood, saliva, semen, and feces are the commonly encountered biological evidence related with pet animals. Collection, preservation, storage, and transportation of biological evidences for laboratory examination are important steps in genetic analysis of pet animals. Negligence during collection and handling of multiple biological samples may often result in cross-contamination resulting in deterioration ultimately debasing the accuracy of analysis. Thus, proper collection and handling are necessary for accurate results (Gaur and Reddy 2017).

19.2.3 DNA Extraction

Isolation of DNA from the collected samples is done through several methods that include PCI extraction method, silica-based extraction, Chelex method, organic extraction, and microconcentration to separate DNA from other non-DNA organic or inorganic PCR-inhibiting material (Lincoln and Thompson 1998).

19.2.4 DNA Quantitation

The isolation step is followed by quantitation of the extracted DNA. The quality and quantity of the isolated DNA are checked through various methods such as UV absorbance-based NanoDrop spectrophotometer, agarose gel electrophoresis, hybridization, and real-time PCR (Gaur and Reddy 2017).

19.2.4.1 Amplification

Extracted DNA fragments are amplified exponentially using polymerase chain reaction (PCR). DNA can be amplified from a very small quantity of sample. The method was developed in 1983 by Kary Mulis and is based on the principle of naturally occurring process of DNA replication during cell division where each

Table 19.3 List of genetic markers useful for DNA analysis of domestic animals

Animals	Locus	Chromosomal location	References
Cattle	BM1818	D23S21	Bishop et al. (1994)
	BM1824	D1S34	Barendse et al. (1994)
	BM2113	D2S26	Sunden et al. (1993)
	CSRM60	D10S5	Baylor College of Medicine Human Genome Sequencing Center (2006)
	CSSM66	D14S31	Barendse et al. (1994)
	ETH3, ETH10	D19S2, D5S3	Solinas-Toldo et al. (1993)
	ETH225	D9S2	Steffen et al. (1993)
	HAUT27	D26S21	Thieven et al. (1997)
	ILSTS006	D7S8	Brezinsky et al. (1993)
	INRA023	D3S10	Vaiman et al. (1994)
	SPS115	D15	Baylor College of Medicine Human Genome Sequencing Center (2006)
	TGLA53, TGLA122, TGLA126, TGLA227	D16S3, D21S6, D20S1, D18S1	Georges and Massey (1992)
	BM2113, BM1824, SPS115, ETH3, ETH125, ETH10, TGLA227, TGLA126, TGLA122		Vázquez et al. (2004)
	STRs: FCB19, FCB, BM410, RM00, TGLA12, BM181, TGLA9, ETH1, ETH15, BM20, ADCY, SPS11, BM182, INRA2, RM06, TGLA12, BR, BM420, ETH22, TGLA221, BM88, BM211, CYP2 **SNP**: BTA_AY849380, BTA_AY863214, BTA_AY842475, BTA_AY939849 BTA_AY842473, BTA_AY860426, BTA_AY842474, BTA_AY851163, BTA_AY916666, BTA_AY773474, BTA_AY942198, BTA_AY842472, BTA_AY914316,		Eenennaam et al. (2007)

(continued)

Table 19.3 (continued)

Animals	Locus	Chromosomal location	References
	BTA_AY853302, BTA_AY853303, BTA_AY761135, BTA_AY850194, BTA_AY857620, BTA_AY941204, BTA_AY943841, BTA_AY856094, BTA_AY841151, BTA_AY851162, BTA_AY919868, BTA_AY929334, BTA_AY858890, BTA_AY937242, BTA_AY776154		
	ILSTS005, ILSTS006, ILSTS011, ILSTS030, ILSTS033, ILSTS034, ILSTS054, INRA005, INRA032, INRA035, INRA063, ETH003, ETH010, ETH152, ETH225, HEL001, HEL005, HEL009, BM1818, BM2113, HAUT024, HAUT027, CSRM60, CSSM66, MM008		Ganapathi et al. (2012)
	BM1824, CSSM08, CSSM33, CSSM66, ETH10, ETH225, ETH3, HEL09, HEL5, ILSTS06, ILSTS11, ILSTS34, ILSTS33, INRA05, INRA35, INRA63, MM12, MM8, TGLA122, TGLA227, TGLA53		Sharma et al. (2015)
	BM1824, CSSM08, CSSM33, CSSM60, CSSM66, ETH3, ETH10, ETH225, HAUT27, HEL1, HEL5, HEL9, ILSTS005, ILSTS006, ILSTS011, ILSTS033, ILSTS034, INRA05, INRA63, INRA35, MM8, MM12, TGLA53, TGLA122, TGLA227		Sodhi et al. (2011)
	BM1818, BM1824, BM2113, CSRM60,	D23S21, D1S34, D2S26, D10S5, D14S31, D19S2,	Van de Goor et al. (2011)

(continued)

Table 19.3 (continued)

Animals	Locus	Chromosomal location	References
	CSSM66, ETH3, ETH10, ETH225, HAUT27, ILSTS006, INRA023, SPS115, TGLA53, TGLA122, TGLA126, TGLA227	D5S3, D9S2, D26S21, D7S8, D3S10, D15, D16S3, D21S6, D20S1, D18S1	
Horse	AHT4, AHT5	24q14, 8	Binns et al. (1995)
	ASB2, ASB17	15q21.3-q23, 2p14-p15	Breen et al. (1997)
	ASB23	3q22.1-q22.3	Irvin et al. (1998)
	CA425 UCDEQ425	28q18	Eggleston-Stott et al. (1997)
	HMS1, HMS2, HMS3, HMS6, HMS7	15, 10, 9, 4, 1q25	Guérin et al. (1994)
	HTG4, HTG6	9, 15q26-q27	Ellegren et al. (1992)
	HTG7, HTG10	4, 21	Marklund et al. (1994)
	LEX3	Xq	Coogle et al. (1996)
	VHL20	30	Van Haeringen et al. (1994)
	VHL20, HTG4, AHT4 HMS7, Lex3, CA425 HTG10, HTG7, HMS3, CA412, HMS2, HTG6, AHT5, HMS6, ASB2		Buitkamp et al. (1999)
	AHT4, AHT5, ASB2, ASB17, ASB23, CA425, HMS1, HMS2, HMS3, HMS6, HMS7, HTG4, HTG6, HTG7, HTG10, LEX3, VHL20		Seo et al. (2016)
Dog	PEZ01, PEZ03, PEZ05, PEZ20, FH2010, FH2079	7, 19, 12, 22, 24, 24	StockMarks kit (n.d)
	PEZ05, PEZ16, PEZ17, PEZ21, FH2001, FH2004, FH2010, FH2107, FH2309, FH2328, FH2361, FH3313, FH3377, VWF.X, PEZ02	12, 27, 4, 2, 23, 11, 24, 3, 1, 33, 33, 19, 3, 27, 17	Tom et al. 2010
	PEZ15, FH2079	16, 24	Berger et al. (2009)
	FH2004, FH2010, FH2361, FH2658, FH3210, FH3241, FH4012, REN214L11, C38, IPAX1, IPAX2, IPAX3, IPAX4	11, 24, 33, 14, 2, 8, 15, 16, 38, X, X, X, X,	van Asch et al. (2009)
	VGL0760, VGL0910, VGL1063, VGL1165, VGL1541, VGL1606, VGL1828, VGL2009,	7: 60065445, 9: 10224058, 10: 63191724, 11: 65356234, 15: 41210435, 16: 6468079, 18: 28419883,	Wictum et al. (2012)

(continued)

Table 19.3 (continued)

Animals	Locus	Chromosomal location	References
	VGL2136, VGL2409, VGL2918, VGL3008	20: 9290711, 21: 36673167, 24: 9197210, 29: 18216971, 30: 8845920	
	FH2848, AHTh171, REN162C04, INU055, AHT137, FH2054, INU030, AHT121, REN169D01, REN247M23, AHTh260, REN54P11, INRA21, CXX279, AHTk253, AHTk211, REN169O18, INU005, Amelogenin	2, 6, 7, 10, 11, 12, 12, 13, 14, 15, 16, 18, 21, 22, 23, 26, 29, 33, X	Ciampolini et al. (2017)
Camel	YWLL08, YWLL09, YWLL38, YWLL44, YWLL59, VOLP03, VOLP08, VOLP10, VOLP32, VOLP67, LCA66, CVRL01, CVRL05, CVRL06, CVRL07, CMS13, CMS17, CMS18, CMS25, CMS50, CMS121	N/A	Mahmoud et al. (2019)
	YWLL08, YWLL09, YWLL38, YWLL44, YWLL59, VOLP03, VOLP08, VOLP10, VOLP32, VOLP67, LCA66, CVRL01, CVRL05, CVRL06, CVRL07, CMS13, CMS17, CMS18, CMS25, CMS50, CMS121	N/A	Mahmoud et al. (2020a, b)
Goat	ILSTS005, MCM527, SRCRSP5, OarFCB128, HUJ616, OarHH47, ILSTS11 DYMS1, BM8024, OarFCB226, OarAE129, OarJMP29, SRCRSP9, MAF214, OarCP34, OarFCB304, MAF209, MAF65	N/A	Mahmoud et al. (2020a, b)
	MAF70, INRA023, SPS113, CSRD247, McM527, ILSTS087, BM6444, P19 (DYA), TCRVB6, DRBP1, ETH10	4, 3, 10, 14, 5, 28, 2, 20, unknown, 23, 5	Bulut et al. (2016)
Pig	S0155, CGA, S0226, SW240, S0002, SW72,	N/A	Charoensook et al. (2019)

(continued)

Table 19.3 (continued)

Animals	Locus	Chromosomal location	References
	S0227, IGF-1, S0005, SW122, S0101, SW632, S0225, SW911, SW951, S0386, S0090, S0068, S0215, SW857, S0355, SW936, S0026, SW1031, S0120, S0218		
	S0155, SW72, S0005, SW122, SW632, S0225, SW951, S0090, SW857, SW936, S0026, SW24, SW787	1, 3, 5, 6, 7, 8, 10, 12, 14, 15, 16, 17, 18	Oh et al. (2014)
Sheep	BM757, BM827, BM1314, BM6506, BM6526, BM8125, CSRD247, CSSM31, CSSM47, HSC, INRA63, MAF214, OarAE129, OarCP20, OarCP34, OarCP49, OarFCB48, OarFCB128, OarHH35, OarHH41, OarHH47, OarHH64, OarJMP8, OarJMP29, OarVH72	N/A	Arora et al. (2011)
	BM757, BM 827, BM6526, BM 8125, CSSM 47, OarAE 129, OarCP 34, OarFCB48, OarHH 35, Oar JMP8, OarJMP29	N/A	Mukesh et al. (2006)
Rabbit	ATP12A, CYTC, MGST3, PRL, STAG1, EXT1, LUM, T, TIAM1, UD14	8, 4, 13, 12, 14, 3, 4, 12, 14, 7	Carneiro et al. (2008)
Buffalo	ETH 10, INRA 5, ILST 029, ILST033, ILST049, ILST052, ETH225, CSSM66	5, 12, 3, 12, 11, 21, 9, 14	Hussain et al. (2017)
	CSRM60, ILSTS026, HEL13, ILSTS030, ILSTS033, ILSTS017, ILSTS019, ILSTS045, ILSTS034, ILSTS058, ILSTS056, ILSTS089, CSSM66, ILSTS036, ILSTS095, ILSTS029, ILSTS028, ILSTS025, ILSTS052, ILSTS031, ILSTS073, ILSTS060, BM1818, ILSTS061, ILSTS068	N/A	Mishra et al. (2007)

(continued)

Table 19.3 (continued)

Animals	Locus	Chromosomal location	References
Cat	F53, FCA723, FCA731, F85, FCA733, FCA740, FCA441, FCA742, F124, FCA749	A1, A1, B1, B1, B2, C1, D3, D4, E1, F2	Menotti-Raymond et al. (2012)
	ATP12A, CYTC, EXT1, LUM, MGST3, PRL, STAG1, TIAM1, UD14, DIAPH2, F9, GPC4, KLHL4, MAOA, PGK1, POLA1	N/A	Carneiro et al. (2011)
	F53, C08, B04, G11, FCA441, D09, F124, C12, C09, F85, D06, SRY	A1, B2, A1, B1, D3, B4, E1, F2, D4, B1, C1, Y	Butler et al. (2002)
	F27, F37, F42, F141, FCA441, FCA559, FCA730, FCA733, FCA734, FCA740, FCA742, FCA744, FCA747, FCA749	N/A	Schury et al. (2014)

strand of DNA acts as template for generation of a new complementary strand. During the process of PCR, the DNA sample to be amplified is subjected to high temperature (approximately 94 °C) for denaturation and separation of DNA strands followed by incubation with DNA polymerase (Taq polymerase), dNTPs, and two oligonucleotide primers. The temperature is then decreased to 55–60 °C for annealing of the primers to the template DNA strand. These primers are short sequences of nucleotide complementary to template DNA located at the 3′ end of both the strands of desired DNA fragment. Subsequent rise in temperature to 72 °C favors activity of DNA polymerase in extension of complementary strand with the help of dNTPs. The primers are oriented facing each other and their position on the template DNA governs the length of the amplified fragments (Linacre and Tobe 2013).

19.2.4.2 Genotyping

The amplified product is then subjected to genotyping that is processed by analyzing the DNA sequence. Various techniques of genotyping including restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), short tandem repeat (STR), and single nucleotide polymorphism (SNP) have been introduced.

Restriction Fragment Length Polymorphisms (RFLP)

Restriction fragment length polymorphisms (RFLPs), developed in 1974, were the first practiced technique for identifying variation within DNA. Such variation in DNA occurs either due to sequence variation or length variation and is detected by the use of restriction enzymes that cleaves DNA strand at specific site. Southern

blotting is used for fragment detection. It includes electrophoretic separation of cleaved DNA fragments on the gel followed by their transfer onto a nylon membrane and finally detection using probe hybridization (Linacre and Tobe 2013). Constant result over time and location due to generation of fragments from specific sites by action of known restriction enzymes explains the high reliability of this technique. Further, high degree of reproducibility, detection of codominance, and selective neutrality add to its advantages. However, extensive labor demand and time consumption, low polymorphisms are its major drawbacks. Also, the process confines detection approach to specific mutations that emerge as the major hindrance in the identification of whole-genome variation in animals (Yang et al. 2013).

Random Amplification of Polymorphic DNA (RAPD)

Random amplification of polymorphic DNA (RAPD) is a PCR-based technique developed in 1990. It is based on the annealing of random oligonucleotide primers at complementary multiple sites on the genomic DNA and their amplification between two priming sites. Designing of specific primer does not demand any prior knowledge of sequence and sequences can be used in various templates. In comparison to RFLP, RADP is a smooth, rapid, and cost-effective technique requiring very small amount of DNA. But, poor reproducibility and reliability due to nonspecific annealing of primers and failure in detection of codominance in F2 generation limit its application (Yang et al. 2013).

Amplification of Fragment Length Polymorphisms (AFLP)

Amplification of fragment length polymorphism (AFLP) introduced in 1993 is based on the combination of restriction enzyme cleaving and PCR amplification. AFLP is considered as a more reproducible advancement of RAPD. Digestion of genomic DNA using restriction enzymes results in numerous DNA fragments called heterosite restriction fragments. These fragments with known sequences are ligated to double-stranded adapter with known sequence at the complementary enzyme-specific cohesive ends or overhangs. Selective amplification of these fragments is carried out using PCR (Linacre and Tobe 2013). Nonspecific amplification in RAPD is resolved in the process of AFLP, increasing its reliability. AFLP also surmounts the tedious, labor-intensive process of RFLP techniques. High sensitivity for high-quality DNA and difficulty in the analysis of impure sample minimize its suitability and certainty for forensic caseworks.

Short Tandem Repeats (STR)

These are natural repeated fragments present in eukaryotic genomes, consisting of 1–6 base pairs (bp) which are repeated again and again following a similar pattern. Its advantages include: (a) They require very less quantity of DNA. (b) Their polymorphism is high. (c) They are codominant markers. (d) They have a better accuracy and higher reproducibility. (e) Different markers can be analyzed in PCR. The basic disadvantages of STRs are: (a) They are time-taking and very expensive. (b) When null alleles appear, due to mutations heterozygotes may be misjudged as homozygotes. (c) Accurate polymorphism may be complicated due to appearance of

stutter bands. (d) Less information about the mutation rates. (e) Information regarding functional trait biodiversity is not provided by them (Yang et al. 2013).

Mitochondrial DNA (mtDNA) Markers

The mitochondrial genome constitutes of an oblique DNA chromosome. Mainly, 36 or 37 genes are present in animal mtDNA: 22 in tRNAs, 2 in rRNAs, and 12 or 13 as multimeric proteins subunits in internal membrane. Further, a noncoding array is also present which plays an important role in transcription and replication of mtDNA fragments and is termed as the control region (CR). The basic property of mtDNA is that its protein is histone-free and it has narrow repairability. Due to this, it has a comparably high mutation fixation rate than nuclear DNA (Arif et al. 2011). It has been categorized into three groups: mitochondrial protein-coding genes, ribosomal DNA, and noncoding fragments. 12S and 16S rRNA segment, cytochrome b (Cytb) gene, cytochrome c oxidase subunit I (COI) gene, and control region (D loop) are majorly used for animal species identification purpose.

Single Nucleotide Polymorphism (SNP)

Developed in 1996, marker of SNP (single nucleotide polymorphism) is a change in single base of a DNA sequence, having substitutes of two available nucleotides. This type of polymorphism consists of single base transversions, transitions, deletions, and insertions. The minor allele frequency shall be equivalent to 1% or greater than this (Kanthaswamy 2015). Presently, these markers are commonly preferred genotyping techniques. It works on the principle of hybridizing exposed fragments of DNA with SNP chips, and then depending on hybridization results, this allele is named. They portray a definite polymorphism in only two alleles of the population as these are biallelic markers. The process of SNP includes amplification of a section of DNA containing that particular SNP or its loci. The fragment to be amplified should be of short length as SNP testing works on highly degraded DNA. During analysis, in SNaPshot™ test, unused free trinucleotides (dNTPs) and the initial unintegrated PCR primers needed to be removed. A couple of enzymes are added at the same time: exonuclease and SAP. The former digests single-stranded DNA and the latter digests the free bases. Only modified bases are used (ddNTPs) in the SNaPshot™ test. As a result, as soon as a single base is added the reaction is aborted. The modified base will be added just next to the SNP base. During sequencing, a dye is attached to the modified base in order to the determination of four bases which has been supplemented. Unlike PCR, this is a definite reaction, as only a single product is contrived in each reaction (Linacre and Tobe 2013). To visualize SNP products, generally 25 cycles are needed. In the process of species evolution and population genetic variations, SNP plays an important role because they are dispersed in both regions of genomes (Vignal et al. 2002).

19.2.5 DNA Sequencing

The Sanger sequencing was started in 1977. It is the process of DNA replication where one strand operates as the template for generating new DNA strand that occurs within the cells. It uses a little volume of both normal bases as well as modified bases (Linacre and Tobe 2013). Sequencer provides an automatic and specific process for genotyping STR and SNP loci simultaneously from a single amplification step of multiplexed PCR. Real-time quantitative PCR (qPCR) based techniques fit the requirements of specificity and sensitivity and are commonly used in human forensics for the purpose of quantification currently. Very less amount of DNA can be amplified and sequenced. Through sequencing, maximum information of polymorphic nature can be dogged. This is due to qPCR technology, which empowers the detailed evaluation of templates of nuclear DNA, which further establishes the success of next nuclear STR and SNP analyses. The same mechanization has also been practiced in quantification of nuclear DNA of domestic animals (Kanthaswamy 2015).

19.3 Future Perspectives

The advancing field of genetics has witnessed progression in sequencing technology. Among the predominant genetic techniques, DNA barcoding is a novel system for identification of species by the comparison of sequence of DNA of one species related with a specific gene or "barcode" to the DNA sequence of the same gene in other species (Bock and Norris 2016). The most commonly used barcode for animal identification is mitochondrial cytochrome c oxidase subunit I (Yang et al. 2018). Whole-genome sequencing is supposed to be another advancing technique in the field of forensic genetics. Mass parallel sequencing or commonly known as next-generation sequencing is a breakthrough that has overcome extremely lengthy and laborious process of Sanger sequencing. Whole-genome sequencing through NGS can reveal comprehensive data among different pet animal's genetic variation as all the variations can be detected within the genome (Yang et al. 2013). A precise, high-throughput sequence data with longer read length can be produced in short time frame (Dunisławska et al. 2017) (M). DNA methylation is another advancing concept that involves analysis of molecular modification that alters the activity of DNA without affecting its sequence. Applications of DNA methylation include determination of tissue type of the animal DNA source, phenotypic identification, and age estimation of the animal (Eilerts 2016).

References

Arif IA, Khan HA, Bahkali AH et al (2011) DNA marker technology for wildlife conservation. Saudi J Biol Sci 18:219–225. https://doi.org/10.1016/j.sjbs.2011.03.002

Arora R, Bhatia S, Mishra BP (2011) Genetic variation and relationship of six Indian sheep breeds adapted to the northwestern arid zone of Rajasthan. Biochem Genet 49:449–457. https://doi.org/10.1007/s10528-011-9421-0

Barendse W, Armitage SM, Kossarek LM, Shalom A, Kirkpatrick BW, Ryan AM, Clayton D, Li L, Neibergs HL, Zhang N et al (1994) A genetic linkage map of the bovine genome. Nat Genet 6:227

Baylor College of Medicine Human Genome Sequencing Center (2006) Bovine Whole Genome Assembly release Btau_3.1. http://www.hgsc.bcm.tmc.edu/projects/bovine/

Berger C, Berger B, Parson W (2009) Canine DNA profiling in forensic casework: the tail wagging the dog. Forensic Sci Rev 21:1–14

Binns MM, Holmes NG, Rolliman A, Scott AM (1995) The identification of polymorphic microsatellite loci in the horse and their use in thoroughbred parentage testing. Br Vet J 151:9–15

Bishop MD, Kappes SM, Keele JW, Stone RT, Sunden SL, Kawkins GA, Toldo SS, Fries R, Grosz MD, Yoo J et al (1994) A genetic linkage map for cattle. Genetics 136:619–639

Bock JH, Norris DO (2016) Forensic plant science. Elsevier, Amsterdam

Breen M, Lindgren G, Binns MM, Norman J, Irvin Z, Bell K, Sandberg K, Ellegren H (1997) Genetical and physical assignments of equine microsatellites—first integration of anchored markers in horse genome mapping. Mamm Genome 8:267–273

Brezinsky L, Kemp SJ, Teale AJ (1993) ILSTS006: a polymorphic bovine microsatellite. Anim Genet 24:73

Buitkamp J, Antes R, Wagner V (1999) DNA profiling in veterinary medicine. In: Epplen JT, Lubjuhn T (eds) DNA profiling and DNA fingerprinting. Birkhäuser, Basel, pp 53–70. https://doi.org/10.1007/978-3-0348-7582-0_4

Bulut Z, Kurar E, Ozsensoy Y et al (2016) Genetic diversity of eight domestic goat populations raised in Turkey. Biomed Res Int 2016:1–6. https://doi.org/10.1155/2016/2830394

Butler JM, David VA, O'Brien SJ, Menotti-Raymond M (2002) The MeowPlex: a new DNA test using tetranucleotide STR markers for the domestic cat. Profiles DNA, Promega Corporation 5 (2):7–10. http://www.promega.com/profiles/502/ProfilesInDNA_502_07.pdf

Carneiro M, Ferrand N, Nachman MW (2008) Recombination and speciation: loci near centromeres are more differentiated than loci near telomeres between subspecies of the European rabbit (Oryctolagus cuniculus). Genetics 181:593–606. https://doi.org/10.1534/genetics.108.096826

Carneiro M, Afonso S, Geraldes A et al (2011) The genetic structure of domestic rabbits. Mol Biol Evol 28:1801–1816. https://doi.org/10.1093/molbev/msr003

Cassidy BG, Gonzales RA (2005) DNA testing in animal forensics. J Wildl Manag 69:1454–1462. https://doi.org/10.2193/0022-541x(2005)69[1454:dtiaf]2.0.co;2

Charoensook R, Gatphayak K, Brenig B, Knorr C (2019) Genetic diversity analysis of Thai indigenous pig population using microsatellite markers. Asian Australas J Anim Sci 32:1491–1500. https://doi.org/10.5713/ajas.18.0832

Ciampolini R, Cecchi F, Spinetti I et al (2017) The use of genetic markers to estimate relationships between dogs in the course of criminal investigations. BMC Res Notes 10:414. https://doi.org/10.1186/s13104-017-2722-6

Coogle L, Bailey E, Reid R, Russ M (1996) Equine dinucleotide repeat polymorphisms at loci LEX002, −003, −004, −005, −007, −008, −009, −010, −011, −013 and −014. Anim Genet 27(2):126–127

Duguay SA (1994) In memory of the journal pioneer.http://www.journalpioneer.com/. Accessed 31 Aug 2018

Dunisławska A, Łachmańska J, Sławińska A, Siwek M (2017) Next generation sequencing in animal science—a review. Anim Sci Paper Rep 35:205–224

Eenennaam ALV, Weaber RL, Drake DJ et al (2007) DNA-based paternity analysis and genetic evaluation in a large, commercial cattle ranch setting 1. J Anim Sci 85:3159–3169. https://doi.org/10.2527/jas.2007-0284

Eggleston-Stott ML, DelValle A, Bautista M, Dileanis S, Wictum E, Bowling AT (1997) Nine equine dinucleotide repeats at microsatellite loci UCDEQ136, UCDEQ405, UCDEQ412,

UCDEQ425, UCDEQ437, UCDEQ467, UCDEQ487, UCDEQ502 AND UCDEQ505. Anim Genet 28(5):370–371

Eilerts DF (2016) Beyond DNA: epigenetics and proteomics in forensic science. Themis Res J Just Stud Forensic Sci 4:175–187

Ellegren H, Johansson M, Sandberg K, Andersson L (1992) Cloning of highly polymorphic microsatellites in the horse. Anim Genet 23(2):133–142

Ganapathi P, Rajendran R, Kathiravan P (2012) Detection of occurrence of a recent genetic bottleneck event in Indian hill cattle breed Bargur using microsatellite markers. Trop Anim Health Prod 44:2007–2013. https://doi.org/10.1007/s11250-012-0171-8

Gaur A, Reddy A (2017) DNA techniques in wildlife forensics (animals): standard operating procedures (SOP). CSIR Centre for Cellular and Molecular Biology, Hyderabad. v + 37 p

Georges M, Massey J (1992) Polymorphic DNA markers in Bovidae. World Intellectual Property Organization, Geneva. Patent application WO PUBL NO 92/13102

Guérin G, Bertaud M, Amigues Y (1994) Characterization of seven new horse microsatellites: HMS1, HMS2, HMS3, HMS5, HMS6, HMS7 and HMS8. Anim Genet 25(1):62

Hussain T, Babar ME, Ali A et al (2017) Microsatellite based genetic variation among the buffalo breed populations in Pakistan. J Vet Res 61:535–542. https://doi.org/10.1515/jvetres-2017-0057

Irvin Z, Giffard J, Brandon RB, Breen M, Bell K (1998) Equine dinucleotide at loci ASB 21, 23, 25 and 37-43. Anim Genet 29(1):67

Kanthaswamy S (2015) Review: domestic animal forensic genetics—biological evidence, genetic markers, analytical approaches and challenges. Anim Genet 46:473–484. https://doi.org/10.1111/age.12335

Lincoln PJ, Thomson J (1998) Forensic DNA profiling protocols. Humana Press, New Jersey

Lincoln PJ, Thomson J (1998) Forensic DNA profiling protocols. Humana Press, New Jersey

Mahmoud AH, Abu-Tarbush FM, Alshaik M et al (2019) Genetic diversity and population genetic structure of six dromedary camel (Camelus dromedaries) populations in Saudi Arabia. Saudi J Biol Sci 27:1384. https://doi.org/10.1016/j.sjbs.2019.11.041

Mahmoud A, Abou-Tarboush F, Alshaikh M, Aljumaah R (2020a) Genetic characterization and bottleneck analysis of Maghateer camel population in Saudi Arabia using microsatellite markers. J King Saud Univ Sci 32:1353–1358. https://doi.org/10.1016/j.jksus.2019.11.027

Mahmoud AH, Farah MA, Rady A et al (2020b) Molecular characterization of goats from Saudi Arabia using microsatellite markers. J King Saud Univ Sci 32:1681–1686. https://doi.org/10.1016/j.jksus.2020.01.001

Marklund S, Ellegren H, Eriksson S, Sandberg K, Andersson L (1994) Parentage testing and linkage analysis in the horse using a set of highly polymorphic microsatellites. Anim Genet 25:19–23

Menotti-Raymond M, David VA, Weir BS, O'Brien SJ (2012) A population genetic database of cat breeds developed in coordination with a domestic cat STR multiplex. J Forensic Sci 57 (3):596–601. https://doi.org/10.1111/j.1556-4029.2011.02040.x

Mishra B, Kataria R, Bulandi S et al (2007) Riverine status and genetic structure of Chilika buffalo of eastern India as inferred from cytogenetic and molecular marker-based analysis. J Anim Breed Genet 126:69–79. https://doi.org/10.1111/j.1439-0388.2008.00759.x

Mukesh M, Sodhi M, Bhatia S (2006) Microsatellite-based diversity analysis and genetic relationships of three Indian sheep breeds. J Anim Breed Genet 123:258–264. https://doi.org/10.1111/j.1439-0388.2006.00599.x

Nishant K, Vrijesh KY, Ajay KR (2017) Wildlife forensic: current techniques and their limitations. J For Sci Crim 5(3):306

Oh J-D, Song K-D, Seo J-H et al (2014) Genetic traceability of black pig meats using microsatellite markers. Asian Australas J Anim Sci 27:926–931. https://doi.org/10.5713/ajas.2013.13829

Oldt R, Ng J, Kanthaswamy S (2016) The application of forensic animal DNA analysis in criminal and civil investigations. In: Amorim A, Budowle B (eds) Security science and technology.

Handbook of forensic genetics. World Scientific, New Jersey, pp 473–491. https://doi.org/10.1142/9781786340788_0019
Panday R, Jha DK, Thapa N, Pokharel BR, Aryal NK (2014) Forensic wildlife parts and their product identification and individualization using DNA barcoding. Open Forensic Sci J 7:6–13. https://doi.org/10.2174/1874402801407010006
Schury N, Schleenbecker U, Hellmann A (2014) Forensic animal DNA typing: allele nomenclature and standardization of 14 feline STR markers. Forensic Sci Int Genet 12:42–59. https://doi.org/10.1016/j.fsigen.2014.05.002
Seo J-H, Park K-D, Lee H-K, Kong H-S (2016) Genetic diversity of Halla horses using microsatellite markers. J Anim Sci Technol 58:40. https://doi.org/10.1186/s40781-016-0120-6
Sharma R, Kishore A, Mukesh M et al (2015) Genetic diversity and relationship of Indian cattle inferred from microsatellite and mitochondrial DNA markers. BMC Genet 16:73. https://doi.org/10.1186/s12863-015-0221-0
Sobrino B, Brión M, Carracedo A (2005) SNPs in forensic genetics: a review on SNP typing methodologies. Forensic Sci Int 154:181–194. https://doi.org/10.1016/j.forsciint.2004.10.020
Sodhi M, Mukesh M, Mishra BP et al (2011) Microsatellite analysis of genetic population structure of zebu cattle (Bosindicus) breeds from North-Western region of India. Anim Biotechnol 22:16–29. https://doi.org/10.1080/10495398.2011.536091
Solinas-Toldo SS, Fries R, Steffen P, Neibergs HL, Barendse W, Womack JE, Hetzel DJ, Stranzinger G (1993) Physically mapped, cosmid-derived microsatellite markers as anchor loci on bovine chromosomes. Mamm Genome 4:720–727
Somanathan A, Mathur K (2017) Application of STR markers in forensics. Int J Adv Res Sci Eng 6:61–79
Steffen P, Eggen A, Dietz AB, Womack JE, Stranzinger G, Fries R (1993) Isolation and mapping of polymorphic microsatellites in cattle. Anim Genet 24:121
Stockmarks Kit (n.d.). https://www.thermofisher.com/order/catalog/product/4307480?SID=srch-srp-4307480
Sunden SLF, Stone RT, Bishop MD, Kappes SM, Keele JW, Beattie CW (1993) A highly polymorphic bovine microsatellite locus: BM2113. Anim Genet 24:69
Thieven U, Toldo SS, Friedi R, Masabanda J, Fries R, Barendse W, Simon D, Harlizius B (1997) Polymorphic CA-microsatellites for the integration of the bovine genetic and physical map. Mamm Genome 8:52
Tom BK, Koskinen MT, Dayton M, Mattila AM, Johnston E, Fantin D, DeNise S, Spear T, Smith DG, Satkoski J, Budowle B, Kanthaswamy S (2010) Development of a nomenclature system for a canine STR multiplex reagent kit. J Forensic Sci 55:597–604
Vaiman D, Mercier D, Moazam-Goudarzi K, Eggen A, Ciampolini R, Lépingle A, Velmala R, Kaukinen J, Varvio SL, Martin P et al (1994) A set of 99 cattle microsatellites: characterization, synteny mapping, and polymorphism. Mamm Genome 5:288–297
van Asch B, Alves C, Gusmão L, Pereira V, Pereire F, António A (2009) A new autosomal STR nineplex for canine identification and parentage testing. Electrophoresis 301:417–423
Van de Goor LHP, van Haeringen WA, Lenstra JA (2011) Population studies of 17 equine STR for forensic and phylogenetic analysis. Anim Genet 42:627–633
Van Haeringen H, Bowling AT, Stott ML, Lenstra JA, Zwaagstra KA (1994) A highly polymorphic horse microsatellite locus: VHL20. Anim Genet 25(3):207
Vázquez JF, Pérez T, Ureña F et al (2004) Practical application of DNA fingerprinting to trace beef. J Food Prot 67:972–979. https://doi.org/10.4315/0362-028x-67.5.972
Vignal A, Milan D, Sancristobal M, Eggen A (2002) A review on SNP and other types of molecular markers and their use in animal genetics. Genet Sel Evol 34(3):275–305. https://doi.org/10.1186/1297-9686-34-3-275

Wictum E, Kun T, Lindquist C, Malvick J, Vankan D, Sacks B (2012) Developmental validation of DogFiler, a novel multiplex for canine DNA profiling in forensic casework. Forensic Sci Int Genet 7(1):82–91. https://doi.org/10.1016/j.fsigen.2012.07.001

Yang W, Kang X, Yang Q, Lin Y, Fang M (2013) Review on the development of genotyping methods for assessing farm animal diversity. J Anim Sci Biotechnol 4(1):2. https://doi.org/10.1186/2049-1891-4-2

Yang F, Ding F, Chen H et al (2018) DNA barcoding for the identification and authentication of animal species in traditional medicine. Evid Based Complement Alternat Med 2018:1–18. https://doi.org/10.1155/2018/5160254

20 DNA Forensics in Combating Illegal Wildlife Trade: Present, Past, and Future Perspectives

Gul Jabin, Vivek Sahajpal, Kailash Chandra, and Mukesh Thakur

Abstract

Wildlife crime has globally reached to billion-dollar industry, which has put wildlife at risk for local extirpation. It is our major duty to combat such blatant industry to maintain ecological stability. Combating wildlife crime and increasing sustainability use of wildlife not only will help in conserving animals but also the mainkind. Wildlife forensics in today's world has become an imperative procedure in monitoring illegal wildlife trade. DNA forensics is a major part in this procedure to fight against wildlife crime and providing stability species diversity. DNA forensics aid in identification of species with no morphological characters, pedigree, population, and gender of different samples seized by law enforcement agencies, thus helping in winding-up of cases efficiently.

Keywords

Wildlife · DNA · Forensics · Wildlife crime · PCR

20.1 Introduction

Biodiversity is imperative in maintaining ecological balance and survivorship of the mankind, but illegal hunting and poaching of wildlife for the commercial benefits have brought the existence of several wild species in question. India is known as one of the mega biodiversity countries in the world possessing ten biogeographic regions that enjoy varying ecologically rich landscape like the floristically rich areas of Northeast India, the northwest Himalayan region, the Western Ghats, and the

G. Jabin · K. Chandra · M. Thakur (✉)
Zoological Survey of India, New Alipore, Kolkata, West Bengal, India

V. Sahajpal
State Forensic Science Lab, Junga, Shimla, Himachal Pradesh, India

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_20

Andaman and Nicobar island with high endemism. Due to encompassing globally rich biodiversity in the South-East Asia, India has also become a hub for illegal wildlife trade providing center position for easy transport of wildlife parts and products from East to West or vice versa across international boundaries. Millions of individual belonging to different species such as pangolin, tigers, elephant, bear, tortoise, and so on are targeted for meat, ornamental body parts, and products used in traditional medicine (Graham-Rowe 2011). Identifying the altered parts and derivatives belonging to wildlife is another serious challenge among the law enforcement agencies to confiscate the materials. However, continuing advancement in the forensic tools used in wildlife identification has been assisting experts and law enforcement agencies in effective implementation of Wildlife (Protection) Act, 1972 of India. DNA forensics deals with species identification based on the investigation of the specific genes (Ogden et al. 2009). In India, species identification of the wildlife materials is authorized by the four government institutions, viz., Zoological Survey of India, Botanical Survey of India, Forest Research Institute, and Wildlife Institute of India as per the Gazette of India–July 2017.

20.2 DNA Forensics

Molecular technology is the most efficient and robust technology used for identification of samples whose morphological characteristics are lost or degraded and cannot be identified by conventional methods. Many studies have shown the applicability of the DNA marker technology in identifying species from various type of seizures such as confiscated meat samples (Verma and Singh 2003; Ghosh et al. 2019; Jabin et al. 2019), cooked and dried meats (Wong et al. 2004), dried shark fins (Chapman et al. 2003), eggshells (Moore et al. 2003), animal hairs (Branicki et al. 2003; Sahajpal and Goyal 2010), bone (Prado et al. 2002), ivory (Wasser et al. 2004, 2007), rhinoceros horns (Hsieh et al. 2003), turtle shell (Lo et al. 2006), feathers (Mukesh et al. 2013), and fish scales (Kumar et al. 2007).

DNA barcoding is a method by which species are identified by generating species-specific data and compared against the reference databases. Wildlife forensic cases like meat or products or animal derivatives are often processed or degraded but molecular techniques allow us to isolate DNA even from such samples and subsequently specific gene is amplified which is then sequenced and compared with public database of NCBI/Genbank. Percent identity gives the information about matched sequence and identity of unknown sample from present databases.

Molecular markers are used in DNA forensic analysis to identify the unknown samples, gender, individual, pedigree, and population. These can be mitochondrial or nuclear gene or repeats which have variation among different species but are generally conserved within species. Use of nuclear markers is not quite developed yet and limited data is available; however, it can be used for individual identification, population, and pedigree analysis. Mitochondrial markers such as cytochrome B, cytochrome C oxidase I, 12S rRNA, 16S rRNA, and so on (Mukesh et al. 2013; Hebert et al. 2003; Thakur et al. 2017) are dominantly used for species identification

in DNA forensics. Mitochondrial markers are inherited from mother to next generation without any recombination process and thus, can assign the unidentified specimen to particular species efficiently even from very low-quality degraded sample.

The primary method involves DNA extraction from biological samples, out of which a certain gene is amplified through PCR. The amplified product is sequenced and then generated sequence is compared with reference databases. A measure of similarity is calculated between the query sequence and the sequence available on database and most similar species is attributed to the unknown sample. Various databases available and commonly used are NCBI/EMBL/DDBJ database (www.insdc.org) and BOLD (www.barcodinglife.com).

20.2.1 Application

DNA forensics has been successfully used in cases which have been forwarded to ZSI for investigation. Nuclear markers have been used in cases where pedigree needs to be understood or to find relationship between different individuals.

For instance, in a recent case, ZSI was requested to undertake genetic analysis of three baby Chimpanzees seized and housed at New Alipore Zoo, Kolkata. Genetic study was done to understand the pedigree relationship among three Chimpanzees. Nuclear markers (STRs) were used for analysis but since, we did not have STRs specifically designed for Chimps, and we attempted heterologous markers developed for human for their possible applicability in Chimps. Analyzing genealogical relationship statistically, we attempted to establish four possible categories, that is, unrelated (U), half-siblings (HS), full siblings (FS), and parent–offspring (PO) with 19 loci, but we found that none of the analyzed individuals was genetically related. This can be interpreted that analyzed Chimps would have been brought from distinct geographical locations and therefore demonstrated different genetic make-up (Thakur et al. 2018).

In another case, we received seized objects, suspected to be originated from Elephant by DRI, Darjeeling, West Bengal, India. It was informed that an elephant was also found dead in Nepal at Budhabare area which was close to the Indo-Nepal border and the concern was raised that the ivory recovered and seized by DRI, Darjeeling might have originated from the said elephant. In this connection of above-mentioned seizure, ZSI received elephant remains (piece of ivory and flesh) from the elephant found dead at Jhapa District, Eastern Nepal. It was asked whether or not it matches with the DNA of the seized ivories by DRI, Siliguri, India. DNA extractions, PCR, genotyping based on fluorescent markers, and scoring of alleles were undertaken. PCR and genotyping were done at least in three replicates to conclude the findings. The results of genetic profiling have been conclusive to provide evidences that the two seizures were from two different individuals (Singh et al. 2019).

Mitochondrial markers were used in identification of species from degraded biological samples with no morphological characteristics such as raw meat, cooked meat, tissue samples, scales, and blood samples. In another case, ZSI received seized

raw meat sample suspected to be civet and another case of seized raw meat suspected to be of deer. Samples were analyzed for identification using mitochondrial markers Cyt b, 12S rRNA, and 16S rRNA. Upon identification, we found the former sample was originated from Asian Palm civet (Schedule II, WPA 1972) and later was of cattle. Thus, gives us enough validation against wildlife criminals as well as innocent (Ghosh et al. 2019; Jabin et al. 2019)

Another case received was roasted meat sample seized at Airport destined to the USA, upon identification, we found that the sample was identical to Oriental house rat (Basu et al. 2020). It was a case of suspense as to why roasted rat will be sent to another country?

20.3 Conclusion

DNA forensics plays a vital role in identification of highly degraded samples and has been implicated in various forensic studies (Verma and Singh 2003; Wong et al. 2004; Sahajpal and Goyal 2010; Thakur et al. 2018; Ghosh et al. 2019). Recent developments in molecular techniques have allowed us to solve many wildlife crime cases efficiently and end the cases with better results for combating wildlife crime.

However, there are some limitations in genetic analysis of some samples in case of hair; DNA can be isolated only when follicle is present at the root of hair. Similarly, some samples which are preserved in formalin cannot be processed well as amplification during PCR is inhibited by formalin. It is also quite difficult to extract DNA from old or ill-stored samples but new techniques are developing to process such samples. New and developed techniques may allow us to better understand our work and reduce limitations in way to curb the wildlife crime. It is expected that the work will help law enforcement agencies to detain those at fault in efficient way and aid in wildlife conservation. Thus, not only maintaining wildlife diversity but also reducing the risk lurking on human survivability by maintaining ecological balance.

References

Branicki W, Kupiec T, Pawlowski R (2003) Validation of cytochrome b sequence analysis as a method of species identification. J Forensic Sci 48:83–87

Basu S, Jabin G, Ghosh A et al (2020) Ascertaining suspected wildlife trade from detained parcels under international shipment. Proc Zool Soc 73:320–323. https://doi.org/10.1007/s12595-019-00312-7

Chapman DD, Abercrombie DL, Douady CJ, Pikitch EK, Stanhope MJ, Shivji MS (2003) A streamlined, biorganelle, multiplex PCR approach to species identification: application to global conservation and trade monitoring of the great white shark, Carcharodon carcharias. Conserv Genet 4:415–425

Ghosh A, Basu S, Khatri H, Chandra K, Thakur M (2019) Ascertaining species of origin from confiscated meat using DNA forensics. DNA Mit B 4(1):329–331

Graham-Rowe D (2011) Endangered and in demand. Nature 480:8–10

Hebert PDN, Ratnasingham S, DeWaard JR (2003) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc R Soc Lond Ser B Biol Sci 270(Suppl 1):96–99
Hsieh HM, Huang LH, Tsai LC, Kuo YC, Meng HH, Linacre A et al (2003) Species identification of rhinoceros horns using the cytochrome b gene. Forensic Sci Int 136:1–11
Jabin G, Ghosh A, Basu S, Khatri H, Singh SK, Chandra K, Sharma LK, Thakur M (2019) Wildlife forensics in nullifying the false accusation: a case to deal with raw meat. DNA Mit B 4 (1):736–739
Kumar R, Singh PJ, Nagpure NS, Kushwaha B, Srivastava SK, Lakra WS (2007) A non-invasive technique for rapid extraction of DNA from fish scales. IJEB 45(11):992–997
Lo CF, Lin YR, Chang HC, Lin JH (2006) Identification of turtle Shell and its preparations by PCR-DNA sequencing method. J Food Drug Anal 14(2):153–158
Moore MK, Bemiss JA, Rice SM, Quattro JM, Woodley CM (2003) Use of restriction fragment length polymorphisms to identify sea turtle eggs and cooked meats to species. Conserv Genet 4:95–103
Mukesh SSK, Shukla M, Sharma LK, Mohan N, Goyal SP, Sathyakumar S (2013) Identification of galliformes through forensically informative nucleotide sequencing (FINS) and its implication in wildlife forensics. J Forensic Res 4:195
Ogden R, Dawnay N, McEwing R (2009) Wildlife DNA forensics-bridging the gap between conservation genetics and law enforcement. Endanger Species Res 9:179–195
Prado M, Franco CM, Fente CA, Cepeda A, Vázquez BI, Barros-Velázquez J (2002) Comparison of extraction methods for the recovery, amplification and species-specific analysis of DNA from bone and bone meals. Electrophoresis 23:1005–1012
Sahajpal V, Goyal SP (2010) Identification of a forensic case using microscopy and forensically informative nucleotide sequencing (FINS): a case study of small Indian civet (Viverricula indica). Sci Justice 50(2):94–97
Singh SK, Jabin G, Basumatary T, Bhattarai GP, Chandra K, Thakur M (2019) Resolving the transboundary dispute of elephant poaching between India and Nepal. Forensic Sci Int Synergy. https://doi.org/10.1016/j.fsisyn.2019.07.007
Thakur M, Javed R, Kumar VP, Shukla M, Singh N, Maheshwari A, Dong Dong W, Zhang YP (2017) DNA forensics in combating food frauds: a case study from China in identifying canned meat labeled as deer origin. Curr Sci 112(12):2449–2442
Thakur M, Chandra K, Sahajpal V, Samanta A, Mitra A (2018) Functional validation of human-specific PowerPlex 21 system (Promega, USA) in chimpanzee (Pan troglodytes). BMC Res Notes 11:695
Verma SK, Singh L (2003) Novel universal primers establish identity of an enormous number of animal species for forensic application. Mol Ecol Notes 3:28–31
Wasser SK, Shedlock AM, Comstock K, Ostrander EA, Mutayoba B, Stephens M (2004) Assigning African elephant DNA to geographic region of origin: applications to the ivory trade. Proc Natl Acad Sci U S A 101:14847–14852
Wasser SK, Mailand C, Booth R, Mutayoba B, Kisamo E, Clark B et al (2007) Using DNA to track the origin of the largest ivory seizure since the 1989 trade ban. Proc Natl Acad Sci U S A 104:4228–4233
Wong KL, Wang J, But PPH, Shaw PC (2004) Application of cytochrome b DNA sequences for the authentication of endangered snake species. Forensic Sci Int 139:49–55

21 The Utility of DNA Barcoding Technology in the Authentication of Medicinal Plants in Illegal Trade: A Critical Review

Mukesh Kumar Thakar and Tina Sharma

Abstract

DNA barcoding technology is the utilization of short segments of DNA for the rapid and accurate identification of closely related species. The technology offers advantages over other taxonomic methods of identification, as it requires very minute tissue specimen. DNA barcoding is a project with mission to provide unique genomic database for the identification of all types of eukaryotes. The utilization of barcode has helped so far to discover new species which were once considered merged with species having similar morphology, but now with the evolution of genetic DNA barcoding sequences, organisms can be differentiated from the closely related species as well as phylogenetic relationships can be discovered. The technique holds potential to resolve fundamental ecological, evolutionary, biological, herbal authentication disputes, as well as answer to unresolved forensic questions. In this chapter, the use of DNA barcoding in wildlife forensics and its application in forensic science are discussed. Various aspects of DNA barcoding technology are highlighted specifically with respect to illegal trade of medicinal plants; identification and authentication of medicinal plants with detailed protocols for extraction, and amplification and sequencing of DNA barcoding are discussed.

Keywords

Wildlife forensics · Wildlife Protection Act · CITES · Forensic botany · DNA barcoding · matK · rbcL

M. K. Thakar (✉)
Department of Forensic Science, Punjabi University, Patiala, Punjab, India

T. Sharma
Department of Forensic Science and Toxicology, Chandigarh University, Mohali, Punjab, India

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_21

21.1 Medicinal Plants in Illegal Trade

Medicinal plants are extracted from wild resources and have been smuggled across the national and international borders to earn instantaneous profits (Douwma and Hollow 2014). The global black market of illegally extracted wild medicinal plants has been emphasized by the international organizations like CITES, UNCTAD, TRAFFIC, and WWF (Fig. 21.1) and national legislation like Wildlife Protection Act 1972 and others shown in Fig. 21.2, which also seeks help of scientific tools of wildlife forensics to implement their regulations to prohibit such trade practices (Wyatt et al. 2018; Dey n.d.; Wellsmith 2011; Haibin and Kunming 1999). The volume of trade and frequency of illegal activities reported in herbal/pharmaceutical industries increase continuously due to which certain plant species are becoming extinct or reaching close to extinction (Butchart et al. 2004).

There are around 60,000 plant species in the world (45,000 from India), which are used as traditional herbal medicine, and constituent of these phytomedicine are useful in curing the serious ailments (Sen et al. 2011; Pan et al. 2014; Rajendran and Basha 2010; Bordoloi and Dutta 2014). According to the global report of

Fig. 21.1 International conventions dedicated for the protection of plants: (**a**) Convention on International Trade in Endangered Species of Wild Fauna and Flora, also known as the Washington Convention, (**b**) Convention on Biological Diversity (CBD) Secretary-General of the United Nations, (**c**) International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA) Secretary-General of the United Nations, (**d**) Convention on the Conservation of European Wildlife and Natural Habitats, also known as the Bern Convention (or Berne Convention), Council of Europe, (**e**) The International Tropical Timber Agreement (ITTA, 1983), International Tropical Timber Agreement, 1994 (ITTA, 1994 or ITTA2)

Fig. 21.2 Official agencies of Government of India regulating wildlife trade in India

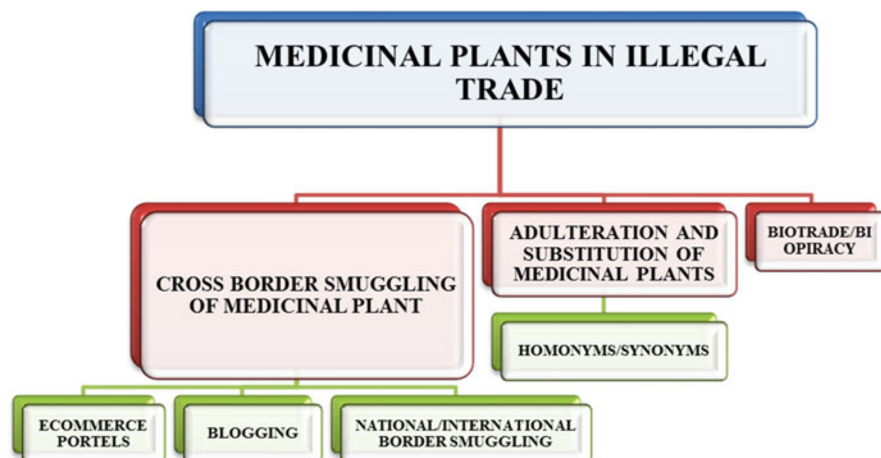


Fig. 21.3 Classification of illegal trade practices of medicinal plants along with different modes of execution

customs code HS 1211 (International Trade Center, United Nations, 2016 study), the estimated trade turnover was around 3.4 billion USD (Lange 2006). The problem becomes complicated when medicinal plants are extracted from the wild in excess amounts beyond the regulations of CITES/Wildlife Protection Act 1972. Because of this reason, many important medicinal plant species are now threatened (e.g., *Trillium govanianum*, *Sacara asoca*) (Vidyarthi et al. 2013; Dawson and Powell 1999). Because of the illegal smuggling of medicinal plants, CITES has now protected around 30,000 plant species under its regulations within three appendices (Warchol 2004). India being one of the signatory of CITES and needs to combat such crime practices and implement CITES regulations actively to conserve the national heritage of India (Reeve 2014).

The herbal sector growth is rapid, and according to data released by Confederation of Indian Industry (CII), gross market size of (end of 2018) Ayurvedic industry has grown up to $4.4 billion (Subrat et al. 2002). Annually the herbal trade industry of India is growing at the rate of 15% with a total turnover of around Rs. 2300 crore per annum and has become the second most active country to export herbal medicine across the globe after China (Wakdikar 2004). It is important to channelize such trade practices by enforcing regulations. The broad classifications of Illegal Trade of Medicinal plants is shown in Fig. 21.3.

It has become important for the enforcement agencies to channelize such trade practices within the framework of legal regulations. To ensure this, identity of herbal plants and their products is the primary requirement. The authentication of confiscated plant material to check its relevance with regulatory legislation needs to be undertaken by identifying the species of plants and their products. Taxonomist has been using morphological features for the identification of plants, yet a number of limitations exist, which need to be combated for proper enforcement of regulations (Coghlan et al. 2012). The whole plant can be easily identified by studying various morphological attributes, but plants are often traded in the form

of processed dried parts or in the form of products. In the processing, most of the morphological attributes are sacrificed and become almost impossible to identify the plants on the basis of their morphological and microscopic characteristics alone (Yao et al. 2010).

21.2 DNA Barcoding as Means of Identification

The land plants worldwide encompass a colossal diversity in terms of form and function. They consist of the seed plants (angiosperms and gymnosperms), along with the bryophytes (mosses, hornworts, and liverworts), ferns, and fern allies. Approximation of the entire species statistics varies to a great extent among authors, but according to a recent estimation, there are approximately 380,000 species of land plants, comprising 3 lakh species belonging to angiosperms, 13 hundred species belonging to gymnosperms, and approximately 12 thousand species belonging to each of bryophytes and ferns/fern allies (Fazekas et al. 2012). The identification has been carried out solely on the basis of taxonomic characteristics. Then the idea of using DNA barcodes for the identification of plants was first given by Kress et al. (2015). After 2 years, the first paper related to the identification of plants with DNA barcoding came in light.

The term “DNA barcoding” had first been coined by Paul Herbert. He mentioned the utility of short DNA segments for the purpose of identification of animal species in his publication entitled “Biological identifications through DNA barcodes.” This paper was published in the Proceedings of the Royal Society of London, Series B: Biological Sciences in 2003. This pioneer publication has become the base of the giant project “Barcode of Life” which emphasizes onto the creation of genetic database. The basic idea of the project was to accumulate genetic sequences, which should include the sequences of candidate barcoding regions. The initial work was progressing with the identification of the members of kingdom Animalia only with COI (cytochrome oxidase I) barcoding region, but now the project has expanded its branches to encompass other kingdoms (plants and others) as well. Despite the kingdom, the common goal of BOL (Barcode of Life) project was to generate a database of known species (Wong and Hanner 2008). The COI barcoding region did not work for the plant tissues; the use of chloroplast and nuclear genome was preferred in plant tissues (Hajibabaei et al. 2007). Probably the cytochrome oxidase 1 (COI) gene in plants shows very slow rate of evolution than in animals (Aubriot et al. 2013). Therefore, the use of mitochondrial genome in case of plant DNA barcoding was not encouraged by various research groups (Hebert et al. 2003; Newmaster et al. 2013). On the contrary, chloroplast genome was found to be significant in identifying plant species with high rate of discriminating ability. Four major plant DNA barcodes have been proposed, and their chronological overview for the exploration of candidate barcoding region of plants has been summarized in Fig. 21.4.

The DNA barcoding technologies have provided a new biological tool to taxonomist and forensic scientists to increase their understanding about the natural world (Fazekas et al. 2012). The species level identification with the use of short segments

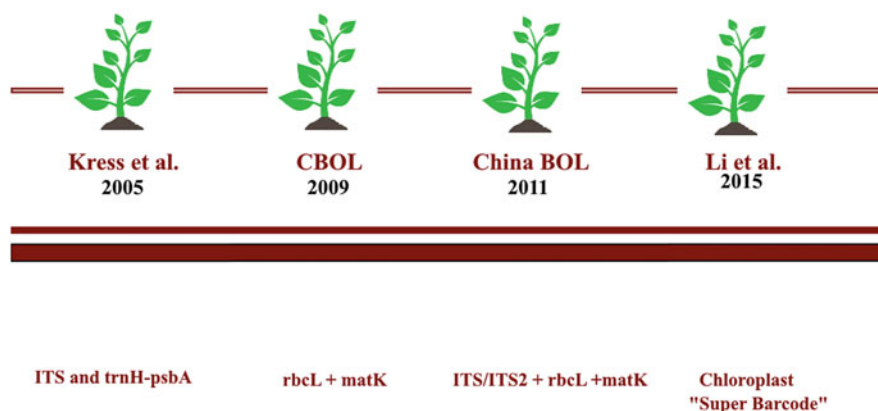


Fig. 21.4 Important milestones in the development of candidate barcoding regions for the identification of land plants

of coding and noncoding sequences of DNA has provided a practical solution to identify plants/products, whenever confiscated at custom borders and sent for forensic analysis (Wong and Hanner 2008; Hajibabaei et al. 2007; Aubriot et al. 2013; Hebert et al. 2003). The advantage of utilizing DNA barcoding technique is that only a small quantity of specimen is required irrespective of morphological and life stage of plant specimen, which is well suitable for forensic examination. With the advancement of DNA barcoding technology to identify plant or its product specimens, taxonomical studies are transitioned from manual phase of identification to the latest "Molecularization of Taxonomy" with a promise of infusing an automation and objectivity to the science of identification (Newmaster et al. 2013; Palhares et al. 2015; Carvalho et al. 2015).

The Consortium for the Barcode of Life is one such project which is dedicated to barcode organisms on the basis of standardized regions of DNA sequences. Specifically for plants, chloroplast genome is the most preferred DNA barcoding to identify plant specimens. The most common chloroplast regions utilized for authentication of plants are *rbcL*, *trnL-F*, *matK*, *ndhF*, and *atpB.* Apart from chloroplast region, *nuclear* DNA holds an important attribute to discriminate closely related species with the help of ITS/ITS2 (Hollingsworth et al. 2011; Kress et al. 2015; Ferri et al. 2015).

21.2.1 Plant DNA Barcoding and Wildlife Forensics

The correct identification and linkage of specimen with known standard form the basis of forensic identification. DNA barcoding is a new informative technology

Table 21.1 List of important forensic research articles recommending DNA barcoding as technique to identify specimen in question

Title of the Study	Recommended candidate barcoding region	Year of publication
DNA barcodes for Mexican Cactaceae, plants under pressure from wild collecting	MatK ITS	2011
An authenticity survey of herbal medicines from markets in China using DNA barcoding	ITS2 sequences	2016
Potential use of DNA barcoding for the identification of tobacco seized from water pipes	MatK + rbcL	2012
Forensic botany II, DNA barcode for land plants: Which markers after the international agreement?	rbcL + matK	2014
DNA identification of *Salvia divinorum* samples	rbcL + trnL	2012
Forensic timber identification: a case study of a CITES listed species, *Gonystylus bancanus* (Thymelaeaceae)	ITS2 + trnH-psba + trnL	2015
Forensic identification of Indian snakeroot (*Rauvolfia serpentina* Benth. Ex Kurz) using DNA barcoding	rps16 intro	2013
Forensic identification of CITES protected slimming cactus (Hoodia) using DNA barcoding	psbA-trnhL, ITS	2013
DNA barcoding for efficient identification of Ixiolirion species (Ixioliriaceae)	ITS, trn intron	2015

which can be used to resolve wildlife forensic cases. The technology can be used to authenticate the confiscated material/s and link the questioned material with species of origin. DNA barcodes in forensics can be utilized to ensure speedy court room trials in illegal wildlife cases with respect to both plants and animals (Hajibabaei et al. 2007). Many studies have been published in the international journals of repute. These studies have recommended different candidate barcoding regions for the identification of economical important medicinal plant species and are enumerated in Table 21.1.

21.2.1.1 Herbal Product Substitution

Plants are basically the source of multidimensional utilities; besides others, they are the richest resource of products that are important component of traditional systems of medicine, modern medicines, food supplements, and folk medicines. Due to lack of best practices for the authentication and identification of ingredients used in herbal products, the illegal practices are prevailing. The taxonomic identification is dependent upon the morphological and anatomical characteristics of plants, but in case of herbal formulation, these characteristics are sacrificed during the process of production. Most of the ingredients become unidentifiable, and this is the major reason that the herbal market lacks accurate quality assurance. As a result, the marketplace is lying face down to contamination and possible product substitution/adulterations. These substitutions or adulterations are not only illegal with respect to national and international legislations but are also harmful to the consumer. The herbal medicinal substitutions have been recorded for many types of medicinal plant species, teas, and

nutraceuticals. Even though high frequencies of illegal trade practices affect the consumer health, yet very limited research studies are available (Vassou et al. 2015). DNA barcodes hold the potential to authenticate the commercial available products and can validate their genuineness and substitution if present. Different studies have been conducted in this regard and are tabulated in Table 21.2.

Despite its proven potential, the technique is not fully utilized in solving wildlife forensic cases. This may be because of lack of instrumental facility and most importantly lack of DNA sequence database specifically medicinal plants in illegal trade.

21.3 Need to Develop DNA Sequence Database

The technology can be used in forensic science laboratories to solve wildlife cases related to smuggling of herbs, etc., but the only limitation it holds is the lack of standardized database of local important medicinal plants (Hajibabaei et al. 2007; Group et al. 2009). The availability of sequences of medicinal plants specifically which are in trade has almost negligible data on portals like NCBI as shown in Fig. 21.5. The lack of standard database restricts the wildlife forensic scientist to authenticate the species in question. There is an urgent need for researchers to design studies to develop DNA sequence database of the species having medicinal values and involved in wildlife trade. The richness in sequence database will not only help in the authentication of the specimen but also will be helpful in determining the geographical origin of the plant species.

CITES organization has asked the forensic science community to answer the following five questions related to wildlife crime:

(a) Identification of the species
(b) Geographical origin
(c) Wild or cultivated species
(d) Age of the trophies
(e) Individualization of the specimen (Cooper and Cooper 2013)

The resolution of this becomes possible only when giant sample size is analyzed with respect to all the above-mentioned factors. This is how the scientists can offer answers to these questions. The DNA barcoding of plants can be a significant milestone in the science of identification, as it can resolve criminal cases with respect to various illegal trade attributes.

Forensic investigators could apply plant DNA barcodes to check smuggling of endangered or restricted medicinal plants and monitor the commercial products (e.g., herbal medicines) for their claimed authenticity. The most imperative defying wildlife forensics is lack of DNA sequence database with respect to medicinal plants in trade. Around 90% of research associated with wildlife forensics is dedicated to animal species and animal produce only. The illegal trade practices with respect to medicinal plants/products or forest produce are not documented properly. As a

Table 21.2 Studies on DNA barcoding technology used to authenticate commercial herbs to evaluate their purity or substitution

Plant specimen	Part used in herbal medicine	Medicinal properties	Publication title	Reference
Sida cordifolia	Root, leaves	Antioxidant, anti-inflammatory, antidiabetic	DNA barcoding for species identification from dried and powdered plant parts: a case study with authentication of the raw drug market samples of *Sida cordifolia*	Vassou et al. (2015)
Peucedanum praeruptorum	Leaves, roots	Expectorant	Molecular authentication of the traditional medicinal plant *Peucedanum praeruptorum* and its substitutes and adulterants by DNA barcoding technique	Zhou et al. (2014)
Gingko biloba	Leaves	Parkinson's disease and Alzheimer's disease	Authentication of *Ginkgo biloba* herbal dietary supplements using DNA barcoding	Little (2014)
Phoenix dactylifera	Fruits	Fruits are antimutagenic and antioxidants	DNA barcoding based on plastid matK and RNA polymerase for assessing the genetic identity of date (*Phoenix dactylifera* L.) cultivars	Enan and Ahamed (2014)
Piper nigrum L. *fruit*		Antimicrobial, antioxidant, anti-inflammatory	DNA barcoding to detect chili adulteration in traded black pepper powder	Parvathy et al. (2014)
Indirubin	Leaves	Treatment of chronic myelocytic leukemia	Rapid identification and verification of indirubin-containing medicinal plants	Hu et al. (2015)
Lonicera	Japonica leaves, flowers	Detoxifying and anti-inflammatory effects	Stability and accuracy assessment of identification of traditional Chinese *materia medica* using DNA barcoding: a case study	Hou et al. (2013)

(continued)

Table 21.2 (continued)

Plant specimen	Part used in herbal medicine	Medicinal properties	Publication title	Reference
Gentiana scabra	Leaves, roots	Hepatoprotective and *Podophyllum hexandrum*	Evaluation of seven DNA barcodes for differentiating closely related medicinal Gentiana species and their adulterants	Wong et al. (2013)
Croton bonplandianum Baill	Seeds	Jaundice, acute constipation, abdominal dropsy	*MatK* gene-based molecular characterization of medicinal plant—*Croton bonplandianum*	Chandramohan et al. (2013)
Scutellaria baicalensis	Root	Hepatitis, jaundice, diarrhea, and inflammatory diseases	DNA barcodes for discriminating the medicinal plant *Scutellaria baicalensis* (Lamiaceae) and its adulterants	Guo et al. (2011)
Ruta graveolens L.	Leaves, stems, flowers	Fertility regulation, menstrual cramps	Authentication of *Ruta graveolens* and ITS adulterant using internal transcribed spacer (ITS) sequences of nuclear ribosomal DNA	Al-Qurainy et al. (2011)
Taxillus chinensis	Branches, leaves	Kidney reinforcement, fertility regulation	Authentication of *Taxillus chinensis* using DNA barcoding technique	Li et al. (2010)
Cinnamomum osmophloeum Kaneh.	Leaves	Antidiabetic, anti-inflammatory	DNA barcoding *Cinnamomum osmophloeum* Kaneh. Based on the partial noncoding ITS2 region of ribosomal genes	Lee et al. (2010)

result, only 10% of research studies present scientific validated tools to resolve authentication problems of plants with the help of DNA technology. Very less data with respect to Indian medicinal plants endemic to the India are available on nucleotide databases like NCBI and BOLD systems. For example, *Trillium govanianum* herb is majorly found at higher altitude regions of Himalayan region. The herb contains chemical constituents like govanoside, steroidal saponins, and compounds like borassoside and pennogenin, which have been known for

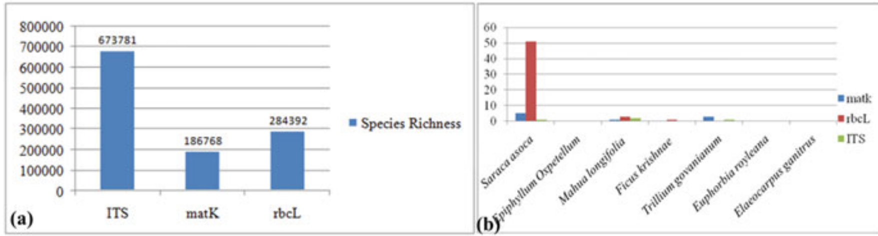


Fig. 21.5 Species sequence availability in terms of three barcoding regions, (**a**) sequences of species available in NCBI gene bank for land plants in totality, (**b**) species in illegal trade and their species richness in NCBI nucleotide bank. The data shown is representing global nucleotide sequence database. *The authors have excluded their own sequences published in NCBI

antioxidant and anticancer activities. Due to surplus illegal extraction from wild resources, the state government has presently banned the extraction of *Trillium govanianum* plant and its products without the permission of regulatory authorities. Till date almost negligible data related to the sequences of *matK* and *rbcL* and ITS barcoding regions of *Trillium govanianum species* (i.e., 12 sequences for *Trillium govanianum*, out of which only one sequence belongs to ITS gene with accession number AB018813.1 from Japan), is available on nucleotide databases like NCBI, Boldv3, DDBJ, and EMBL. The authors have submitted one sequence of ITS region to the NCBI nucleotide database in 2017.

The Consortium for the Barcode of Life suggested the development of more DNA barcode databases of medicinal plants in trade with specific candidate barcoding regions, which is urgently required and the need of the hour. This type of databases could provide a scientific tool to authenticate species in question as well as help in the regulation of CITES and other national and international legislations (Aubriot et al. 2013; Valentini et al. 2009).

21.4 Candidate Barcoding Region for Identification of Plant Specimen

The selection of candidate DNA barcode is a challenging facet in the identification of plants as each and every region exhibits variable discriminatory abilities. Unlike animal COI barcoding region, no conclusive barcoding region has given complete discretion with respect to identification of plants DNA barcode. An ideal candidate barcode must possess three main attributes:

- The first attribute is "universality," i.e., it should be universally present among all the population of species in question.
- The second attribute is "Amplification Success Rate," i.e., the segment should be amplifiable with standard universal/specific primers.

- Lastly, the third major attribute is conserved as well as variable regions which will ultimately showcase intraspecific and interspecific variations, respectively (Hebert et al. 2003).

DNA barcoding in simpler words is the utilization of short segments of DNA sequences (600–800 bps) which are universal and amplifiable and show discrimination ability. So this proposed method of identification empowers the biodiversity studies, forensic analysis, and crime scene investigation related to wild flora and fauna (Hajibabaei et al. 2007; Hollingsworth et al. 2011; Pawlowski et al. 2012; Kress and Erickson 2007). These attributes are not evenly distributed among all plant groups; therefore, it is expected that the resolution at the species level will be reasonably good in some groups and quite poor in others.

21.4.1 Single-Locus Versus Multilocus Barcode Approach

Different validation studies conducted for the authentication and selection of candidate barcoding gene have shown chloroplast genome to be best suitable region to identify land plants. Chloroplast genome further can be classified as protein coding region which includes (*matK*, *rbcL*, *rpoC1*, and *psbA-trnH*) and noncoding (ITS, *psbA-trnH*), which has been summarized by Chen et al. (2010). Although chloroplast genome exhibits higher amplification success, barcoding regions in chloroplast genome exhibits very low substitution rate, which ultimately confines its boundaries to be used as single super barcode. The species-lineage reconstructions in case of hybridization and introgression events or linear sorting events shows higher error rate. Specifically, the barcoding region of chloroplast rbcL has higher universality but lacks discriminatory power in contrast to that matK shows higher discrimination ability (Li et al. 2015; Gao et al. 2010; Newmaster et al. 2006).

21.4.1.1 "rbcL" of Chloroplast Genome

The size of coding region of chloroplast genome "rbcL" is 599 bp at 5′ of the gene specifically located at 1–599 bp with primer sites. *rbcL* is the most potential barcoding region due its smaller size and precise amplification rate with very less error rate. The only drawback of using this barcoding region is its sequence divergence with respect to interspecific and intraspecific population of plants. Although the barcode cannot be utilized as single barcode, it has its full potential to work in compliance with other barcoding region which could provide the lacking discrimination ability. The barcoding region was tested with various chloroplast and nuclear barcoding regions, so that two locus barcode approaches may give better discrimination than using the region as single barcode. The best combination agreed upon by scientific community is *rbcL* and *matK*. The barcoding region *rbcL* is also considered the backbone of plant DNA barcoding. Technically, both *rbcL* and *matK* are coding sequences of chloroplast genome and thus can be easily translated to amino acids to check any editing and assembly errors or accurate sequence orientation even the presence or absence of pseudo genes. All these sequence qualities

provide ample number of informatics opportunities to study character-based data to correlate the sequence variation with taxonomic groups and the geographical location of specimen (Newmaster et al. 2006; Pang et al. 2010; Li et al. 2011).

21.4.1.2 "matK" of Chloroplast Genome

The size of the coding region of chloroplast genome is 1500 bp in length located in between *trnK*. The matK region is also known as maturase kinase which encodes intron maturase. The major function of maturase is to splice the introns present sites overall shown in complete plastid genome sequencing of *Arabidopsis thaliana*. This barcoding region is considered equivalent to COI of animal mitochondrial genome, which is used as a single locus barcoding region for the identification of animals. Due to rapid evolution of *matK* region, it constitutes very high value of discrimination rate among the land plants. The only drawback of this region is the size of *matK* region, which makes the amplification quite cumbersome for some plant species of gymnosperms although it shows significant success with angiosperm plants. The reason behind the low amplification rate is the utilization of primers designed for angiosperms. The common primers considered for successful amplification are 3F/1R and 390/1326R. Several researchers have taken an initiative to design universal primers on the basis of clade specificity. The standardization of PCR conditions is based on species in question with respect to *matK* amplification individually. The universal primer cocktails applicable universally to all type of land plants to reduce the risk of amplification failure are major vertical researches in the arena of DNA barcoding (Li et al. 2011; Yu et al. 2011; Dunning and Savolainen 2010; Sass et al. 2007).

21.4.1.3 "ITS" of Nuclear Genome

The size of ITS region varies with respect to the genus of the plant. Overall the length of the ITS region is around 556–770 bp for Angiosperms, Conferrals, Cycadales, Ginkgoales, and Gnetales which shows the size of 975–3125 bp. In contrast, 5.8S rDNA + ITS2 show a length variation of 375–450 bp, considering the fact that the variation in gymnosperms varies with respect to the length of ITS1.

Till date, the maximum length recorded for ITS gene is with family Pinaceae, which exhibit ITS region of 1550–3125 bp long. Interspacer transcribe spacer region of nuclear genome was proposed by *China Plant DNA Barcode Meeting in 2011*, to be used as universal marker for the identification of land plants. The advantage of this barcoding region is related to the variability of sequence, which provides the opportunity of delineating closely related species. But later on this barcode was suggested to be used as a supplementary barcode along with *matK* and *rbcL*. In some species, the amplification of ITS is cumbersome, so ITS2 is used for easy amplification. Most of the plant species constitute haplotypes of plastid, thus making ITS/ITS2 excellent choice for identification. The limitation of ITS region is the divergent paralogous copies within individual resulting in incomplete coordinated evolution. These types of sites require consistent scoring method for polymorphic sites which are practically difficult to reproduce across laboratories. Moreover, the divergent copies often results in unreadable sequence. Some plant species being

fungal endophytes are prone to fungal contamination, thus making the sequencing difficult to interpret. Due to ITS regions' length, the chances of amplification success are very less. So, Gonzalez et al. in their study concluded to use ITS2 region in spite of using of ITS1–5.8S-ITS2. Due to its paralogy and enormous polymorphic sites, the sequence quality must be thoroughly quantified to gain discriminatory ability (Yao et al. 2010; Group et al. 2011; Techen et al. 2014).

21.4.1.4 Other Barcoding Region

trnH-psbA

The size of *trnH-psbA* varies from 100 to 1000 bp depending upon conifers and monocots. The bryophytes on the other hand show less than 100 bp sequence length. This region is the spacer region of plastid, which shows high level of discrimination ability along with high amplification success rate. The spacer region is often paired with matK or rbcL in terms of multilocus barcode approach to identify land plants. In case of *Ficus* and *Alnus* genus, high rate of discrimination ability as well as amplification success rate was observed. In case of genus *Quercus* and *Salix*, enhanced resolution with species delineation property has been observed. The only limitation of utilizing this region is the presence of microinversions, i.e., small-scale genetic inversion. These regions can lead to overestimation of genetic differences present between plant species, which should be detected, reoriented, or detached from the data (Group et al. 2009).

The Second International Conference of Barcode of Life conducted by CBOL Plant working group recognized other spacer genes for the purpose of identification. These include spacers like *atpF-atpH* and *psbK-psbI*. The spacer region "*psbK-psbI*" exhibits high discrimination ability, but low universality. Ki-Joong investigated floral species of South Korea and claimed positive amplification and sequencing results in case by both spacer genes (atpF-atpH and psbK-psbI) for the purpose of identification (Group et al. 2009).

The intergenic spacer between *trnL* and *trnF* is *trnL* intron which is used for species identification due to the presence of P6 loop, a small stem loop structure. The major strength of this region is the flanking variable loop of 10–143 bp also exhibiting conserved primary sites. This small-sized barcode can be useful to study the DNA, which is highly degraded and can be utilized to study complex samples like faces with the help of latest technology of next-generation sequencing (Shokralla et al. 2012).

Considering the significance of the eccentricity of the bar coding region and the performance of barcoding regions in different land plants, there will always remain species which would be better resolved by some other barcoding region (Lahaye et al. 2008). Due to this conclusive attribute, plant DNA barcoding community is still searching for single locus barcoding region which would be applicable to all land plants.

Lahaye et al. (2008) projected that *matK* individually can be the sole barcode to identify land plants, but because of low amplification success rate, the proposal was not accepted widely. Further validation of other chloroplast region combination of

seven markers showcased the potential applicability as barcoding region. These were *rpoC1*, *rpoB*, *trnH-psbA*, *rbcL*, *atpF-H*, *psbK-I*, and *matK* (Krawczyk et al. 2014). A variety of combinations of these markers were presented at the second IBOL conference held in Taipei, but no agreement was reached. No single barcoding region in plants was able to solve the ultimate goal of identification. So two barcodes, instead of one, were utilized for the identification. Later, multilocus barcodes in different combination were proposed for the accurate and precise identification without nay ambiguity (Han et al. 2013). Recently, Han et al. have proposed ITS2 in place of ITS as reliable barcoding region which can act as single locus universal plant DNA barcode (Hollingsworth 2008). The claims were rejected by various researchers because in 2008 a paper published in nature has claimed *ycf5* as potential segment of DNA in plant to work as candidate barcode (Yao et al. 2009). Several studies have observed that high discrimination rate persists for *psbA-trnH* (Yao et al. 2009; Bruyns et al. 2006; Ma et al. 2010) or for *matK* (Fazekas et al. 2008).

21.4.2 Multilocus Barcoding Approach

In multilocus barcode approach, two or more loci in combination are used and are preferred than using (single-DNA region) single-locus barcode (Fazekas et al. 2009). Some of the loci exhibit low discrimination rate, this may be because of low rate of nucleotide substitution, hybridization, polyploidy, speciation, and transitions. Furthermore, species defined on very constricted taxon concepts may be one of the explanation for showcasing such low levels of intraspecific gene flow for plastid markers (Theodoridis et al. 2012).

Theodoridis et al. (2012) also recommended the use of *matK* and *psbA-trnH* barcoding region for the identification of medicinal plants belonging to family Labiatae (Lamiaceae). In contrast, Schori and Showalter (2011) observed that identification parameters and choice of candidate barcoding region vary with each and every species of plant tested. The success rate of one barcoding region may not completely suitable for other medicinal plant species. Muellner et al. (2011) analyzed family mahogany (Meliaceae) and recommended the use of ITS/ITS2 barcode for successful identification. Nithaniyal et al. (2014) demonstrated the low discriminatory rate and limited opportunities to identify closely related species of timber specimens of Andhra Pradesh and Tamil Nadu with standard barcoding region (*matK* and *rbcL*). The observations were further recommended by Bolson et al. (2015) which analyzed tree species of high economical value.

21.4.2.1 Best Candidate Barcoding Region (*rbcL* + *matK*)

The in silico study using bioinformatics analytical tools concludes that *rbcL* + *matK* is the best multilocus barcode combination for the identification of all types of land plants. The combination was also recommended by the Consortium for the Barcode of Life 2009 to be used as universal barcode for the identification of land plants. But later on, with the advancement in research parameters and amplification success of

other barcoding region in the plant genome, more opportunities were presented than the present combination (Burgess et al. 2011; Cuénoud et al. 2002; Saarela et al. 2013).

21.5 Protocols for the Study of Plant DNA

Although specifically for wildlife forensic cases, chain of custody should be maintained, and the samples confiscated should be relocated to forensic science laboratories for the purpose of identification. In lieu of that, authors mentioned in Fig. 21.6 liked to propose a brief idea about the schematic process of specimen confiscation, collection, and preservation with the help of FTA© card. The DNA barcoding facility in forensic science laboratory can open new horizons in wildlife forensic investigations. Standard procedure for DNA barcoding has been

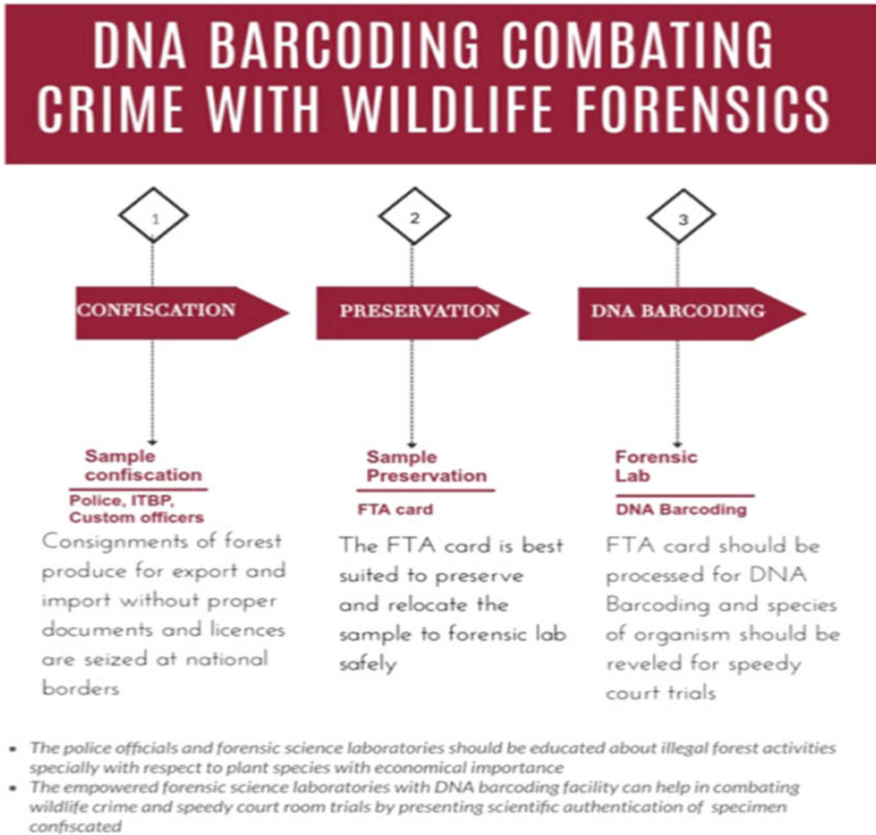


Fig. 21.6 Sample collection, preservation, and utilization of DNA barcoding for the authentication of specimen confiscated by police officials

standardized all over the globe and has also been available on the official site of CBOL/BOLD.

21.5.1 DNA Extraction

For the extraction of DNA, column-based commercial kits or standard CTAB method can be preferred. For column-based method, sample up to 100 mg is placed into micro-centrifuge of total capacity of 1.5 mL. The second step involves the addition of cell lysis buffer and incubation which will help in dissolution of cell wall and other components of cells so that the DNA can be extracted. For example, approximately 100 mg of hexaploid genome of freshly homogenized wheat would constitute 30 μg DNA, contrary to that species Arabidopsis with a smaller diploid genome will yields 3 μg DNA. The third step is usually related to the filtration of the extracted matter with the help of column. Further, after filtration, the DNA binding conditions are provided so that DNA should get attached to silica membrane. The attached DNA is then lastly eluted out from the silica membrane. The extracted DNA is preferably preserved with the help of TE buffer (Hajibabaei et al. 2005).

The extraction of plant DNA as compared to animal DNA is quite complex because of the presence of secondary metabolites. In case of thick and succulent leaves, sometimes extraction is complicated and often compromised because of the presence of polyphenols and laxatives. Apart from the above-mentioned complexities, waxy leave tissues and laxative present in cell sap present a very challenging matrix for DNA extraction (Rogers and Bendich 1989). However, the published literature related to DNA barcoding suggests the utilization of column-based kits for good quality of DNA.

21.5.2 PCR Amplification

Species-specific primer or universal DNA barcoding primers based on plant group could be utilized to run a polymerase chain reaction (PCR). The important primers that can be used for barcoding studies are described in Table 21.3. The known positive and negative controls (a mixture containing no DNA template) should be included in gel electrophoresis to ensure the quality of the reagents being used (Hajibabaei et al. 2007; Ferri et al. 2015; Vassou et al. 2015; Zhou et al. 2014; Little 2014; Group et al. 2009). The amplified DNA amplicon should be purified with the help of purification kits like Exo SAP-IT (GE Healthcare) which particularly consists of two hydrolytic enzymes, Exonuclease I and Shrimp Alkaline Phosphatase (SAP) (Werle et al. 1994). The enzyme buffer is used specifically to remove the unwanted PCR products like extraneous DNA, dNTPs, so that these extraneous matters do not contribute as inhibitor. The amplified PCR product should be incubated with enzymes at 80 °C for 15 min, so that purified PCR product could be further sequenced with the help of sequencing kit by following the manufactures protocol.

Table 21.3 Universal primers of candidate barcoding regions used for identification of species

Sl. No.	Candidate barcoding region	Direction type	Primer sequences	Reference
1.	rbcL			
	a-f	Forward	ATGTCACCACAAACAGAGACTAAAGC	Hasebe et al. (1994)
	a-r	Reverse	CTTCTGCTACAAATAAGAATCGATCTC	Hasebe et al. (1994)
	1F	Forward	ATGTCACCACAAACAGAAAC	Lledo et al. (1998)
	724R	Reverse	TCGCATGTACCTGCAGTAGC	Lledo et al. (1998)
2.	matK			
	390F	Forward	CGATCTATTCATTCAATATTTC	Yu et al. (2011)
	1326R	Reverse	TCTAGCACACGAAAGTCGAAGT	Yu et al. (2011)
	matK_1F	Forward	GAACTCGTCGGATGGAGTG	Gernandt et al. (2005)
	matK_1R	Reverse	GAGAAATCTTTTTCATTACTACAGTG	Gernandt et al. (2005)
	matK_2F	Forward	CGTACTTTTATGTTTACAGGCTAA	Gernandt et al. (2005)
	matK_2R	Reverse	TAAACGATCCTCTCATTCACGA	Gernandt et al. (2005)
3.	ITS			
	ITS2	Forward	GCTGCGTTCTTCATCGATGC	Cheng et al. (2016)
	ITS3	Reverse	GCATCGATGAAGAACGCAGC	Cheng et al. (2016)
	ITS5	Forward	GGAAGTAAAAGTCGTAACAAGG	Cheng et al. (2016)
	ITS4	Reverse	TCCTCCGCTTATTGATATGC	Cheng et al. (2016)
4.	accD 3			
	accD 1		AGTATGGGATCCGTAGTAGG	Lee et al. (2004)
	accD 3	Reverse	TTTAAAGGATTACGTGGTAC	Lee et al. (2004)

(continued)

Table 21.3 (continued)

Sl. No.	Candidate barcoding region	Direction type	Primer sequences	Reference
	accD 2	Forward	GGRGCACGTATGCAAGAAGG	Lee et al. (2004)
	accD 4	Reverse	TCTTTTACCCGCAAATGCAAT	Lee et al. (2004)
5.	ycf5 ycf5 1	Forward	GGATTATTAGTCACTCGTTGG	Vijayan and Tsou (2010)
	ycf5 3	Forward	ACTTACGTGCATCATTAACCA	Vijayan and Tsou (2010)
	ycf5 2	Forward	ACTTTAGAGCATATATTAACTC	Vijayan and Tsou (2010)
	ycf5 4	Reverse	CCCAATACCATCATACTTAC	Vijayan and Tsou (2010)
6.	trnH trnH2	Forward	CGCGCATGGTGGATTCACAATCC	Vijayan and Tsou (2010)
	psbAF	Reverse	GTTATGCATGAACGTAATGCTC	Vijayan and Tsou (2010)
	trn H (GUG)	Forward	ACTGCCTTGATCCACTTGGC	Vijayan and Tsou (2010)
	psb A	Reverse	CGAAGCTCCATCTACAAATGG	Vijayan and Tsou (2010)
7.	atpF atpF	Forward	ACTCGCACACACTCCCTTTCC	Vijayan and Tsou (2010)
	atpH	Reverse	GCTTTTATGGAAGCTTTAACAAT	Vijayan and Tsou (2010)
8.	psbK	Forward	TTAGCCTTTGTTTGGCAAG	Vijayan and Tsou (2010)
	psbI	Reverse	AGAGTTTGAGAGTAAGCAT	Vijayan and Tsou (2010)
9.	trnL–F trnL-c	Forward	CGAAATCGGTAGACGCTACG	Vijayan and Tsou (2010)
	trnL-d	Reverse	GGGGATAGAGGGACTTGAAC	Vijayan and Tsou (2010)

(continued)

Table 21.3 (continued)

Sl. No.	Candidate barcoding region	Direction type	Primer sequences	Reference
10.	rpoB rpoB 1	Forward	AAGTGCATTGTTGGAACTGG	Vijayan and Tsou (2010)
	rpoB 3	Reverse	CCGTATGTGAAAAGAAGTATA	Vijayan and Tsou (2010)
11.	rpoC1 rpoC1 1	Forward	GTGGATACACTTCTTGATAATGG	Vijayan and Tsou (2010)
	rpoC1 3	Reverse	TGAGAAAACATAAGTAAACGGGC	Vijayan and Tsou (2010)

21.5.3 DNA Sequencing

The amplified products can be sequenced by using BigDye Terminator v3.1 Cycle Sequencing Kit by following manufacturer's protocol, and the results can be matched with blast search available with NCBI or BOLD website to verify the similarity of obtained sequences with known sequences (Yao et al. 2009; Bruyns et al. 2006; Ma et al. 2010). The input query (sequences) generates a number of similarity indices. In case of certain plant species, not even a single sequence data is available on NCBI or BOLD website, which makes it quite difficult to interpret the unknown species. Research groups all around the world is emphasizing to generate the database of the economically important plant species, which can be included in such database portals for easy access and comparison by various organizations.

21.6 Implications and Challenges

DNA barcoding has started spreading its branches and making available the facilities to evaluate evidences to authenticate wildlife trophies, medicinal plants, and botanical evidences collected from scene of crimes. The DNA barcoding can also be attempted to analyze the samples of diatoms and phytoplanktons received in cases of drowning, postmortem interval analysis with entomological evidences, and even the toxicological evidences with respect to poisonous plant samples. Moreover, on a larger scale, it also helps the ecologists, conservationists, foresters, agriculturalists, and customs and quarantine officers (Fig. 21.7).

In 2016, CITES organization has asked the forensic scientific community to answer five questions related to wildlife crime, which includes

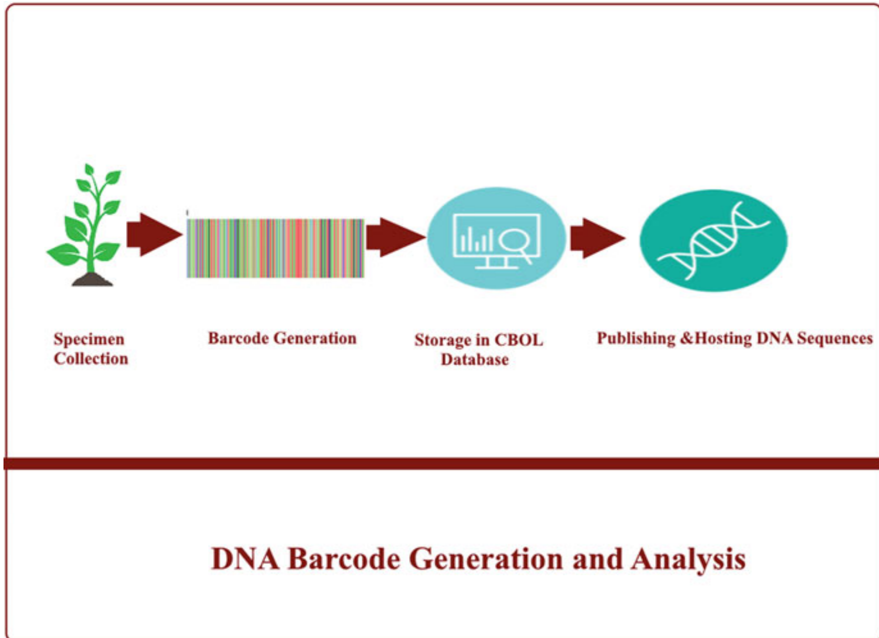


Fig. 21.7 Utilization of DNA barcoding technology for the identification of specimen as well as sequence data freely available on gene bank data portals to be used by researchers and general public

(a) Identification of the species
(b) Geographical origin
(c) Wild or cultivated or captive species
(d) Age of the trophies
(e) Individualization of the specimen (Linacre and Tobe 2011)

The major challenge related to plant DNA barcoding is still unanswered that is "agreement on candidate barcoding region" which would be universally applicable to plant species (Group et al. 2009). With the advent of time, the studies related to DNA barcoding are increasing in number, consequently the number sequences submitted to GenBank portals are also increasing. Yet due to limited availability of regional DNA sequence database, scientific community is unable to find answer related to geographical locations of plant specimens confiscated (Wasser et al. 2008).

The legal aspects are bound not only with species of origin but also with the nature of their extraction. The question of their nature, i.e., whether the specimen is being extracted from the wild resources or cultivated is still unanswered and needs attention (Halward et al. 1991). Lastly, the foremost important challenge is the extraction of DNA from dried specimens (Hajibabaei et al. 2006), mixture of specimens (Mishra et al. 2016), and specimens with contaminants which requires special emphasizes (Newmaster et al. 2013; Cowan and Fay 2012).

21.7 Conclusions

DNA barcoding is the utilization of short segment of DNA (candidate barcoding region) to identify and can even differentiate the closely related species of the plants and animals. The discriminatory ability of DNA barcoding is intensely dependent upon the selection of candidate barcoding region with regard to its interspecific and intraspecific variations. Presently, DNA barcoding technology is based on multilocus barcode approach, i.e., two locus (*rbcL* + *matK*) standard barcode for all land plants. Although, some studies have recommended single locus barcode ITS2 as universal candidate barcode region for the authentication of specimens. The study of short nucleotide polymorphism from nucleotide dataset generated from varied geographical locations can significantly give new turn to identification of geographical region with molecular approach.

The practical applicability of DNA barcoding in forensics is only possible by equipping the Forensic Science Laboratories with necessary equipment and appropriate awareness programs for the training of their officials. As it is rightly said that in near future, DNA sequence database will become available, but the next challenge will be to store and organize the data to be utilized in numerous applications.

References

Al-Qurainy F, Khan S, Ali MA, Al-Hemaid FM, Tarroum M, Ashraf M (2011) Authentication of Ruta graveolens and ITS adulterant using internal transcribed spacer (ITS) sequences of nuclear ribosomal DNA. Pak J Bot 43(3):1613–1620

Aubriot X, Lowry PP, Cruaud C, Couloux A, Haevermans T (2013) DNA barcoding in a biodiversity hot spot: potential value for the identification of Malagasy Euphorbia L. listed in CITES Appendices I and II. Mol Ecol Resour 13(1):57–65

Bolson M, de Camargo Smidt E, Brotto ML, Silva-Pereira V (2015) ITS and trnH-psbA as efficient DNA barcodes to identify threatened commercial woody angiosperms from southern Brazilian Atlantic rainforests. PLoS One 10(12):e0143049

Bordoloi R, Dutta KN (2014) A review on herbs used in the treatment of diabetes mellitus. J Pharm Chem Biol Sci 2:86–92

Bruyns PV, Mapaya RJ, Hedderson TJ (2006) A new subgeneric classification for Euphorbia (Euphorbiaceae) in southern Africa based on ITS and psbA-trnH sequence data. Taxon 55 (2):397–420

Burgess KS, Fazekas AJ, Kesanakurti PR, Graham SW, Husband BC, Newmaster SG, Percy DM, Hajibabaei M, Barrett SC (2011) Discriminating plant species in a local temperate flora using the rbcL+ matK DNA barcode. Methods Ecol Evol 2(4):333–340

Butchart SH, Stattersfield AJ, Bennun LA, Shutes SM, Akçakaya HR, Baillie JE et al (2004) Measuring global trends in the status of biodiversity: red list indices for birds. PLoS Biol 2(12): e383

Carvalho DC, Palhares RM, Drummond MG, Frigo TB (2015) DNA barcoding identification of commercialized seafood in South Brazil: a governmental regulatory forensic program. Food Control 50:784–788

Chandramohan A, Divya SR, Dhanarajan MS (2013) MatK gene based molecular characterization of medicinal plant—Croton bonplandianum Baill. Int J Biosci Res 2:1–7

Chen S, Yao H, Han J, Liu C, Song J, Shi L, Zhu Y, Ma X, Gao T, Pang X, Luo K (2010) Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS One 5 (1):e8613

Cheng T, Xu C, Lei L, Li C, Zhang Y, Zhou S (2016) Barcoding the kingdom Plantae: new PCR primers for ITS regions of plants with improved universality and specificity. Mol Ecol Resour 16(1):138–149

Coghlan ML, Haile J, Houston J, Murray DC, White NE, Moolhuijzen P et al (2012) Deep sequencing of plant and animal DNA contained within traditional Chinese medicines reveals legality issues and health safety concerns. PLoS Genet 8(4):e1002657

Cooper JE, Cooper ME (2013) Wildlife forensic investigation: principles and practice. CRC Press, Boca Raton

Cowan RS, Fay MF (2012) Challenges in the DNA barcoding of plant material. In: Plant DNA fingerprinting and barcoding. Humana Press, Totowa, NJ, pp 23–33

Cuénoud P, Savolainen V, Chatrou LW, Powell M, Grayer RJ, Chase MW (2002) Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid rbcL, atpB, and matK DNA sequences. Am J Bot 89(1):132–144

Dawson IK, Powell W (1999) Genetic variation in the Afromontane tree *Prunus africana*, an endangered medicinal species. Mol Ecol 8(1):151–156

Dey SC (n.d.) The convention on international trade in endangered species of wild fauna and flora (CITES)

Douwma G, Hollow T (2014) TRAFFIC's vision. TRAFFIC Bull 26(1)

Dunning LT, Savolainen V (2010) Broad-scale amplification of matK for DNA barcoding plants, a technical note. Bot J Linn Soc 164(1):1–9

Enan MR, Ahamed A (2014) DNA barcoding based on plastid matK and RNA polymerase for assessing the genetic identity of date (*Phoenix dactylifera* L.) cultivars. Genet Mol Res 13 (2):3527–3536

Fazekas AJ, Burgess KS, Kesanakurti PR, Graham SW, Newmaster SG, Husband BC, Percy DM, Hajibabaei M, Barrett SC (2008) Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. PLoS One 3(7):e2802

Fazekas AJ, Kesanakurti PR, Burgess KS, Percy DM, Graham SW, Barrett SC, Newmaster SG, Hajibabaei M, Husband BC (2009) Are plant species inherently harder to discriminate than animal species using DNA barcoding markers? Mol Ecol Resour 9:130–139

Fazekas AJ, Kuzmina ML, Newmaster SG, Hollingsworth PM (2012) DNA barcoding methods for land plants. In: DNA barcodes. Humana Press, Totowa, NJ, pp 223–252

Ferri G, Corradini B, Ferrari F, Santunione AL, Palazzoli F, Alu M (2015) Forensic botany II, DNA barcode for land plants: which markers after the international agreement? Forensic Sci Int Genet 15:131–136

Gao T, Yao H, Song J, Zhu Y, Liu C, Chen S (2010) Evaluating the feasibility of using candidate DNA barcodes in discriminating species of the large Asteraceae family. BMC Evol Biol 10 (1):324

Gernandt DS, López GG, García SO, Liston A (2005) Phylogeny and classification of Pinus. Taxon 54(1):29–42

Group CPW, Hollingsworth PM, Forrest LL, Spouge JL, Hajibabaei M, Ratnasingham S, van der Bank M, Chase MW, Cowan RS, Erickson DL, Fazekas AJ (2009) A DNA barcode for land plants. Proc Natl Acad Sci U S A 106(31):12794–12797

Group CPB, Li DZ, Gao LM, Li HT, Wang H, Ge XJ, Liu JQ, Chen ZD, Zhou SL, Chen SL, Yang JB (2011) Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. Proc Natl Acad Sci U S A 108 (49):19641–19646

Guo X, Wang X, Su W, Zhang G, Zhou R (2011) DNA barcodes for discriminating the medicinal plant Scutellaria baicalensis (Lamiaceae) and its adulterants. Biol Pharm Bull 34(8):1198–1203

Haibin W, Kunming G (1999) Wildlife trade in Yunnan Province, China, at the border with Vietnam. TRAFFIC Bull 18:21–30

Hajibabaei M, deWaard JR, Ivanova NV, Ratnasingham S, Dooh RT, Kirk SL, Mackie PM, Hebert PD (2005) Critical factors for assembling a high volume of DNA barcodes. Philos Trans R Soc B Biol Sci 360(1462):1959–1967

Hajibabaei M, Smith MA, Janzen DH, Rodriguez JJ, Whitfield JB, Hebert PD (2006) A minimalist barcode can identify a specimen whose DNA is degraded. Mol Ecol Notes 6(4):959–964

Hajibabaei M, Singer GA, Hebert PD, Hickey DA (2007) DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. Trends Genet 23(4):167–172

Halward TM, Stalker HT, Larue EA, Kochert G (1991) Genetic variation detectable with molecular markers among unadapted germ-plasm resources of cultivated peanut and related wild species. Genome 34(6):1013–1020

Han J, Zhu Y, Chen X, Liao B, Yao H, Song J, Chen S, Meng F (2013) The short ITS2 sequence serves as an efficient taxonomic sequence tag in comparison with the full-length ITS. BioMed Res Int 2013:741476

Hasebe M, Omori T, Nakazawa M, Sano T, Kato M, Iwatsuki K (1994) rbcL gene sequences provide evidence for the evolutionary lineages of leptosporangiate ferns. Proc Natl Acad Sci U S A 91(12):5730–5734

Hebert PD, Cywinska A, Ball SL, Dewaard JR (2003) Biological identifications through DNA barcodes. Proc R Soc Lond Ser B Biol Sci 270(1512):313–321

Hollingsworth PM (2008) DNA barcoding plants in biodiversity hot spots: progress and outstanding questions. Heredity (Edinb) 101(1):1–2

Hollingsworth PM, Graham SW, Little DP (2011) Choosing and using a plant DNA barcode. PLoS One 6(5):e19254

Hou D, Song J, Shi L, Ma X, Xin T, Han J, Xiao W, Sun Z, Cheng R, Yao H (2013) Stability and accuracy assessment of identification of traditional Chinese materia medica using DNA barcoding: a case study on Flos Lonicerae Japonicae. BioMed Res Int 2013(1):549037

Hu Z, Tu Y, Xia Y, Cheng P, Sun W, Shi Y, Guo L, He H, Xiong C, Chen S, Zhang X (2015) Rapid identification and verification of indirubin-containing medicinal plants. Evid Based Complementary Altern Med 2015:484670

Krawczyk K, Szczecińska M, Sawicki J (2014) Evaluation of 11 single-locus and seven multilocus DNA barcodes in Lamium L. (Lamiaceae). Mol Ecol Resour 14(2):272–285

Kress WJ, Erickson DL (2007) A two-locus global DNA barcode for land plants: the coding rbcL gene complements the non-coding trnH-psbA spacer region. PLoS One 2(6):e508

Kress WJ, García-Robledo C, Uriarte M, Erickson DL (2015) DNA barcodes for ecology, evolution, and conservation. Trends Ecol Evol 30(1):25–35

Lahaye R, Van der Bank M, Bogarin D, Warner J, Pupulin F, Gigot G, Maurin O, Duthoit S, Barraclough TG, Savolainen V (2008) DNA barcoding the floras of biodiversity hotspots. Proc Natl Acad Sci U S A 105(8):2923–2928

Lange DAGMAR (2006) International trade in medicinal and aromatic plants: actors, volumes and commodities. Frontis:155–170

Lee SS, Jeong WJ, Bae JM, Bang JW, Liu JR, Harn CH (2004) Characterization of the plastid-encoded carboxyltransferase subunit (accD) gene of potato. Mol Cells 17(3):422–429

Lee SC, Chiou SJ, Yen JH, Lin TY, Hsieh KT, Yang JC (2010) DNA barcoding Cinnamomum osmophloeum Kaneh. Based on the partial non-coding ITS2 region of ribosomal genes. J Food Drug Anal 18(2)

Li Y, Ruan J, Chen S, Song J, Luo K, Lu D, Yao H (2010) Authentication of Taxillus chinensis using DNA barcoding technique. J Med Plants Res 4(24):2706–2709

Li FW, Kuo LY, Rothfels CJ, Ebihara A, Chiou WL, Windham MD, Pryer KM (2011) rbcL and matK earn two thumbs up as the core DNA barcode for ferns. PLoS One 6(10):e26597

Li X, Yang Y, Henry RJ, Rossetto M, Wang Y, Chen S (2015) Plant DNA barcoding: from gene to genome. Biol Rev 90(1):157–166

Linacre A, Tobe SS (2011) An overview to the investigative approach to species testing in wildlife forensic science. Investig Genet 2(1):2

Little DP (2014) Authentication of *Ginkgo biloba* herbal dietary supplements using DNA barcoding. Genome 57(9):513–516

Lledo MD, Crespo MB, Cameron KM, Fay MF, Chase MW (1998) Systematics of Plumbaginaceae based upon cladistic analysis of rbcL sequence data. Syst Bot 23:21–29

Ma XY, Xie CX, Liu C, Song JY, Yao H, Luo K, Zhu YJ, Gao T, Pang XH, Qian J, Chen SL (2010) Species identification of medicinal pteridophytes by a DNA barcode marker, the chloroplast psbA-trnH intergenic region. Biol Pharm Bull 33(11):1919–1924

Mishra P, Kumar A, Nagireddy A, Mani DN, Shukla AK, Tiwari R, Sundaresan V (2016) DNA barcoding: an efficient tool to overcome authentication challenges in the herbal market. Plant Biotechnol J 14(1):8–21

Muellner AN, Schaefer H, Lahaye R (2011) Evaluation of candidate DNA barcoding loci for economically important timber species of the mahogany family (Meliaceae). Mol Ecol Resour 11(3):450–460

Newmaster SG, Fazekas AJ, Ragupathy S (2006) DNA barcoding in land plants: evaluation of rbcL in a multigene tiered approach. Botany 84(3):335–341

Newmaster SG, Grguric M, Shanmughanandhan D, Ramalingam S, Ragupathy S (2013) DNA barcoding detects contamination and substitution in North American herbal products. BMC Med 11(1):222

Nithaniyal S, Newmaster SG, Ragupathy S, Krishnamoorthy D, Vassou SL, Parani M (2014) DNA barcode authentication of wood samples of threatened and commercial timber trees within the tropical dry evergreen forest of India. PLoS One 9(9):e107669

Palhares RM, Drummond MG, Brasil BDSAF, Cosenza GP, Brandão MDGL, Oliveira G (2015) Medicinal plants recommended by the world health organization: DNA barcode identification associated with chemical analyses guarantees their quality. PLoS One 10(5):e0127866

Pan S-Y, Litscher G, Gao S-H, Zhou S-F, Yu Z-L, Chen H-Q, Zhang S-F, Tang M-K, Sun J-N, Ko K-M (2014) Historical perspective of traditional indigenous medical practices: the current renaissance and conservation of herbal resources. Evid Based Complement Alternat Med 2014:1

Pang X, Song J, Zhu Y, Xie C, Chen S (2010) Using DNA barcoding to identify species within Euphorbiaceae. Planta Med 76(15):1784–1786

Parvathy VA, Swetha VP, Sheeja TE, Leela NK, Chempakam B, Sasikumar B (2014) DNA barcoding to detect chilli adulteration in traded black pepper powder. Food Biotechnol 28 (1):25–40

Pawlowski J, Audic S, Adl S, Bass D, Belbahri L, Berney C, Bowser SS, Cepicka I, Decelle J, Dunthorn M, Fiore-Donno AM (2012) CBOL protist working group: barcoding eukaryotic richness beyond the animal, plant, and fungal kingdoms. PLoS Biol 10(11):e1001419

Rajendran R, Basha NS (2010) Antimicrobial activity of crude extracts and fractions of *Premna serratifolia* Linn., root. Med Plants 2(1):1–6

Reeve R (2014) Policing international trade in endangered species: the CITES treaty and compliance. Routledge, UK

Rogers SO, Bendich AJ (1989) Extraction of DNA from plant tissues. In: Plant molecular biology manual. Springer, Dordrecht, pp 73–83

Saarela JM, Sokoloff PC, Gillespie LJ, Consaul LL, Bull RD (2013) DNA barcoding the Canadian Arctic flora: core plastid barcodes (rbcL+ matK) for 490 vascular plant species. PLoS One *8* (10):e77982

Sass C, Little DP, Stevenson DW, Specht CD (2007) DNA barcoding in the cycadales: testing the potential of proposed barcoding markers for species identification of cycads. PLoS One 2(11): e1154

Schori M, Showalter AM (2011) DNA barcoding as a means for identifying medicinal plants of Pakistan. Pak J Bot 43:1–4

Sen S, Chakraborty R, De B (2011) Challenges and opportunities in the advancement of herbal medicine: India's position and role in a global context. J Herb Med 1(3–4):67–75

Shokralla S, Spall JL, Gibson JF, Hajibabaei M (2012) Next-generation sequencing technologies for environmental DNA research. Mol Ecol 21(8):1794–1805

Subrat N, Iyer M, Prasad R (2002) The ayurvedic medicine industry: current status and sustainability. Ecotech Services, London, and International Institute for Environment and Development, New Delhi

Techen N, Parveen I, Pan Z, Khan IA (2014) DNA barcoding of medicinal plant material for identification. Curr Opin Biotechnol 25:103–110

Theodoridis S, Stefanaki A, Tezcan M, Aki C, Kokkini S, Vlachonasios KE (2012) DNA barcoding in native plants of the Labiatae (Lamiaceae) family from Chios Island (Greece) and the adjacent Çeşme-Karaburun Peninsula (Turkey). Mol Ecol Resour 12(4):620–633

Valentini A, Pompanon F, Taberlet P (2009) DNA barcoding for ecologists. Trends Ecol Evol 24 (2):110–117

Vassou SL, Kusuma G, Parani M (2015) DNA barcoding for species identification from dried and powdered plant parts: a case study with authentication of the raw drug market samples of *Sida cordifolia*. Gene 559(1):86–93

Vidyarthi S, Samant SS, Sharma P (2013) Dwindling status of *Trillium govanianum* Wall. ex D. Don-A case study from Kullu district of Himachal Pradesh, India. J Med Plants Res 7 (8):392–397

Vijayan K, Tsou CH (2010) DNA barcoding in plants: taxonomy in a new perspective. Curr Sci 99:1530–1541

Wakdikar S (2004) Global health care challenge: Indian experiences and new prescriptions. Electron J Biotechnol 7(3):02–03

Warchol GL (2004) The transnational illegal wildlife trade. Crim Justice Stud 17(1):57–73

Wasser SK, Joseph Clark WILLIAM, Drori O, Stephen Kisamo EMILY, Mailand C, Mutayoba B, Stephens M (2008) Combating the illegal trade in African elephant ivory with DNA forensics. Conserv Biol 22(4):1065–1071

Wellsmith M (2011) Wildlife crime: the problems of enforcement. Eur J Crim Policy Res 17 (2):125–148

Werle E, Schneider C, Renner M, Völker M, Fiehn W (1994) Convenient single-step, one tube purification of PCR products for direct sequencing. Nucleic Acids Res 22(20):4354

Wong EHK, Hanner RH (2008) DNA barcoding detects market substitution in North American seafood. Food Res Int 41(8):828–837

Wong KL, But PH, Shaw PC (2013) Evaluation of seven DNA barcodes for differentiating closely related medicinal Gentiana species and their adulterants. Chin Med 8(1):16

Wyatt T, Johnson K, Hunter L, George R, Gunter R (2018) Corruption and wildlife trafficking: three case studies involving Asia. Asian J Criminol 13(1):35–55

Yao H, Song JY, Ma XY, Liu C, Li Y, Xu HX, Han JP, Duan LS, Chen SL (2009) Identification of Dendrobium species by a candidate DNA barcode sequence: the chloroplast psbA-trnH intergenic region. Planta Med 75(06):667–669

Yao H, Song J, Liu C, Luo K, Han J, Li Y, Pang X, Xu H, Zhu Y, Xiao P, Chen S (2010) Use of ITS2 region as the universal DNA barcode for plants and animals. PLoS One 5(10):e13102

Yu J, Xue JH, Zhou SL (2011) New universal matK primers for DNA barcoding angiosperms. J Syst Evol 49(3):176–181

Zhou J, Wang W, Liu M, Liu Z (2014) Molecular authentication of the traditional medicinal plant *Peucedanum praeruptorum* and its substitutes and adulterants by DNA-barcoding technique. Pharmacogn Mag 10(40):385

DNA Barcoding in Forensic Mycology: Concepts, Limitations, and Future Prospects

22

Spriha Sharma, Jaskirandeep Kaur Jossan, and Rajinder Singh

Abstract

In forensic investigations, mycological evidence can be an important marker for ascertaining the cause of death, time since death, and location of the corpse. Moreover, many hallucinogenic mushrooms are also encountered in illegal black markets. In both the scenarios, it is of utmost significance to identify the fungal species in question. Species identification can be performed using conventional (taxonomic) methods as well as by using DNA barcoding. Conventional methods such as morphological identification suffer from serious limitations in both poisoning and illegal trade cases. In mushroom poisoning cases, degradation and contamination occurs due to cooking, gastric juices, vomit, and fecal matter, whereas in cases of illegal trade, confiscated samples are found in powdered form. This leads to the absence of identification features which are required for the taxonomic identification. One prevalent alternative to counter this problem is "DNA barcoding," which unlike taxonomic identification does not depend on sample morphology. In this chapter, an attempt has been made to discuss various facets of DNA barcoding of fungal evidence along with the recent advancements made in this field in the past few decades. In spite of wide applications of DNA barcoding in other fields, this technique has largely remained underexplored in forensic mycology due to the lack of skilled mycologists and limited awareness of current methods among the criminal investigators. Recent advancements in fungal identification on the cadavers have preferred DNA-based methods over conventional techniques to determine time since death by analyzing the fungal growth rate. High-throughput sequencing technology (HTS) has also improved the identification capabilities of DNA barcoding in fungal species identification.

S. Sharma · J. K. Jossan · R. Singh (✉)
Department of Forensic Science, Punjabi University, Patiala, Punjab, India

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_22

Keywords

DNA barcoding · Forensic mycology · Fungal evidence · Forensic analysis

22.1 Introduction

Forensic mycology can be defined as the application of mycological evidence (including blights, molds, mildews, mushrooms, plant and human pathogens, rust, and smuts) in the criminal investigation. Mycological evidence has been documented to determine the cause of death in mushroom poisoning cases, to counter illegal trade of hallucinogenic preparations and to determine time since death and location of corpse using traces of fungi. Throughout the world, a number of studies have been reported in which fungal evidence has proved useful in solving criminal cases and its testing in the courts (Hawksworth and Wiltshire 2011; Piazza et al. 2018). Despite its many forensic significance, forensic mycology has been a lesser-explored area and is rarely employed as evidence in matters pertaining to criminal cases. The reason for the limited use of forensic mycology in a criminal investigation is the lack of awareness among crime investigators and unavailability of skilled experts.

22.1.1 Brief History and Epidemiology

Mushrooms are ubiquitous and, therefore, are consumed worldwide for their nutritional value. But, sometimes, they can be the cause of accidental poisoning deaths due to misidentification of edible and inedible species while foraging wild mushrooms. The incidences of mushroom poisonings have been reported in different parts of the world since ancient times. About 5000 mushroom species have been identified and recorded, and about 50–100 species of mushroom are known to cause poisoning effects and deaths when consumed. The first report of mushroom poisoning dates back to 430 BC by Euripides, where the author mentioned about the fatal poisoning case of his family (Berger and Guss 2005). Another case of *Amanita* mushroom poisoning was reported in which a Gulf Coast family died of fulminant hepatic failure after consumption of wild mushrooms (O'Brien and Khuu 1996). Recently, a retrospective study on the meta-analysis of several mushroom poisoning cases has been conducted worldwide, and a significant increase in the frequency of mushroom poisoning cases was observed over the years (Diaz 2017). Erguven et al. (2007) analyzed 39 pediatric patients from 1994 to 2004 of mushroom poisoning and reported that 59% of patients were females and 41% were males. In Italy from 1996 to 2000, about 10,000 cases of mushroom poisoning were reported and about 2400 of these cases showed a long incubation, 22 cases resulted in the death of patients, and nine required liver transplants due to severe hepatic insufficiency. Also, in northern Italy from 2005 to 2008, 15% of poisoning cases have been caused by *Lepiota*, *Amanita*, and *Inocybe* (Epis et al. 2010). Due to the lack of global data, it is

difficult to pinpoint the actual number of deaths due to mushroom poisoning; however, it is safe to assume that the number is substantially higher than 100 deaths/year due to the high number of underreported cases (White et al. 1993; Santi et al. 2012). Gurbuz et al. (2015) reported poisoning cases in Turkey during spring and autumn seasons from the year 2011 to 2014 in which 78.4% poisoning was due to ingestion of misidentified mushrooms. In southern China, 102 mushroom poisoning cases were investigated from 1994 to 2012 with an overall mortality of 21.48% (Chen et al. 2012). In case of hallucinogenic mushrooms, also known as "magic mushrooms," the oldest record has been found on the wall paintings in the southern part of Algeria. In 3500 BC, the symbolic paintings depict a Shaman dancing with a mushroom in his hands, suggesting a mental linkage. During the bronze age (3000 BC to 500 AD), many pieces of mushroom-shaped bronzeware and wall paintings were recovered (Matsushima et al. 2009). The magic mushrooms have later been reported to be used by Indo-Mexicans for religious, medical (physical and mental), and recreational purposes since prehistoric times (Wasson 1959). With time the recreational use of hallucinogenic mushrooms has become a serious problem all over the world (Keller et al. 1999; Tranchida et al. 2018). And in case of fungi as trace evidence, in 1955, it was proposed for the first time that it can be used as a forensic tool for the determination of time since death and for the location of the corpse. Hawksworth and Wiltshire (2011) have compiled various case studies reported by a number of authors in which fungal growth had been utilized for the identification purposes. Recent studies by Tranchida et al. (2016, 2018) have examined and identified the fungal biota grown on and under the surface of human cadaver.

22.1.2 Forensic Significance of Fungal Evidence

Generally, forensic mycology has four facets comprising (1) cause and manner of death, (2) determination of postmortem interval, (3) location of the corpse, and (4) biological warfare. However, to the best of author's knowledge, no case of biological warfare using fungi have been reported; therefore, this aspect will not be discussed in detail in the present review. For more insights, readers are advised to read an excellent review by Hawksworth and Wiltshire (2011). In this chapter, we will be focusing on first three facets of forensic mycology.

Mushrooms are fungi with large fruiting bodies which are often consumed as delicacies all over the world. However, many mushrooms can either prove to be toxic or hallucinogenic if consumed mistakenly. In such cases, their identification can help in ascertaining the cause and manner of death. Genus *Amanita*, *Lepiota*, and *Galerina* are familiar to humans as certain species are well-known for causing poisoning that often results in death. Among these members, genus *Amanita* is responsible for around 90% of global deaths due to mushroom poisoning which is primarily due to accidental consumption of misidentified wild mushrooms (Lima et al. 2012; Garcia et al. 2015; Zhou et al. 2017). On the other hand, hallucinogenic mushrooms are also confiscated in cases of illegal trade, mostly in the form of

powders or extracts. Therefore, the possession, sale, and purchase of psychedelic species of *Psilocybe* and *Panaeolus* have been regulated in many countries (Gausterer et al. 2014). Many genera containing hallucinogenic compounds include the *Amanita, Conocybe*, *Galerina*, *Gymnopilus*, *Inocybe*, *Panaeolus*, *Pluteus*, and among all these, *Psilocybe* and *Panaeolus* are the most commonly used hallucinogenic mushrooms. All these members are known to synthesize the controlled substances such as psilocin and psilocybin, and in the illegal market, magic mushrooms are often seized in intact or fresh state (Linacre et al. 2002; Stra et al. 2014; Tsujikawa et al. 2003, 2007; Nagy and Veress 2016). Cultivation and possession of mushrooms that contain psilocin and/or psilocybin is illegal in many countries (Gonmori et al. 2011; Margarida et al. 2015; Zuber et al. 2011).

One can appreciate that fungal colony on, or associated with, human cadavers can give indications of time since death by studying the growth pattern. But the reliability of any estimates will depend on the accuracy of the identification of the fungus, the storage methods for the body, and the availability of data on the temperature and humidity at the site. The fungi that have proved of most interest in estimating postmortem interval are neither specialized medically important fungi nor ones restricted to dead human tissues, but rather are decomposer or spoilage fungi that can directly colonize the surface of corpses and body parts after death. In this sense, forensic mycology is also considered as an auxiliary method in the determination of time since death just like forensic entomology. For PMI determination, fungal evidence can be used as an alternative to forensic entomology when flies are absent and fungal growth is prominent (Zuber et al. 2011; Hosukler et al. 2018). In literature, there is no criminal case reported on the role of fungi in determining the location of corpse; however, the presence of certain groups of fungi like ammonia fungi and post-putrefaction fungi can act as “grave markers” that can serve as a tool for the estimation of the postmortem interval and in location of corpse (Hosukler et al. 2018; Cecilia and Noemi 2017; Tibbett and Carter 2003; Ishii et al. 2006; Bellini et al. 2015). Various fungal species reported for ascertaining time since death include *Cladosporium* sp., *Fusarium* sp., *Geotrichum candidum*, *Hormodendron* sp., *Mortierella* sp., and *Penicillium chrysogenum*, *Aspergillus chevalieri*, *Gliocladium* sp., *Aspergillus terreus*, *Mucor*, and *Pythinium*, and only *Hebeloma radicosum* and *Hebeloma syriense* have been mentioned in literature to be associated with corpse (Hawksworth and Wiltshire 2011).

In all the above-mentioned scenarios, it is imperative to identify the fungal species associated with various crime cases. The species identification is usually based upon the observation of the macroscopic and microscopic characteristics which include unique structural features of gills and spores. However, these characteristics could be misleading or unidentifiable, especially when the fungal evidence is found in a morphologically compromised form such as in raw, cooked, gastric aspirates, and stool samples. Thus, the lack of morphological features becomes a hindrance for species-level identification. As a substitute to morphological and physicochemical studies, the potential of molecular methods is well recognized. DNA-based methods, which are standardized, rapid, and offer reliable species identification brought a revolution in the 1990s, which can be used by

non-taxonomists as well to solve cases. Recent advances have demonstrated the value of DNA-based methods in the identification of toxic mushrooms (Karlson-Stiber and Persson 2003; Bresinsky and Besl 1990; Enjalbert et al. 2002; Gicquel et al. 2014; Kang et al. 2015; Parnmen et al. 2016; Li et al. 2014; Himmelmann et al. 2001), hallucinogenic mushrooms (Linacre et al. 2002; Sarwar and Mcdonald 2003; Nugent and Saville 2004), and fungal traces (Tranchida et al. 2016, 2018; Cecilia and Noemi 2017; Schwarz et al. 2014; Wiltshire et al. 2014; Tranchida et al. 2014).

22.2 DNA-Based Identification in Forensic Mycology

In forensic cases, rapid identification of poisonous and hallucinogenic mushrooms consumed by individuals is required for three reasons: to provide proper medical treatment, to know the cause of death, and to confiscate the hallucinogenic mushrooms in cases of illegal trade. As processed, cooked, and eaten mushrooms do not retain their original shape, genetic strategies for species identification provide a promising alternative to morphological and taxonomical identification (Kowalczyk et al. 2015; Harper et al. 2011).

A number of papers published from 2000 to 2018 on DNA-based identification have been summarized in Fig. 22.1. Certain fungal genera like psychoactive fungi contain most common hallucinogens known as “psilocin” and “psilocybin” which are known to be 50 times more potent than mescaline. The phylogenetic relationship of “magic mushrooms” collected in the form of dry powders from the local market of Japan has been studied to establish an easy identification system by TaqMan polymerase chain reaction (PCR) method (Maruyama et al. 2006). In Italy, police seized hallucinogenic mushrooms cultivated in “grow-kits,” and DNA-based taxonomic identification of basidiospores was performed for the species identification (Gambaro et al. 2016). The analysis of hallucinogenic fungal genera *Psilocybe* and

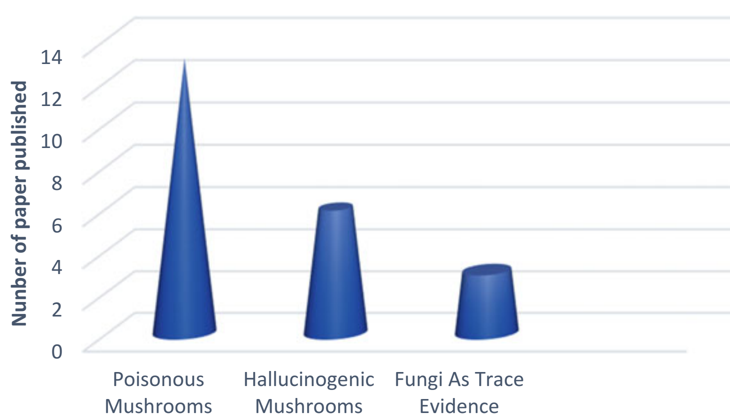


Fig. 22.1 Distribution of the number of papers published on DNA-based identification of forensically important fungal species

Panaeolus have been done through random amplification of polymorphic DNA (RAPD) and amplified fragment length polymorphism (AFLP) techniques which were the first DNA-based taxonomic tools to examine the entire genome of an organism in one reaction (Linacre et al. 2002; Lee et al. 2000). Also, molecular identification of 221 fungal evidentiary materials (dried mushrooms, stomach contents, feces) of hallucinogenic and other poisonous mushrooms has proved to be a better identification method than the sporological examination for the analysis in clinical and forensic purposes (Zuber et al. 2011). So, there is a need to ascertain the fungal species of these substances which are generally seized in different forms by adopting DNA-based techniques.

The viability of a direct PCR approach on species identification from diluted gastric contents from a forensic case has been checked for the detection of deadly poisonous European species of the genus *Amanita*, namely *A. phalloides*, *A. virosa*, and *A. verna*. PCR primers were designed to target discriminatory polymorphic sequences located on the nuclear ribosomal DNA. A series of experiments were conducted that followed the course of mushroom food processing and consumption. Test samples including homogenized mixed mushrooms (raw, fried, digested) and fecal preparations were subjected to analysis by direct PCR. Target amplification by direct PCR was successful with raw, fried, and digested mushrooms. However, fecal samples were pre-processed by using short protocol along with modified PCR mixture. With this approach, highly diluted *A. phalloides* traces in spiked stool specimens provided confirming molecular evidence in clinical cases of suspected mushroom poisoning (Gausterer et al. 2014). For the quick and accurate identification of poisonous mushrooms of genus *Amanita,* the specific multiple oligonucleotide macroarray probes were used which are based on the sequences of highly variable ITS region of rDNA (Harper et al. 2011). Maeta et al. (2008) reported a rapid system of poisonous mushroom identification using real-time PCR system from cooked mushrooms. A conventional PCR-based approach was developed for the specific detection of the glyceraldehyde-3-phosphate dehydrogenase (*gapdh*) gene fragment from *A. phalloides*. However, this technique is relatively time-consuming, *gapdh* gene is a low-copy target which produces low assay sensitivity, and the authors did not present an application to clinical samples. Single-plex real-time PCR method was developed for the specific detection of four poisonous mushrooms (including *A. phalloides*) and tested these assays with DNA extracts from a variety of different sources (dried mushrooms, pasta with mushrooms, cooked mushrooms, and gastric aspirates mixed with dried mushrooms) (Epis et al. 2010).

22.3 Basic Components of DNA Barcoding in Forensic Mycology

The identification of fungal genera using taxonomic keys is often impossible in cases of degraded samples and has traditionally been the specialized area of taxonomists, providing a nomenclatural backbone and a key prerequisite for numerous biological studies but has a limitation in analyzing degraded samples. The idea of a

standardized molecular identification system without such limitations emerged progressively after 2003 with the development of "DNA barcoding" method for species identification in animals. To create a universal system based on standardized molecular approach for species identification, The Barcode of Life Data (BOLD) system was initiated. This project enables the acquisition, storage, analysis, and publication of DNA barcode records (Frézal and Leblois 2008). DNA barcoding provided facilities and platform needed to establish a database of all organisms. Cytochrome c oxidase 1 (CO1 or *COX1*) gene sequence has emerged as an appropriate DNA region for providing high resolution within insects on various taxonomic levels (Hebert et al. 2003; Hebert and Gregory 2005; Dasmahapatra and Mallet 2006; Savolainen et al. 2005; Ratnasingham and Hebert 2007; Fišer Pečnikar and Buzan 2014). DNA barcoding utilizes a short-standardized region between 400 and 800 base pairs known as "barcodes" to identify not only interspecific but also intraspecific variations. However, it has been established that *cox1* is unsuitable for fungal identifications.

There are very few studies on DNA barcoding of fungi in forensic samples, and other related studies have included commercially available mushroom products with medicinal benefits (Frézal and Leblois 2008; Bridge 2002). To successfully perform DNA barcoding of fungi found as evidence in poisoning cases, three prerequisites must be fulfilled. First, a set of genetic markers that can be easily amplified from most if not all fungal species using universal primers along with low intraspecific and high interspecific divergence (Dentinger et al. 2010). Fungal species identification through DNA barcoding technique is a promising alternative to morphological and physicochemical studies. In fungi, the most important target region is nuclear DNA encoding ribosomal (nrDNA), arranged in the form of tandem repeats. The number of copies of nrDNA genes in mushrooms varies from 60 to 220, depending on the species. Fragment encoding ribosomal RNA are separated by noncoding fragments, which are removed after the transcription phase. The internal transcribed spacers (ITS1 and ITS2) and non-transcribed intergenic spacers (IGS) are characterized by the presence of polymorphism of length and sequence, which make them highly useful for species identification (Innis et al. 1990). Specific primers serve in the amplification of both ITS regions, including the 5.8S coding region, as well as species-specific primer targeting ITS. Khaund and Joshi (Khaund and Joshi 2014) studied wild edible mushroom species using multiloci molecular characterization using Small Sub-Unit (SSU), Internal Transcribed Spacer (ITS), RNA Polymerase II Largest Subunit (RPB1) and RPB2 markers. The authors reported that species identity generated by the ITS marker matched accurately with morphological characteristics/appearance of the specimens, indicating the ITS region as a reliable barcode for identifying wild edible mushrooms. Mouhamadou et al. (2008) investigated the molecular evolution of the V6 and V9 domain of the mitochondrial SSU-rDNA to evaluate the use of these sequences for DNA barcodes in the Basidiomycota division. The sequenced species possessed different V9 sequences due to point mutations and insertion/deletion events. They reported V9 domain as an alternative molecular marker because it has an adequate divergence level and is easily amplifiable for the identification of the fungal kingdom. However,

Table 22.1 Summary of common markers used in DNA barcoding methods for forensically important fungal species

Locus	Primer sequence (5′–3′)	Orientation	Reference
ITS 1	TCCGTAGGTGAACCTGCGG	Forward	White et al. (1990), Oda et al. (1999, 2004), Gonzalez et al. (2002), Zhang et al. (2004, 2010), Cai et al. (2014), Jansson et al. (2018), Nilsson et al. (2012), Sugawara et al. (2016) Fu et al. (2017), Weiß et al. (1998)
ITS 4	TCCTCCGCTTATTGATATGC	Reverse	
ITS 1	TCCGTAGGTGAACCTGCGG	Forward	
ITS 2	GCTGCGTTCTTCATCGATGC	Reverse	
ITS 3	GCATCGATGAAGAACGCAGC	Forward	
ITS 4	TCCTCCGCTTATTGATATGC	Reverse	
ITS 5	GGAAGTAAAAGTCGTAACAAGG	Forward	
ITS 4	TCCTCCGCTTATTGATATGC	Reverse	
nrLSU (nuclear large subunit)	LROR—ACCCGCTGAACTTAAGC	Forward	Cai et al. (2014) Vilgalys and Hester (1990)
	LR5—ATCCTGAGGGAAACTTC	Reverse	
rpb2 (RNA polymerase II)	Am-6—TGGGGSSTGGTRTGYCCTGC	Forward	Zhang et al. (2004, 2010), Cai et al. (2014)
	Am-7—*CCCATKGCTTGTTTRCCCATGGC*	Reverse	
ef 1-α (translation elongation factor 1-alpha)	983—GCYCCYGGHCAYCGTGAYTTYAT	Forward	Zhang et al. (2004, 2010), Cai et al. (2014).
	1567—ACHGTRCCRATACCACCRAT	Reverse	
β-Tubulin (Btub)	Am—AAGCGGAGCRGGTAACAAYTGG	Forward	Zhang et al. (2004, 2010), Cai et al. (2014)
	Am—ACRAGYTGGTGRACRGAGAGYG	Reverse	

many researchers have reported ITS as “universal primer pair” for the amplification of fungal sequence. The details of the common markers used as DNA barcode for the species identification of fungal evidence are given in Table 22.1.

The second basic requirement is DNA isolation and high-throughput sequencing analysis to address the problem of misidentified sequences. The first database called the Ribosomal Database Project (RPD) initially included nuclear small subunit (SSU) rRNA gene sequence. Recent additions to RPD include a database for ITS and large subunit (LSU) rRNA genes. To further enhance the sequence accuracy, a number of other databases focused on specific marker sequences have been built. The most common is UNITE (https://unite.ut.ee) database which is a web-based database and sequence management environment developed for ITS sequences of fungi. The UNITE database now acts as GenBank for all fungal ITS sequences from environmental samples (Hibbett et al. 2016; Schoch et al. 2014). Various other standard protocols have been developed that are also available in the form of commercial kits such as QIAquick PCR purification kit (Qiagen) and sequencing by BigDye Terminator sequencing reagents (Kowalczyk et al. 2015; Raja et al.

2017a). ITSx, a Perl-based software tool to extract ITS1, 5.8S gene, and ITS2, and full-length ITS sequences have been introduced from both Sanger and high-throughput sequencing (HTS) data sets. A number of ITSx pave the way for more sensitive Basic Local Alignment Search Tool (BLAST) searches, sequence clustering operations, and are also useful for amplicon-based next-generation sequencing.

The third component is a database containing DNA sequences of the genetic markers for the majority of species to be identified. The desired properties of DNA barcodes are clearly defined: (1) the DNA fragment must be nearly identical in specimens of the same species but different between individuals of different species, (2) the section must be standardized (the same section should be used in different taxonomic groups), and (3) the marker must be robust, with conservative primer building sites that allow it to be readily amplified and sequenced.

For fungi, ITS regions work well, but sequence should be confirmed with GenBank identifications. DNA taxonomy employing phylogenetic analysis is useful for placing an unknown sequence into an existing classification in an evolutionary framework. Researchers are still trying to find a single segment of DNA suitable for the identification of all taxa. Despite several years of work in this area, such a region has not been identified and a single universal DNA barcoding marker is unlikely to exist. Figure 22.2 demonstrates the systematic approach to be followed for the fungal species identification. If samples are encountered in a fresh and intact state, morphological features can be used; however, DNA extraction methods, PCR amplification, and sequencing must be used for the absolute identification. When the samples are highly degraded, direct PCR and real-time PCR analysis can be helpful in species identification. In both the cases, ITS which is marked as universal primer pair in case of fungal genera is to be employed.

22.4 Technical Issues Related to the Forensic Analysis of Fungal Evidence

22.4.1 Issues Specific to Fungi

22.4.1.1 Use of Morphological Characteristics for the Identification of Fungi

A survey of different studies performed on the identification of fungal species for the years 2000–2015 revealed that about 31% studies reported the identification of fungal species is based either on morphology only or no method for identification was used (Raja et al. 2017a). Often, morphological identification can be challenging, confusing, time-consuming, misleading, and even inaccurate even for trained mycologists.

22.4.1.2 Lack of Available Reference Sequences

It is a consensus that to perform species identification through DNA profiling, good, authentic, and reliable reference sequences are very important. Moreover, these sequences must come from well-identified specimens including full voucher

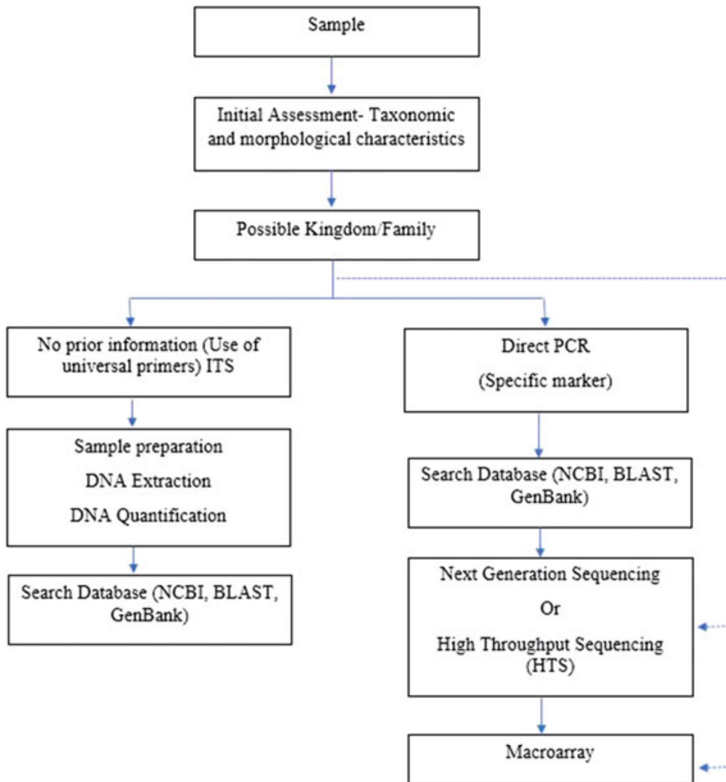


Fig. 22.2 Outline of the workflow for DNA barcoding methods

information and geo-reference data (Blackwell 2011). This may either be due to non-availability of authentic samples or due to availability of old or highly degraded samples for DNA profiling.

One solution to the above-mentioned problem is to use epitype. An epitype is a specimen or illustration selected to serve as an interpretative type when the holotype, lectotype, or previously designated neotype, or all original material associated with a validly published name, is demonstrably ambiguous and cannot be critically identified for purposes of the precise application of the name of a taxon. It is mandatory to cite the holotype, lectotype, or nectotype which is supported by designated epitype (Hyde and Zhang 2008).

22.4.1.3 Lack of Universal Marker for Different Fungal Species

Identification of fungi species, which is the second-largest eukaryotic kingdom, is extremely limited due to the absence of a universal DNA barcode (Schoch et al. 2012; Kress and Erickson 2007). For a barcode locus to be considered universal, it must possess significant variation at species level along with a conserved flanking site for the development of universal PCR primers that can be used in a wide number

of taxonomic groups. Moreover, universal barcodes must be of short sequence length such that they can be easily extracted and amplified using current techniques. It has been reported that *CO1*, which is considered universal barcode in case of animals produces reliable results only in few clads of closely related species. In many cases, *CO1* produces inconsistent results and cloning is required (Prakash et al. 2017).

Alternatives to *CO1* include use of other barcode markers such as ITS marker, *rbcL* marker. However, the studies suggest that these markers are also limited to identifying a group of fungi groups. Another alternative to above-mentioned markers is the use of nuclear cistron for identification. The 18S subunit of nuclear rRNA is used in phylogenetic identification and classification (Raja et al. 2017b).

22.4.1.4 Lack of Databases and Mistakes in Published Studies of Fungal DNA Sequences

Well-curated databases with substantial accurate sequence data play an important role in the identification of fungal species. This need has led to explosion in available fungal databases. Such databases include (1) clinical, biochemical, morphology, and taxonomy-based databases, (2) gene sequence-based databases, (3) fungal strain genotyping databases, and (4) genome-based databases (Hawksworth and Wiltshire 2011). However, many studies contributed in such databases are submitted with insufficient taxonomic identification (Koljalg et al. 2014). Often, either taxonomic annotations are not up to date or most of the fungal species described have not been sequenced. Moreover, many species are unnamed or partially named. Therefore, it is crucial to develop a database based on DNA sequencing for proper and reliable identification of fungal species.

The frequency and monitoring of database is another problem associated with the databases. Often, the databases have usability issues and are not updated frequently, which increases limited data sharing. Another problem faced by the databases is that if inconsistencies in grant support. Grants are required for the development and advancement of accurately curated and networked databases; however, in lack of sufficient grants, the databases are not properly maintained. Also, there is no database available related to the succession rate of fungi which poses another challenge to crime investigators.

22.4.2 Issues Specific to Forensic Identification of Fungal Evidence

22.4.2.1 Less Use of Fungal Evidence in Crime Investigation

Fungal evidence in different forms have a great potential to be used in providing convictions; however, the potential is wasted because of lack of awareness and neglect on part of crime scene investigating officers and skilled mycologists (Hawksworth and Wiltshire 2011).

22.4.2.2 Degradation of Fungal DNA Due to Various Processing Methods and Digestion

In forensic cases, the fungal evidence encountered are often found in processed forms. Processing methods of fungal evidence include cooking, boiling, baking, and frying. In such cases, the DNA of fungi might get degraded, which might raise difficulties in the identification of respective fungi. Although cooking and other processing methods degrade DNA, many studies in other organisms including plants suggest that DNA can be successfully extracted. Often in accidental and homicidal cases, the edible mushrooms are generally consumed with poisonous mushrooms. In case of death due to poisoning, stomach contents are analyzed for species determination as mushrooms are poorly digested. In such cases, degradation of fungal DNA can be due to the digestion of edible and poisonous mushrooms in the stomach by acids and digestive juices. In such cases, it becomes difficult to use DNA for the identification of fungi.

22.4.2.3 Lack of Sufficient Studies

The problems mentioned above especially degradation and contamination of DNA of fungal evidence lead to the problem of lack in studies-related effect of different parameters such as cooking and digestion of extraction and identification of DNA. No information is currently available about the particular role of enzymatic and metabolic activities of species obtained from cadavers (Lima et al. 2012).

22.5 Conclusions

The use of forensic mycology in ascertaining the cause and time of death along with the location of corpses has increased in the last decade; however, its full potential is yet to be explored. Forensic mycology stands on the pillars of taxonomic identification and DNA barcoding. With various limitations faced by the taxonomic identification, it is imperative to use DNA barcoding for absolute identification. DNA barcoding has the potential to identify fungal species encountered in various forms (cooked and gastric aspirates) effectively and accurately. Many studies have identified mixed-species using sequencing methods. However, less work has been reported related to DNA-based identification in cases of fungi found as trace evidence to determine time since death and location of the corpse. There is a need to examine more related cases from different parts of the world having varying climatic conditions, so that the pattern and rate of fungal growth can be determined accurately. Also, due to the lack of proper database sequences, forensic mycology is at a developing stage. Where morphological data is inconclusive, fungal evidence can be analyzed through DNA barcoding for the species identification. DNA barcoding is set to facilitate improvement in the existing identification methods, and it will transform the field of forensic mycology.

Conflict of Interest Declared none.

References

Bellini E, Ambrosio E, Di Piazza S, Gabbrielli M, Zotti M (2015) microfungi on human corpse in the skeletisation stage. In: Seventh European academy of forensic science congress, Prague, 6–11 Sept, Czech Republic

Berger KJ, Guss DA (2005) Mycotoxins revisited: Part I. J Emerg Med 28(1):53–62. https://doi.org/10.1016/j.jemermed.2004.08.013

Blackwell M (2011) The fungi: 1,2,3... 5.1 million species? Am J Bot 98:426–438. https://doi.org/10.3732/ajb.1000298

Bresinsky H, Besl A (1990) A colour atlas of poisonous fungi: a handbook for pharmacists, doctors, and biologists. Wolf Publishing, London

Bridge P (2002) The history and application of molecular mycology. Mycologist 16:90–99. https://doi.org/10.1017/S0269-915X(02)00201-X

Cai Q, Tulloss RE, Tang LP, Tolgor B, Zhang P, Chen ZH, Yang ZL (2014) Multi-locus phylogeny of lethal amanitas: implications for species diversity and historical biogeography. BMC Evol Biol 14:143. https://doi.org/10.1186/1471-2148-14-143

Cecilia TM, Noemi CM (2017) The mycology as forensics tool. Adv Tech Biol Med 05:10–11. https://doi.org/10.4172/2379-1764.1000226

Chen Z, Zhang P, Zhang Z (2012) Investigation and analysis of 102 mushroom poisoning cases in Southern China from 1994 to 2012. Fungal Divers 64:123–131. https://doi.org/10.1007/s13225-013-0260-7

Dasmahapatra KK, Mallet J (2006) Taxonomy: DNA barcodes: recent successes and future prospects. Heredity (Edinb) 97:254–255. https://doi.org/10.1038/sj.hdy.6800858

Dentinger BTM, Margaritescu S, Moncalvo JM (2010) Rapid and reliable high-throughput methods of DNA extraction for use in barcoding and molecular systematics of mushrooms. Mol Ecol Resour 10:628–633. https://doi.org/10.1111/j.1755-0998.2009.02825.x

Diaz JH (2017) Amatoxin-containing mushroom poisonings: species, toxidromes, treatments, and outcomes. Wilderness Environ Med 29:111–118. https://doi.org/10.1016/j.wem.2017.10.002

Enjalbert F, Rapior S, Nouguier-Soule J, Guillon S, Amouroux N (2002) Treatment of amatoxin poisoning: 20-year retrospective study. J Toxicol Clin Toxicol 40:715–757. https://doi.org/10.1081/CLT-120014646

Epis S, Matinato C, Gentili G, Varotto F, Bandi C, Sassera D (2010) Molecular detection of poisonous mushrooms in different matrices. Mycologia 102(3):747–754. https://doi.org/10.3852/09-124

Erguven M, Yilmaz O, Deveci M, Aksu N, Dursun F, Pelit M, Cebeci N (2007) Mushroom poisoning. Indian J Pediatr 74:847–852

Fišer Pečnikar Z, Buzan EV (2014) 20 years since the introduction of DNA barcoding: from theory to application. J Appl Genet 55:43–52. https://doi.org/10.1007/s13353-013-0180-y

Frézal L, Leblois R (2008) Four years of DNA barcoding: current advances and prospects. Infect Genet Evol 8:727–736. https://doi.org/10.1016/j.meegid.2008.05.005

Fu X, Fu B, He Z, Gong M, Li Z, Chen Z (2017) Acute renal failure caused by Amanita oberwinklerana poisoning. Mycoscience 58:121–127. https://doi.org/10.1016/j.myc.2016.12.004

Gambaro V, Roda G, Visconti GL, Arnoldi S, Casagni E, Dell'Acqua L, Farè F, Paladino E, Rusconi C, Arioli S, Mora D (2016) DNA-based taxonomic identification of basidiospores in hallucinogenic mushrooms cultivated in "grow-kits" seized by the police: LC-UV quali-quantitative determination of psilocybin and psilocin. J Pharm Biomed Anal 125:427–432. https://doi.org/10.1016/j.jpba.2016.03.043

Garcia J, Costa VM, Costa AE, Andrade S, Carneiro AC, Conceição F, Paiva JA, de Pinho PG, Baptista P, de Lourdes Bastos M, Carvalho F (2015) Co-ingestion of amatoxins and isoxazoles-containing mushrooms and successful treatment: a case report. Toxicon 103:55–59. https://doi.org/10.1016/j.toxicon.2015.06.012

Gausterer C, Penker M, Krisai-greilhuber I, Stein C, Stimpfl T (2014) Rapid genetic detection of ingested Amanita phalloides. Forensic Sci Int Genet 9:66–71. https://doi.org/10.1016/j.fsigen.2013.11.002

Gicquel T, Lepage S, Fradin M, Tribut O, Duretz B, Morel I (2014) Amatoxins (α- and β-amanitin) and phallotoxin (phalloidin) analyses in urines using high-resolution accurate mass LC – MS technology. J Anal Toxicol 38(6):335–340. https://doi.org/10.1093/jat/bku035

Gonmori K, Fujita H, Yokoyama K, Watanabe K, Suzuki O (2011) Mushroom toxins: a forensic toxicological review. Forensic Toxicol 29(2):85–94. https://doi.org/10.1007/s11419-011-0115-4

González V, Arenal F, Platas G, Esteve-Raventós F, Peláez F (2002) Molecular typing of Spanish species of Amanita by restriction analysis of the ITS region of the DNA. Mycol Res 106:903–910. https://doi.org/10.1017/S0953756202006305

Gurbuz S, Oguzturk H, Turgut K, Turtay MG, Guven T (2015) Mushrooms: the velvety poison. Acta Medica Mediterr 31:947–951

Harper KA, Smart CD, Davis RM (2011) Development of a DNA-based macroarray for the detection and identification of Amanita species. J Forensic Sci 56(4):1003–1009. https://doi.org/10.1111/j.1556-4029.2011.01739.x

Hawksworth DL, Wiltshire PEJ (2011) Forensic mycology: the use of fungi in criminal investigations. Forensic Sci Int 206:1–11. https://doi.org/10.1016/j.forsciint.2010.06.012

Hebert PDN, Gregory TR (2005) The promise of DNA barcoding for taxonomy. Syst Biol 54:852–859. https://doi.org/10.1080/10635150500354886

Hebert PDN, Cywinska A, Ball SA, de Waard JR (2003) Biological identifications through DNA barcodes. Proc Biol Sci 270:313–321. https://doi.org/10.1098/rspb.2002.2218

Hibbett D, Abarenkov K, Koljalg U, Opik M, Chai B, Cole J, Wang Q, Crous P, Robert V, Helgason T, Herr JR, Kirk P, Lueschow S, O'Donnell K, Nilsson RH, Oono R, Schoch C, Smyth C, Walker DM, Porras-Alfaro A, Taylor JW, Geiser DM (2016) Sequence-based classification and identification of Fungi. Mycologia 108(6):1049–1068. https://doi.org/10.3852/16-130

Himmelmann A, Mang G, Schnorf-Huber S (2001) Lethal ingestion of stored Amanita phalloides mushrooms. Swiss Med Wkly 131(41–42):616–617

Hosukler E, Erkol Z, Petekkaya S, Gundogdu V, Samurcu H (2018) Fungal growth on a corpse: a case report. Rom J Leg Med 26:158–161. https://doi.org/10.4323/rjlm.2018.158

Hyde KD, Zhang Y (2008) Epitypification: should we epitypify? J Zhejiang Unic Sci B 9:842–846. https://doi.org/10.1631/jzus.B0860004

Innis MA, Gelfand DH, Sninsky JJ, White TJ (1990) PCR protocols: a guide to methods and applications. Academic Press, Inc., Harcourt Brace Jovanovich, Publishers, New York

Ishii K, Hitosugi M, Kido M, Yaguchi T, Nishimura K, Hosoya T, Tokudome S (2006) Analysis of fungi detected in human cadavers. Legal Med 8:188–190. https://doi.org/10.1016/j.legalmed.2005.12.006

Jansson D, Wolterink A, Bergwer L, Hough P, Geukens K (2018) Source attribution pro fi ling of five species of Amanita mushrooms from four European countries by high resolution liquid chromatography-mass spectrometry combined with multivariate statistical analysis and DNA-barcoding. Talanta 186:636–644. https://doi.org/10.1016/j.talanta.2018.03.069

Kang E, Cheong KY, Lee MJ, Kim S, Shin GT, Kim H, Park IW (2015) Severe but reversible acute kidney injury resulting from Amanita punctata poisoning. Kidney Res Clin Pract 34:233–236. https://doi.org/10.1016/j.krcp.2015.05.007

Karlson-Stiber C, Persson H (2003) Cytotoxic fungi – an overview. Toxicon 42:339–349. https://doi.org/10.1016/S0041-0101(03)00238-1

Keller T, Schneider A, Regenscheit P, Dirnhofer R, Rücker T, Jaspers J, Kisser W (1999) Analysis of psilocybin and psilocin in *Psilocyne subcubensis* Guzmán by ion mobility spectrometry and gas chromatography-mass spectrometry. Forensic Sci Int 99(2):93–105. https://doi.org/10.1016/s0379-0738(98)00168-6

Khaund P, Joshi SR (2014) DNA barcoding of wild edible mushrooms consumed by the ethnic tribes of India. Gene 550:123–130. https://doi.org/10.1016/j.gene.2014.08.027

Koljalg U, Nilsson RH, Abarenkov K, Tedersoo L, Taylor AF, Bahram M, Bates ST, Bruns TD, Bengtsson-Palme J, Callaghan TM, Dauglas B, Drenkhan T, Eberhardt U, Duenas M, Grebenc T, Griffith GW, Hartmann M, Kirk M, Kohout P, Larsson E, Lindahl BD, Lucking R, Martin MP, Matheny PB, Nguyen NH, Niskanen T, Oja J, Peay KG, Peintner U, Peterson M, Poldmaa K, Saag L, Saar I, Schubler A, Scott JA, Senes C, Smith ME, Suija A, Taylor DL, Telleria MT, Weiss M, Larsson KH (2014) Towards a unified paradigm for sequence-based identification of fungi. Mol Ecol 22:5271–5277. https://doi.org/10.1111/mec.12481

Kowalczyk M, Sekuła A, Mleczko P, Olszowy Z, Kujawa A, Zubek S, Kupiec T (2015) Practical aspects of genetic identification of hallucinogenic and other poisonous mushrooms for clinical and forensic purposes. Croat Med J 56:32–40. https://doi.org/10.3325/cmj.2015.56.32

Kress WJ, Erickson DL (2007) A two-locus global DNA barcode for land plants: the coding *rbc*L gene complementa the non-coding *trn*H-*psb*A spacer region. PLoS One 2(6):e508. https://doi.org/10.1371/journal.pone.0000508

Lee JCI, Cole M, Linacre A (2000) Identification of members of the genera Panaeolus and Psilocybe by a DNA test: a preliminary test for hallucinogenic fungi. Forensic Sci Int 112:123–133. https://doi.org/10.1016/S0379-0738(00)00181-X

Li P, Deng W, Li T (2014) The molecular diversity of toxin gene families in lethal Amanita mushrooms. Toxicon 83:59–68. https://doi.org/10.1016/j.toxicon.2014.02.020

Lima AD, Costa Fortes R, Carvalho Garbi Novaes MR, Percário S (2012) Poisonous mushrooms; a review of the most common intoxications. Nutr Hosp 27(2):402–408. https://doi.org/10.3305/nh.2012.27.2.5328

Linacre A, Cole M, Lee JC (2002) Identifying the presence of “magic mushrooms” by DNA profiling. Sci Justice 42(1):50–54. https://doi.org/10.1016/S1355-0306(02)71797-8

Maeta K, Ochi T, Tokimoto K, Shimomura N, Maekawa N, Kawaguchi N, Nakaya M, Kitamoto Y, Aimi T (2008) Rapid species identification of cooked poisonous mushrooms by using real-time PCR. Appl Environ Microbiol 74(10):3306–3309. https://doi.org/10.1128/AEM.02082-07

Margarida A, Félix C, de Maria LB, Pinho PG, Carvalho M (2015) The hallucinogenic world of tryptamines: an updated review. Arch Toxicol:1151–1173. https://doi.org/10.1007/s00204-015-1513-x

Maruyama T, Kawahara N, Yokoyama K, Makino Y, Fukiharu T, Goda Y (2006) Phylogenetic relationship of psychoactive fungi based on rRNA gene for a large subunit and their identification using the TaqMan assay (II). Forensic Sci Int 163:51–58. https://doi.org/10.1016/j.forsciint.2004.10.028

Matsushima Y, Eguchi F, Kikukawa T, Matsuda T (2009) Historical overview of psychoactive mushrooms. Inflamm Regener 29(1):47–58. https://doi.org/10.2492/inflammregen.29.47

Mouhamadou B, Carriconde F, Gryta H, Jargeat P, Manzi S, Gardes M (2008) Molecular evolution of mitochondrial ribosomal DNA in the fungal genus Tricholoma: barcoding implications. Fungal Genet Biol 45:1219–1226. https://doi.org/10.1016/j.fgb.2008.06.006

Nagy J, Veress T (2016) HPLC analysis of hallucinogenic mushroom alkaloids (psilocin and psilocybin) applying hydrophilic interaction chromatography (HILIC). J Forensic Res 7:356. https://doi.org/10.4172/2157-7145.1000356

Nilsson RH, Tedersoo L, Abarenkov K, Ryberg M, Kristiansson E, Hartmann M, Schoch CL, Nylander JAA, Bergsten J, Porter TM, Jumpponen A, Vaishampayan P, Ovaskainen O, Hallenberg N, Bengtsson-Palme J, Eriksson KM, Larsson KH, Larsson E, Kõljalg U (2012) Five simple guidelines for establishing basic authenticity and reliability of newly generated fungal ITS sequences. MycoKeys 4:37–63. https://doi.org/10.3897/mycokeys.4.3606

Nugent NG, Saville BJ (2004) Forensic analysis of hallucinogenic fungi: a DNA-based approach. Forensic Sci Int 140:147–157. https://doi.org/10.1016/j.forsciint.2003.11.022

O’Brien BL, Khuu L (1996) A fatal Sunday brunch: Amanita mushroom poisoning in a gulf coast family. Am J Gastroentrol 91:581–583

Oda T, Tanaka C, Tsuda M (1999) Molecular phylogeny of Japanese Amanita species based on nucleotide sequences of the internal transcribed spacer region of nuclear ribosomal DNA. Mycoscience 40(1):57–64

Oda T, Tanaka C, Tsuda M (2004) Molecular phylogeny and biogeography of the widely distributed Amanita species, *A. muscaria* and *A. pantherina*. Mycol Res 108:885–896. https://doi.org/10.1017/S0953756204000620

Parnmen S, Sikaphan S, Leudang S, Boonpratuang T, Rangsiruji A, Naksuwankul K (2016) Molecular identification of poisonous mushrooms using nuclear ITS region and peptide toxins: a retrospective study on fatal cases in Thailand. J Toxicol Sci 41:65–76. https://doi.org/10.2131/jts.41.65

Piazza SD, Zotti M, Barranco R, Cecchi G, Greco G, Ventura F (2018) Post-mortem fungal colonization pattern during 6 weeks: two case studies. Forensic Sci Int 289:e18–e23. https://doi.org/10.1016/j.forsciint.2018.05.037

Prakash PY, Irinyi L, Halliday C, Chen S (2017) Online databases for the taxonomy and identification of pathogenic fungi and proposal for a cloud-based dynamic data network platform. J Clin Microbiol 55:1011–1024. https://doi.org/10.1128/JCM.02084-16

Raja HA, Miller AN, Pearce CJ, Oberlies NH (2017a) Fungal identification using molecular tools: a primer for the natural products research community. J Nat Prod 80(3):756–770. https://doi.org/10.1021/acs.jnatprod.6b01085

Raja HA, Baker TR, Little JG, Oberlies NH (2017b) DNA barcoding for identification of consumer-relevant mushrooms: a partial solution for product certification? Food Chem 214:383–392. https://doi.org/10.1016/j.foodchem.2016.07.052

Ratnasingham S, Hebert PDN (2007) Barcoding, bold: the barcode of life data system. Mol Ecol Notes 7:355–364. https://doi.org/10.1111/j.1471-8286.2006.01678.x. www.barcodinglife.org

Santi L, Maggioli C, Mastroroberto M, Tufoni M, Napoli L, Caraceni P (2012) Acute liver failure caused by amanita phalloides poisoning. Int J Hepatol 2012:487480. https://doi.org/10.1155/2012/487480

Sarwar M, Mcdonald JL (2003) A rapid extraction and GC/MS methodology for the identification of psilocyn in mushroom/chocolate concoctions. Microgram J 1:177–183

Savolainen V, Cowan RS, Vogler AP, Roderick GK, Lane R (2005) Towards writing the encyclopedia of life: an introduction to DNA barcoding. Philos Trans R Soc Lond Ser B Biol Sci 360:1805–1811. https://doi.org/10.1098/rstb.2005.1730

Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. PNAS 109(16):6241–6246. https://doi.org/10.1073/pnas.1117018109

Schoch CL, Robbertse B, Robert V, Vu D, Cardinali G, Irinyi L, Meyer W, Nilsson RH, Hughes K, Miller AN, Kirk PM, Abarenkov K, Aime MC, Ariyawansa HA, Bidartondo M, Boekhout T, Buyck B, Cai Q, Chen J, Crespo A, Crous PW, Damm U, De Beer ZW, Dentinger BTM, Divakar PK, Dueñas M, Feau N, Fliegerova K, García MA, Ge ZW, Griffith GW, Groenewald JZ, Groenewald M, Grube M, Gryzenhout M, Gueidan C, Guo L, Hambleton S, Hamelin R, Hansen K, Hofstetter V, Hong SB, Houbraken J, Hyde KD, Inderbitzin P, Johnston PR, Karunarathna SC, Kõljalg U, Kovács GM, Kraichak E, Krizsan K, Kurtzman CP, Larsson KH, Leavitt S, Letcher PM, Liimatainen K, Liu JK, Lodge DJ, Luangsa-Ard JJ, Lumbsch HT, Maharachchikumbura SSN, Manamgoda D, Martín MP, Minnis AM, Moncalvo JM, Mulè G, Nakasone KK, Niskanen T, Olariaga I, Papp T, Petkovits T, Pino-Bodas R, Powell MJ, Raja HA, Redecker D, Sarmiento-Ramirez JM, Seifert KA, Shrestha B, Stenroos S, Stielow B, Suh SO, Tanaka K, Tedersoo L, Telleria MT, Udayanga D, Untereiner WA, Uribeondo JD, Subbarao KV, Vágvölgyi C, Visagie C, Voigt K, Walker DM, Weir BS, Weiß M, Wijayawardene NN, Wingfield MJ, Xu JP, Yang ZL, Zhang N, Zhuang WY, Federhen S (2014) Finding needles in haystacks: linking scientific names, reference specimens and molecular data for fungi. Database (Oxford) 2014:bau061. https://doi.org/10.1093/database/bau061

Schwarz P, Dannaoui E, Gehl A, Felske-zech H, Birngruber CG, Dettmeyer RB, Verhoff MA (2014) Molecular identification of fungi found on decomposed human bodies in forensic autopsy cases. Int J Legal Med 129(4):785–791. https://doi.org/10.1007/s00414-014-1118-6

Stra J, Kallifatidis B, Borovic J (2014) Fluorescent random amplified microsatellites (F-RAMS) analysis of mushrooms as a forensic investigative tool. Forensic Sci Int Genet 9:25–32. https://doi.org/10.1016/j.fsigen.2013.10.009

Sugawara R, Yamada S, Tu Z, Sugawara A, Suzuki K (2016) Rapid and reliable species identification of wild mushrooms by matrix assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS). Anal Chim Acta 934:163–169. https://doi.org/10.1016/j.aca.2016.05.056

Tibbett M, Carter DO (2003) Mushrooms and taphonomy: the fungi that mark woodland graves. Mycologist 17(1):20–24. https://doi.org/10.1017/S0269915X03001150

Tranchida MC, Centeno ND, Cabello MN (2014) Soil fungi: their potential use as a forensic tool. J Forensic Sci 59(3):785–789. https://doi.org/10.1111/1556-4029.12391

Tranchida MC, Centeno ND, Stenglein SA (2016) First record of *Talaromyces udagawae* in soil related to decomposing human remains in Argentina. Rev Argent Microbiol 48(1):86–90. https://doi.org/10.1016/j.ram.2015.10.002

Tranchida MC, Berruezo LMB, Stenglein SA, Cabello MN (2018) Mycobiota associated with human cadavers: first record in Argentina. J Can Soc Forensic Sci 51(2):39–47. https://doi.org/10.1080/00085030.2018.1463131

Tsujikawa K, Kanamori T, Iwata Y, Ohmae Y, Sugita R, Inoue H, Kishi T (2003) Morphological and chemical analysis of magic mushrooms in Japan. Forensic Sci Int 138:85–90. https://doi.org/10.1016/j.forsciint.2003.08.009

Tsujikawa K, Kuwayama K, Miyaguchi H, Kanamori T (2007) Determination of muscimol and ibotenic acid in Amanita mushrooms by high-performance liquid chromatography and liquid chromatography-tandem mass spectrometry. J Chromatogr B Analyt Technol Biomed Life Sci 852(1–2):430–435. https://doi.org/10.1016/j.jchromb.2007.01.046

Vilgalys R, Hester M (1990) Rapid genetic identification and mapping of enzymatically amplified ribosomal DNA from several *Cryptococcus* species. J Bacteriol 172(8):4238–4246

Wasson R (1959) The hallucinogenic mushrooms of Mexico: an adventure in ethnomycological exploration. Trans New York Acad 21(Series II). https://doi.org/10.1111/j.2164-0947.1959.tb00681.x

Weiß M, Yang ZL, Oberwinkler F (1998) Molecular phylogenetic studies in the genus Amanita. Can J Bot 76:1170–1179

White TJ, Bruns TD, Lee SB, Taylor JW (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics, PCR protocols: a guide to methods and applications. Academic, San Diego, pp 315–322

White RL, Wentzell PD, Beasy MA, Clark DS, Grund DW (1993) Taxonomy of Amanita mushrooms by pattern recognition of amino acid chromatographic data. Anal Chim Acta 277:333–346. https://doi.org/10.1016/0003-2670(93)80446-R

Wiltshire PEJ, Hawksworth DL, Webb JA, Edwards KJ (2014) Palynology and mycology provide separate classes of probative evidence from the same forensic samples: a rape case from southern England. Forensic Sci Int 244:186–195. https://doi.org/10.1016/j.forsciint.2014.08.017

Zhang L, Yang J, Yang Z (2004) Molecular phylogeny of eastern Asian species of Amanita (Agaricales, Basidiomycota): taxonomic and biogeographic implications. Fungal Divers 17:219–238

Zhang P, Chen ZH, Xiao B, Tolgor B, Bao HY, Yang ZH (2010) Lethal amanitas of East Asia characterized by morphological and molecular data. Fungal Divers 42:119–133. https://doi.org/10.1007/s13225-010-0018-4

Zhou Q, Tang SS, He ZM, Luo T, Chen ZH, Zhang P (2017) Amatoxin and phallotoxin concentrations in Amanita fuliginea: influence of tissues, developmental stages and collection sites. Mycoscience 58:267–273. https://doi.org/10.1016/j.myc.2017.03.003

Zuber A, Kowalczyk M, Sekuła A, Mleczko P, Kupiec T (2011) Methods used in species identification of hallucinogenic and other poisonous mushrooms in forensic investigations. Z Zagadnien Nauk Sadowych 86:151–161

Part V

Advances in Forensic DNA Typing

Applications of Next-Generation Sequencing in Forensic Field

23

María Saiz, Luis Javier Martínez-Gonzalez, Jose Antonio Lorente, and Maria Jesus Alvarez-Cubero

Abstract

Next-generation sequencing (NGS) technology has developed and improved rapidly in the recent years, and many research fields use this methodology for its application. In the forensic field, new strategies and opportunities have emerged by the use of NGS which could be useful in analyzing multiple loci in certain samples and scenes. NGS overtakes the limitations of using the most common technologies such as capillary electrophoresis, STR, or SNP markers analysis for human genetic DNA typing. Nowadays, there are more equipment, analysis strategies, or even regions of analysis such as microbiome, miRNAs, or other markers. Here we report the evolution of NGS in forensic field, the workflow, and the opportunities of these methodologies in forensic research.

M. Saiz
Laboratory of Genetic Identification, Department of Legal Medicine, Toxicology and Physical Anthropology, Faculty of Medicine, University of Granada, Granada, Spain

L. J. Martínez-Gonzalez
Centre for Genomics and Oncological Research, GENYO (University of Granada-Andalusian Government—Pfizer), Granada, Spain

J. A. Lorente (✉)
Laboratory of Genetic Identification, Department of Legal Medicine, Toxicology and Physical Anthropology, Faculty of Medicine, University of Granada, Granada, Spain

Centre for Genomics and Oncological Research, GENYO (University of Granada-Andalusian Government—Pfizer), Granada, Spain
e-mail: jose.lorente@genyo.es

M. J. Alvarez-Cubero
Centre for Genomics and Oncological Research, GENYO (University of Granada-Andalusian Government—Pfizer), Granada, Spain

Department of Biochemistry & Molecular Biology III, Faculty of Medicine, University of Granada, Granada, Spain

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_23

Keywords

NGS · STRs · Epigenetic · miRNA · Microbiome

23.1 Introduction

23.1.1 What Is It?

Within the area of biomedicine, the techniques associated with molecular genetics, highlighting DNA sequencing, have enormously advanced in the last two decades. This big progress in molecular biology techniques led to next-generation sequencing (NGS) development. New sequencing technologies are characterized by performing the massive parallel sequencing (MPS), where thousands of reads are produced at the same time, reads coming from DNA fragment (Williams 2012).

Main technologies of NGS equipment are principally massive parallel chemical reactions, detection systems, and computational approaches to analyze sequences. These groundbreaking scientific advances have drastically decreased sequencing cost and reduced the time to a few days. Now, the amount of data derived from NGS platform is facing another challenge associated with the problem of millions of confusing results (Wong et al. 2013). Nevertheless, these know-hows have the potential to endorse another revolution in sequencing and its use (directly sequencing single molecules like DNA, RNA or even protein) and the great challenge to the substantially increase of data storage and processing (Han et al. 2014; van Dijk et al. 2014).

NGS technologies are principally applied in biomedical areas which have allowed finding the bases of some rare syndromes. Even so, developments in prenatal diagnosis are not comparable with the ones in genetics identification where there is a great variability of samples and not enough NGS data for each evidence type. Moreover, they also help to improve research in cancer diagnosis and genetically heterogeneous disorders as well as common diseases. Also, NGS has been frequently applied into the whole mitochondrial sequencing in forensic sciences. Furthermore, when a low amount of DNA is available, NGS has also found great achievements, such as in noninvasive prenatal diagnosis and in many forensic cases. So, the advances that have occurred in prenatal diagnosis can be transferred to the forensic field. Fetus DNA can be found in maternal blood, even though it is of low quality, quantity, and its difficulty to be distinguished from maternal DNA. So, the huge power of NGS seems striking for noninvasive prenatal testing. It is possible to detect trisomies with noninvasive sample collection (Buermans and den Dunnen 2014). In the last decade, advanced research approaches to circulating cell-free fetal DNA have made it possible to determine fetal rhesus D (RhD) genotype, fetal sex, aneuploidies, detection, and microdeletions of inherited paternal monogenic disorders. NGS will allow detection from hereditary mutations that cause genetic diseases (Breveglieri et al. 2019).

We should consider some ancient DNA (aDNA) reviews where the developments of DNA isolation methods combined with and short sequences, generated by NGS, have helped to increase its accessibility (Hofreiter et al. 2015). The elevated grade of sample degradation and the limited amount of DNA are still the principal obstacle for applying NGS systems in forensic fields. However, there are some studies that have also considered enriching aDNA (ancient DNA) using custom project-specific SNP panels or detecting specific genomic areas (Alvarez-Cubero et al. 2017).

23.1.2 Evolution

In 1977, Frederick Sanger announced a new method to study the DNA molecule known as Sanger sequencing; this new sequencing method was based on a chain-termination technique (Sanger et al. 1977). This was the primary technology of sequencing or the "first generation" for research or diagnostic sequencing applications (Collins et al. 2004; Liu et al. 2012). In 2003, the "Human Genome Project" concludes all the human genome thanks to semiautomatic capillary electrophoresis sequencers and Sanger technology (Collins et al. 2004; International Human Genome Sequencing Consortium 2004). The "Human Genome Project" verified that whole-genome sequencing (WGS) could be performed, but not routinely, due to its high economic and time costs. NGS has accelerated the development of WGS, decreasing the effort and high cost associated (Liu et al. 2012).

In 2005, the first NGS system appears as 454, it was carried out and launched by 454 and acquired by Roche in 2007, a year before Solexa developed a Genome Analyzer which would be purchased by Illumina afterwards. The same year Agencourt announced its new equipment, SOLiD (Sequencing by Oligo Ligation Detection). In 2011, Life Technologies introduced a new system based on the detection of hydrogen ions, the Ion Torrent Personal Genome Machine (PGM). These were, and still are, the most extended NGS sequencers in all fields, with Illumina instruments at the top. All of them share a satisfactory accuracy, a high throughput and a lower cost compared to Sanger sequencing (Liu et al. 2012). All these NGS platforms (categorized as Second Generation of sequencer, 2G) share the basic idea to perform a ligation of DNA fragments with specific adapter sequences. When DNA has adapters at both ends (libraries), they are linked to a surface (solid surface or microsphere), then amplified and sequenced by diverse methods, resulting in millions of reads.

Third-generation (3G) instruments are the evolution of NGS platforms. 3G systems use single-molecule, DNA or RNA, templates without PCR, due to requiring less starting material, and produce longer readings. 3G platforms include PacBio (Pacific Biosciences) and HeliScope (Helicos Biosciences) (Ku and Roukos 2013). An enzymatic template replication system is used to sequence individual clonal molecules. Genomic DNA template is captured, and the incorporation of a new base is controlled by breakup of fluorescent dye-linked pyrophosphate in the volume-limited observation window (Gut 2013). 3G instruments had some sensitivity problems, especially in some genome regions. PacBio had no sequencing problems

in regions with 100% GC content of CGG trinucleotide repeat expansions; these regions are very interesting in forensic studies (Loomis et al. 2013). Nowadays, long-read technologies are removing initial limitations in throughput and accuracy and also expanding their application domains in genomics.

4G platforms are also available such as Oxford Nanopore Technologies (ONT). Nanopore sequencing has evolved rapidly and recently. This technology is based on the straight reading of individual molecules incorporating nanopore technology to single-molecule-Seq (Ku and Roukos 2013). In forensic sciences, the precision is very important, for specific studies, low accuracy in 4G sequences can be enough to make a difference so we need higher accuracy technologies. PCR does not occur in single-molecule systems; therefore, 4G technologies do not produce artifacts like unbalanced amplification and GC-bias, and as a result, they produce homogeneous coverage and span GC-rich regions. Besides, they may span duplications or repeated structures that cannot be determined using short read sequencing (Buermans and den Dunnen 2014).

The possibility of in situ sequencing seems to be the next step in the evolution. Carrying out an experiment in an autopsy room or at a crime scene opens the doors to discovering RNA at a specific moment, cell, or tissue (Ku and Roukos 2013).

23.2 Workflow and Main Platforms of NGS

23.2.1 Workflow in NGS Sample Processing

NGS instruments need hard protocols for samples analysis. Library preparation goes from a multistep procedure starting with genomic template fragmentation, end repair, purifications, and adapters ligation, but it can become a very complex process like enrichment of targeted region of DNA or RNA. The common steps within the workflow of the library construction, which are described in Fig. 23.1, include (a) DNA/RNA fragmentation, (b) fragmented products end repairing and

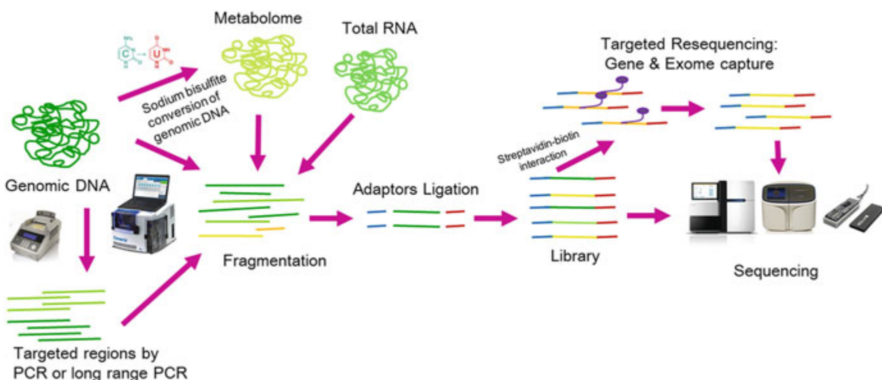


Fig. 23.1 Workflow in NGS sample preparation

A-tailing, (c) adapters ligation, (d) size selection, and (e) if necessary, amplification and, generally, after fragmentation, RNA library will require a new reverse transcription step.

23.2.2 Main NGS Platforms

Currently, there are a wide variability of NGS instruments, Ion Torrent (semiconductor sequencing technology) platform: PGM; Proton; S5, S5 plus; and S5 prime (managing 50Gb/day of sequences up to 600 bp). PacBio systems: RS II and Sequel System, Advance genomics with single-molecule real-time (SMRT) sequencing which is the longest read lengths available (average > 10,000 bp and some reads >60,000 bp), but both systems and sequences are expensive (Sequel System—PacBio n.d.). In recent years, DNBSEQ technology has broken into sequencers system, MGI company offers to world trade instruments like DNBSEQ-G400, DNBSEQ-G50, and DNBSEQ-T7. But after all, the highest percentage of sequencing data are produced by Illumina instruments: iSeq 100, MiniSeq, MiSeq Series, Nexseq 550 Series, HiSeq 2000, and Novaseq 6000 (ordered by instrument cost and generally, just the opposite, cost of sequencing). All of them produce short paired-end sequences (2 × 150, except Miseq 2 × 300 pb) (Sequencing Platforms | Compare NGS platforms (benchtop, production-scale) n.d.). Among Illumina systems, we highlight the MiSeq FGx™, it is the first equipment fully validated for forensic science which is able to address thousands of forensic markers in a single test, through targeted sequencing of forensically relevant STRs and SNPs loci. Illumina kits combined with MiSeq FGxTM allow forensic laboratories to obtain genetic profiles from degraded, mixed, and limited DNA samples. This equipment produces high-quality, efficient forensic DNA profiles which assure sharing these profiles with others laboratories; moreover, this system performs mitochondrial sequencing which is also very demanded in forensic community. Finally, there are many other instruments but with an infrequent use, such as Oxford Nanopore systems: PromethION, GridION X5, MinION, and SmidgION; Helicos technologies; and seqll (tSMS technology, true single-molecule sequencing) (Products n.d.).

New technologies of sequencing allow an accurate processing of huge number of samples, including a sensitivity and specificity detection of differences on a population level, the discovery of novel responsible variants, and the verification of benign single/multiple-nucleotide polymorphisms. New studies such as human microbiome analysis, small RNA, noncoding RNAs, methylome analysis, ChIP sequencing, or methylated DNA immunoprecipitation force to realize new molecular designs to produce libraries for small inserts, repeated sequences, complex elements, etc. Thanks to this great development in NGS technologies and strategies, it will also improve their applications into forensic sciences and human identification.

Another field of technological development with relevance in forensic field is the single-cell sequencing studies. For NGS library preparation, it needs nanogram amounts of DNA, but a cell just contains 6–7 pg of DNA (Sanchez-Cespedes et al.

1998), so the whole-genome amplification (WGA) of the original DNA chains is a decisive step for single-cell sequencing. Methodologies developed for single-cell WGA, such as combine displacement pre-amplification and PCR amplification (indicated as PicoPLEX WGA Kit; Rubicon Genomics, Ann Arbor, MI, USA), degenerate-oligonucleotide-primed polymerase chain reaction (DOP-PCR) (such as marketed as WGA4 kit; Sigma-Aldrich, St. Louis, MO, USA), and multiple displacement amplification (MDA) (marketed as REPLI-g Single Cell Kit; QIAGEN, Germantown, MD, USA), allow us to examine the complete genome of a single cell, producing a library from this small amount of DNA. It is something incredible when talking about a crime scene sample or in the case of ancient remains.

In addition to the WGA workflow with single-cell also entails cell separation and equipment characterization. In the near future, these systems may be very useful for cases of multiple violations or in disaster victims' identifications. Now we could find a varied number of single-cell platforms. One of the first to appear and currently very widespread is the C1 (Fluidigm Corporation, CA, USA). C1 prepares single-cell templates DNA or mRNA sequencing, epigenetics, or miRNA expression. Its principal applications are identify rare cell types, survey cells diversity, or characterize cellular functions (Fluidigm | Single-Cell Analysis n.d.). Chromium Controller (10× Genomics, Inc., CA, USA) is a high-throughput automated barcoding and library construction for sequencing applications, equipment that coupled to Illumina sequencers, and allows the analysis of all types of nucleic acids, highlighting the analysis of single cell CNVs; (Home—10× Genomics n.d.). BD Rhapsody™ System (Becton, Dickinson and Company, NJ, USA) profiles the gene expression of thousands of single cells; using predesigned or customized assays, it reduces the experimentation time and the sequencing cost (BD Rhapsody™ Single-Cell Analysis System—BD n.d.). Cyto-Mine® (Sphere Fluidics Limited; CA, UK) is able to process up to ten million of varied mammalian cells in not more than half a daytime. Its workflow lets discriminating the screening of single cells to discover less common candidates (Products | Single Cell Analysis System | Cyto-Mine® n.d.). Finally, a solution that starts from Bio Rad digital PCR, ddSEQ™ Single-Cell Isolator, it is part of Illumina® Bio-Rad® Single-Cell Sequencing Solution, this technology co-encapsulates single cells and barcodes into sub-nanoliter droplets, where cell lysis and barcoding occurs (ddSEQ™ Single-Cell Isolator | Life Science Research | Bio-Rad n.d.).

23.3 Applications in Human Identification and Forensic Sciences

The use of NGS methodologies in "de novo" analysis of whole genomes (like the whole mitochondrial DNA sequence of *Neophema chrysogaster*) and in repeating the sequencing of human genomes (Human Genome Project with 1092 genomes of human samples from various populations, was developed by the use of a mixture of exome sequencing and low-coverage whole-genome) has been extensively proven. Hence, it is easy to visualize the possible uses of NGS in forensic sciences, avoiding

the principal technical limits of present technologies like capillary electrophoresis and the high rate of degraded or restricted samples of forensic cases (Irwin et al. 2011a, b; Miller et al. 2013).

Several current studies have integrated NGS data obtained from human hair generating relevant data in metagenomic analyses. A study based on forty-two extracts of DNA obtained from pericranium and pubic hairs produced approximately 80 thousand of sequences that, after post control analyses and filtering, almost 40,000 were accessible. This type of studies would be relevant for the identification of victims of sexual crimes and rapist when there are several evidences (Tridico et al. 2014). There are not many attentions in researching of Indels and CNVs, although repeats cover up nearly the half of the human genome, and STRs just the 15%. The reason is that pyrosequencing were just developed in platforms with enough read length to sequence the nucleus of STR loci applied in forensic genetics, and in these days, the majority of the forensic literature with NGS STRs data were developed using pyrosequencing methodologies. In not may reports, sequencing by synthesis and recently also semiconductor sequencing were also applied for the construction of libraries directly by PCR or using adapter ligation.

23.3.1 Short Tandem Repeats (STRs) and Single-Nucleotide Polymorphisms (SNPs)

DNA profile by using STR typing is the most common method for forensic analysis, and DNA sequencing has usually been used to indicate the origin and relationship between DNA samples in small quantity by the analysis of hypervariable section of mitochondrial DNA (mtDNA) (Berglund et al. 2011).

STR study has become a very common methodology for human forensic cases (Irwin et al. 2011a, b; Rockenbauer et al. 2014). STRs are normally studied by multiplexed PCR continued by capillary electrophoresis (CE)-based separation (Jobling and Gill 2004). Moreover, normal STR typing gives enough discrimination power for many of the current applications, and it has been proven that just using routine STRs can solve mistakes in relationship analysis (Li et al. 2012; Tsai et al. 2013). Furthermore, it is not capable to distinguish in the production stutter data produced from mixtures (Butler et al. 2004; McNevin et al. 2005), and dye artifacts are usually found (Zeng et al. 2015). The use of autosomal SNPs for human classification in forensic identification also has several problems. Bi-allelic SNPs are not as polymorphic as multi-allelic STRs. In DNA mixtures, analysis from several individuals SNPs are not very informative, but the analysis of a incremented number of SNPs as well as the use of multiple genotyping methodologies may neutralize this event (Kayser and de Knijff 2011). Some SNP panels have been used for individual identification; Fudan ID Panel covers 175 SNP markers and has been improved and validated by Ion Torrent™ PGM (Li et al. 2017).

In recent times, MPS is very concerned for the forensic genetic community. It offers the option to discover several hundred to thousand markers at the same time. Using large-scale sequencing offers multiple enhancements to forensic science.

Lately, NGS has been applied to study nuclear microsatellite markers, where random sequencing of a small fraction of the genome has been proven to produce high density of potential microsatellite loci in a low cost and rapid process. The DNA commission of the International Society for Forensic Genetics has just taken some considerations for using typing, analysis, and naming MPS STRs for forensic uses (Parson et al. 2016).

Furthermore, it will be possible to complement STR data with other informative forensic markers obtained by large-scale sequencing (Berglund et al. 2011). New sequencing technologies present a completely new paradigm for sequence data generation, contributing the chance to sequence up to millions of individual DNA fragment (like in DNA mixtures), and the elevated throughput sequencing when compared with Sanger sequencing with a much lower cost. Using new instruments to detect STR amplicons will make it much more discriminative for genetic identification, and with a degraded DNA input of just 31 pg complete profiles could be generated, also we could obtain partial informative profiles out of 5 pg of DNA (Zeng et al. 2015; Elena et al. 2016).

SNPs and Indels could be found in the repeated unit of STRs or in the flanking region, if nucleotide substitutions, deletions, or insertions are sequenced, but these changes are not detectable by CE analysis (Rockenbauer et al. 2014). NGS identifies the number of repetitions presented in the STRs, but also the sequence of the polymorphisms presented in them (Novroski et al. 2016). Moreover, new technologies allow to generate individual sequences of the different alleles present in a STR amplicon mixture (Van Neste et al. 2012). More identifiable alleles lead to more statistical power study and decrease the necessary number of STR loci that is necessary to be analyzed to solve a case with a high probability of coincidence. In criminal cases, it may also be easier to solve sample mixtures, when the same alleles appear in CE analyses and may be latter identified with more information by NGS. The possibility to discriminate between persons with identical allele lengths as far as the characterization of variations could be crucial in forensic or affiliation analysis. Statistical power will be increased using NGS technologies with a reduced number of SNPs or STRs (Rockenbauer et al. 2014).

The STR Sequencing Project, known as STRSeq, was started to simplify the characterization of sequence-based alleles at STR loci targeted in commercial kits (Gettings et al. 2017). This tool entails a curated database of sequence diversity at identification of STR loci, in addition to the nomenclature, essential elements, and variations described in accordance with the current guidelines (Parson et al. 2016).

So, the power of new sequencing technologies can only be used by also including complex STRs, analyzing both SNPs and STRs in parallel. NGS instrument provides a possibility for creating an all-in-one multiplex with important identification markers that include, e.g., STRs, SNPs, InDels, and mtDNA markers (Irwin et al. 2011a, b). In recent years, some biotech corporations are developing focused NGS-Panel into the forensic area using the same genomics areas applied till now, helping researchers to study their samples instead of customizing their own panels and primers. Currently, Thermo Fisher has designed some kits focused on forensic samples, Precision ID Panels, all of them based on Ion PGM System: Precision ID

GlobalFiler™ NGS STR Panels and GlobalFiler™ PCR Amplification Kit, NGS and STRs kits respectively, including the same 21 autosomal STRs and amelogenin sex markers. However, this kit adds STR sequence motifs to the usual STR profile, isometric heterozygotes, and known SNPs in flanking regions (Wang et al. 2017). In addition, three other panels have been developed to investigate about the bio-geographical origin of the samples and discriminate individuals based on SNP analysis and mitochondrial DNA whole genome. The Illumina® ForenSeq™ DNA Signature kit includes 200 genetic loci (autosomal STRs, Y-STRs and X-STRs, ancestry informative SNPs, identity informative SNPs, and phenotypic informative SNPs). As the size of the targeted amplicons range from 64 to 231 bp for SNPs and 61 to 430 bp for STRs, this kit has been postulated as a good choice in MPS analysis of degraded human remains (Almohammed et al. 2017). Promega Corporation has developed the PowerSeq™ Systems Prototype Auto/Y analyzing 22 autosomal STR markers, amelogenin, and 23 Y-STR markers designed specifically for NGS on the MiSeq FGx® System (Montano et al. 2018; Silva et al. 2018). Many evaluation studies have been done in order to assess sensitivity, reproducibility, mixtures, and concordance as well as the analysis of casework and ancient DNA samples (Fattorini et al. 2017; Jäger et al. 2017; Just et al. 2017; Xavier and Parson 2017).

Thanks to the genome association studies, GWAs, it has been possible to identify SNPs for the characterization of the phenotype, indicating their ethnic origin and physical characteristics. These data are very useful in criminal cases because it can help recognize an individual, which will later be confirmed compared with a reference sample. Moreover, the capability to predict facial morphology and estimate individual-specific appearance through DNA is a very important goal to identify unknown persons (Irwin et al. 2011a, b; Rockenbauer et al. 2014). In order to achieve this great precision, it is most likely that it cannot be done with a few markers; on the contrary, a large set of genetic markers will be necessary. So, NGS and high-density arrays are useful for genotyping large numbers of SNPs, which can successfully deal with DNA degraded and in low quantity (Kayser and de Knijff 2011). Currently, a good application of NGS is discovering forensic SNP markers (Seo et al. 2013). TruSeqTM ChIP protocol has been modified and developed to detect 160 human identification, phenotypic SNPs, and ancestry; this new method is called TruSeqTM Forensic Amplicon (Warshauer et al. 2015a). This protocol is less labor-intensive than other techniques. In addition, low input DNA (1 ng) is required for library preparation (Warshauer et al. 2015a, b). Additionally, a genome-wide SNP array, containing 906,600 autosomal SNPs, Y-SNPs, and mitochondrial DNA has been developed by Affimetrix (Bridges et al. 2011).

The ability to produce a huge volume of data (in Novaseq more than 30 billion short reads per run) is the key advance of NGS technologies, obtaining these genetics data economically (Metzker 2009). Developing a software for the analysis of forensic NGS data will be one of the main challenges in genetics identification; it is needed because it is not possible to analyze the enormous amount of data manually. The application tool must be consistent and validated before the software can be applied in real case work. *Forenseq Universal Analysis Software* has been developed by Illumina, specifically designed to support forensic genomics

applications. Recently, there has been developed a new web application, *toaSTR*. It is a user-friendly tool for STR allele calling in massively parallel sequencing data independent of the system and the forensic kit used (Ganschow et al. 2017). The program differentiates automatically isoalleles from stutter and artefacts on a sequence bases; it facilitates an automatic allele calling with minimal need for review (Ganschow et al. 2017).

23.3.2 Lineage Markers

NGS has been used to discover systematically Y-STRs. 4500 Y-STRs have been genotyped, and the mutation rates have been estimated for 702 of them (Willems et al. 2016). Furthermore, sequence-based studies are beginning to reveal the full extent of structural variants (CNVs and inversions) and the great acceptance for gene loss on the Y chromosome compared with the autosomes (Massaia and Xue 2017) as far as sequence microvariance among Y-STRs (Warshauer et al. 2015a, b; Kwon et al. 2016; Iacovacci et al. 2017).

On the other hand, in Y-chromosome, the most interesting finding of NGS is the discovery of new Y-SNPs. These previously unknown Y-SNPs are relevant for Y-chromosomal phylogenies and haplogroup nomenclatures used for anthropological research and forensic identification. This finding has direct influence over phylogenies trees, appearing as new Y-chromosomal categories. Nevertheless, to include new Y-SNPs and updating and having a global consensus in forensic and anthropological sciences, a lot of work is still needed; most of the Y-chromosome recent project use Y phylogeny and Y-SNPs defined by Karafet et al. (2008). Due to the development of NGS analysis, new lineages have been described, and there is a need in the updating of the Y-chromosomal tree (Larmuseau et al. 2015). A new haplogroup has been described, A00, as the deepest-rooting known haplogroup in the Y-chromosome tree that diverged 275 thousand years ago (Mendez et al. 2013).

Many populations have been sequenced for several Y-SNPs in order to include all Y-haplogroups present and update the Y-Chromosome Consortium tree (Ochiai et al. 2016; Choi et al. 2017; Gao et al. 2017; Larmuseau et al. 2017). All the information derived from these studies has been accumulated, and the Y-chromosomal phylogenetic tree has been updated for criminal purposes (Van Geystelen et al. 2013).

mtDNA is present in higher number than nuclear DNA (nDNA), mitochondria vary in number and also mtDNA copies vary in mitochondria. So the probability of recovering useable DNA data is augmented in degraded samples that fail to yield helpful nDNA typing results (Parson et al. 2013). mtDNA is also informative of maternal biogeographic ancestry principally obtained by the HV1 and HV2 regions (Kayser and de Knijff 2011).

The study of entire mtDNA control region offers a random match probability (RMP) of 1 in 120 and has provided valuable evidence in many cases. The extension to examine the complete mitochondrial genome is a logical consequence and required goal to maximize the information content of mtDNA analyses (Irwin

et al. 2011a). A study on Chinese Han population has determined that the RMP probability decreases 4.12% when analyzing the mitochondrial Genome (mtGenome) compared to control region analyses (Zhou et al. 2016).

New sequencing technologies have the potential to radically increase sample throughput, workflow efficiency, and detection resolution, so helping to get reliable and accurate entire mtGenome information (Yang et al. 2014; Zhou et al. 2016). Analyzing the whole mitochondrial genome in individuals with the same control region sequences will recognize different samples by the analysis of personal polymorphisms in the coding region not studied yet by traditional methods (Holland et al. 2011).

There is a substantial number of different variants on mtGenome, mainly due because there are dozens of mtDNA copies in each mitochondrion and up to hundreds of mitochondria in some tissues. Studying all mutations presented in mtDNA in tissues could be slightly arduous. If we compare NGS data with those analyzed by capillary electrophoresis, we could see the percentage of variables found using each technique, and how many of these mutations are homoplasmic or heteroplasmic. Several studies developed by NGS show us that this methodology allowed detecting a high percentage of mutations. These mutations can also be presented in homoplasmic or heteroplasmic states. Thanks to a study performed with 20 cases, it was possible to recognize more than 400 individual nucleotide substitutions that include four heteroplasmic variants confirmed by capillary electrophoresis with high equivalence (98%) (Zaragoza et al. 2010). While Sanger sequencing can detect heteroplasmy at a threshold of 10–20% but not resolve the variants, MPS is able to both detect and conclude heteroplasmy at levels of 1–2% (Rathbun et al. 2017; Gallimore et al. 2018); heteroplasmy can be detected accurately in 1 per 10,000 mtGenome copies with Illumina GAII (He et al. 2010). A recent study that analyzed six pairs of adult monozygotic twins has observed point heteroplasmies in five sets of twins, and a single nucleotide variant was detected in four sets. This results give an evidence for the hypothesis that variants of the mtGenome could be a biomarker to distinguish monozygotic twins from each other (Wang et al. 2015).

New sequencing methods are transforming data generation and have the potential to generate whole mitochondrial genomes profiles from even highly degraded specimens. It is quite simple and cost-effective. The study of mtDNA coding region data will be included in many routine forensic caseworks, and its study will not be dictated by the quantity or quality of the sample. This data allow accessibility to population databases that can be used to determine uncommonness mitochondrial haplotypes.

Recently, Precision ID mtDNA Whole Genome Panel kit has been tested to study the mtGenome of three skeletons dated to about early eighteenth to mid-nineteenth century. The three skeleton samples had the same whole mtDNA sequences with 38 mutations as compared to the rCRS, being assigned to haplogroup D4a1c. These results suggested that the three skeletons might belong to the same maternal line (Hashiyada et al. 2017). A recent study performed on chemically treated, degraded, high-quality, and nonhuman samples showed NGS methods to be exceedingly

sensitive, capable of generating entire mtGenome data from samples that failed to yield reliable sequences with standard PCR-based techniques (Marshall et al. 2017).

Many software packages have been developed for mtDNA profile generation. Both mitoSAVE (King et al. 2014) and GeneMarker HTS (Holland et al. 2017) generate haplotypes consistent with current forensic nomenclature guidelines and apply used-defined thresholds for profile reporting from high-quality samples. On the other hand, AQME automates mtDNA analysis from sequence data to forensic profiles and offers mtDNA haplogroup assignment. It has shown to produce accurate forensic profiles for high-quality, degraded, and chemically treated samples (Sturk-Andreaggi et al. 2017). In addition, MitoSuite supports quality check of alignment data, building consensus sequences, variant annotation, detection of heteroplasmic sites and haplogroup classification, contamination, and base substitution patterns (Ishiya and Ueda 2017).

23.3.3 Microbiome

16S rRNA and metagenomic data "The Microbiome Project" enlisted around 300 people at two places in the USA, and proof in 15–18 body positions that characterizes skin, urogenital tract, gut, and oral cavity; studying them indicated that the microbial variety on the human body site was more similar to the similar human part on a different individual than to other body part on the same person (Clarke et al. 2017).

The human microbiota consists of 10–100 trillion symbiotic microbial cells harbored by each person and the genes these cells port. There are as many bacteria in our body as human cells (Hampton-Marcell et al. 2017). Understanding these genomes, we could understand the microbiome impacts on human health, but it is also interesting as innovative tool of genetic identification (Ursell et al. 2012). Two facts make microbiome as a relevant point for forensic identification: (a) the high number of these bacterial cells in a human body; (b) the subspecies level of these bacteria appears to be unique to each person. All of this offers an excellent chance to find a new identifiable marker unique to each person (Hampton-Marcell et al. 2017). There are unique methodologies that fit and classify some interesting taxa. In forensic field, microbial analysis focuses on identifying precise bacterial strains related to terrorism, illness, pollution, microbial postmortem variations, and trace evidence signs. The current use of NGS of taxonomically and/or phylogenetically useful genomic regions (i.e., 16S rRNA, 18S rRNA, and ITS) has optimized the excellent potential use of microbiome to estimate the postmortem intervals, to identify clandestine graves, and to unify people spaces or objects analyzing their skin microbes. But postmortem intervals are one of the most improved forensic investigations using metabolomics, by giving novel biomarkers linked to chronological variations after death, mainly to the high diversity and strength to unstable conditions to microbes (Ursell et al. 2012).

Currently, there are different focuses on NGS, mass spectrometry, and computational methods for improving the ability to characterize the microbial diversity

(microbiome); increasing the new applications of microbiome analysis such as, oral microbiome. Such profile may be unique to each individual and thereby useful for investigating genetics purposes. Microbiome will offer relevant data about each person's latest actions and about the place they live (Wong et al. 2013). Skin microbiome is also very personal; it is recognized that two individual's hands can vary around a 80% by the types of microbe discovered there. They are also good trace evidence, unique by person but can be easily transferred to objects associated with a given person, such as a computer keyboard, computer mouse, and cell phone, among others. Moreover, the structure of personalized skin microbial communities is stable over time although variable by season; and highly specific of person's gender and lifestyle (Metcalf et al. 2017). Nowadays, there is a commercial kit that are based on skin microbiome profiling for forensic human identification, such as hidSkinPlex, comprising of 286 bacterial (and phage) family-, genus-, species-, and subspecies-level markers (Schmedes et al. 2018).

Necrobiome or thanatomicrobiome (defined as the community of organisms implicated in the decomposition events of a human body) is also a novel strategy used in forensic caseworks, which improves the data of forensic entomology and the obtained data from the physical description of corpses (Hyde et al. 2013; Pechal et al. 2014). NGS analysis of the necrobiome could establish the stage of decomposition, its succession, as well as important bacteria taxa from these communities. Moreover, they are also useful tools as postmortem interval estimator (Guo et al. 2016) and details about paleoforensic allowing to deepen in the genomic development of DNA microbes, pathogens as well as microbioma of putrefaction (Gorgé et al. 2016). Recent data included *Clostridium* spp. prevailed at long postmortem intervals (up to 10 days) and that these Gram-positive, anaerobic extremophiles also prevailed at shorter postmortem intervals (4 h) (Javan et al. 2017).

Lastly, details about geographical microbiome locations could also be of interest in forensic field. It is known that soil samples from hundreds of sites in the area have a specific profile, so any remains of soil sample in a scene of crime could be informative (Hampton-Marcell et al. 2017).

There are many advances to date in NGS and microbiome analysis; however, many things should be implemented in the near future and much work to be done such as including microbiome in databases such as CODIS could prove an effective method to lowering crime rates and clearing cases (Hampton-Marcell et al. 2017).

23.3.4 Epigenetic Analysis

It was in 1942 when Waddington coined the term "epigenetics," which he defined as changes in phenotype without changes in genotype. Nowadays, we know that epigenetics mechanisms interact with gene expression pathways and programs that can canalize different cell-type identities, including details of histone modifications and DNA methylation (Allis and Jenuwein 2016).

Many recent researches of epigenetic are focused on diseases such as cancer or inflammation, because epigenomic events are relevant for cellular reprograming

(Allis and Jenuwein 2016). For example, an analysis demonstrated that KMT2D epigenetically activates PI3K/Akt path and EMT by linking both LIFR and KLF4 serving as a supposed epigenetic-based target for treating prostate cancer (Lv et al. 2017). There are also details in non-cancer diseases such as the analysis of patterns of DNA methylation and posttranslational histone modifications in chronic inflammatory diseases (Fogel et al. 2017).

In contrast, epigenetic has been explored slowly in forensic field. New researches are focused in the "epigenomic fingerprint" using epigenomic prediction of lifestyle and environmental exposures to improve DNA characterization in criminal cases where problems in DNA profiles identification are, complementing its predictions (Vidaki and Kayser 2017).

Many of the studies in epigenetic in forensic field are focused on markers for the prediction of tissue-identification and monozygotic twin bias, by using DNA methylation methodologies (methylation-sensitive restriction endonuclease, bisulfite modification, methylation-CpG linking protein, and third-generation sequencing methods, among others) (Kader and Ghai 2015).

In relation to the use of epigenetic in the estimation of human age, there is a study which analyses a total number of 27 CpG sites at three genetic loci (SCGN, Secretagogin; KLF14, Kruppel-Like Factor 14; and DLX5, distal-less homeobox 5 gene) in relation to their methylation status with age (samples range from 5 to 73 years). Finally, it is discovered that specific CpGs in SCGN and KLF14 can be used as potential epigenetic markers to estimate age using saliva and blood specimens, data with relevant interest in forensic field. However, there are still many challenges in finding universal DNA methylation markers using any body fluid, because many studies reported that DNA methylation is tissue specific. By the moment, just certain markers such as ELOVL2, which was describe as relatively stable age predictors for some cell types (Alghanim et al. 2017).

There are also data reporting the use of epigenetic for the discrimination of identical twins, demonstrating that there are twin-differentially methylated sites that can be useful in the case of twin genetic identification, that nowadays is currently impossible (Vidaki et al. 2017).

Many of the studies have been developed with great quantities of DNA, but promising reports also exist with few quantities of initial DNA (100 pg), which is relevant for forensic sciences (Yang et al. 2014). The application of epigenetic is in the initial steps in the forensic field, but promise results are described according to this data.

23.3.5 MicroRNA Analysis

In 1993 Lee was the first to describe microRNAs (miRNAs). They are defined as small RNA molecules encoded in the genome of plants and animals with highly conserved regions; their main role is to regulate the expression of genes (Lee et al. 1993). There are many data about their role in cancer development and evolution. Dr. Croce's group was the first to highlight miRNAs role in B-cell chronic

lymphocytic leukemia cells; describing for the first time miR-15a and miR-16-1 in this cancer. Currently it is known that the role of miRNAs in cancer is wide, and they can act in different ways depending on their target genes. miRNA could function as either oncogene or tumor suppressor under certain circumstances. Recent researches are focused in identifying miRNA profile in exosomes of cancer (Peng and Croce 2016).

miRNAs have also been recently introduced in forensic field. One of the principal significant things is their little size (8–22 nucleotides in extent) which highlights them as the most promising molecules for avoiding degradation and tissue-specificity or very tissue-different expression. They are fitting for forensic detection, body fluid classification, and postmortem interval (PMI) deduction analysis (Yang et al. 2014).

There are data exposing a set of miRNAs as validated ones for human identification in several body fluids (blood, saliva, semen, menstrual blood, and vaginal secretions) and skin such as miR10b, miR203, miR374, miR451, and miR943 as well as proper candidate housekeeping genes like miR26b, miR92, miR144, and miR484 for normalizing miRNA gene expression data (Sirker et al. 2017). Recent publications have also indicated that miR-451a and miR-142-3p were observed in venous profiles and miR-205 absent; in menstrual blood samples, miR-451a, miR-141-3p, and miR-205; miR-891a, miR-10b, miR-142-3p, and miR-205 were observed in all semen samples; in saliva, miR-142-3p and miR-205 were observed (Mayes et al. 2018).

There are also efforts in identifying miRNA organ-specific markers such as hsa-miR-219a-5p, hsa-miR-122-5p, hsa-miR-205-5p, hsa-miR-208b-3p, and hsa-miR-206 for the identification of brain, liver, skin, heart, and skeletal muscle, respectively. Further, hsa-miR-9-5p and hsa-miR-124-3p as well as hsa-miR-499a-5p, hsa-miR-1-3p, and hsa-miR-133a-3p were found to be promising markers for the identification of brain and muscle in general, respectively (Sauer et al. 2017). As the usefulness of miRNAs in the forensic field to determine the origin of the evidence is more and more proven, there has been developed some strategies to multiplex the analysis of miRNA for body fluid identification (Mayes et al. 2018). However, many challenges are still open for its wide application in human identification and database updating.

References

Alghanim H, Antunes J, Silva DSBS, Alho CS, Balamurugan K, McCord B (2017) Detection and evaluation of DNA methylation markers found at SCGN and KLF14 loci to estimate human age. Forensic Sci Int Genet 31:81–88. https://doi.org/10.1016/j.fsigen.2017.07.011

Allis CD, Jenuwein T (2016) The molecular hallmarks of epigenetic control. Nat Rev Genet 17:487. https://doi.org/10.1038/nrg.2016.59

Almohammed E, Zgonjanin D, Iyengar A, Ballard D, Devesse L, Sibte H (2017) A study of degraded skeletal samples using ForenSeq DNA Signature™ Kit. Forensic Sci Int Genet Suppl Ser 6:e410–e412. https://doi.org/10.1016/j.fsigss.2017.09.158

Alvarez-Cubero MJ, Saiz M, Martínez-García B, Sayalero SM, Entrala C, Lorente JA, Martinez-Gonzalez LJ (2017) Next generation sequencing: an application in forensic sciences? Ann Hum Biol 44(7):581–592. https://doi.org/10.1080/03014460.2017.1375155

BD Rhapsody™ Single-Cell Analysis System—BD (n.d.). https://www.bd.com/en-us/offerings/capabilities/genomics/bd-rhapsody-single-cell-analysis-system. Accessed 7 June 2018

Berglund EC, Kiialainen A, Syvänen AC (2011) Next-generation sequencing technologies and applications for human genetic history and forensics. Investig Genet 2(1):23

Breveglieri G, D'Aversa E, Finotti A, Borgatti M (2019) Non-invasive prenatal testing using fetal DNA. Mol Diagn Ther 23:291–299. https://doi.org/10.1007/s40291-019-00385-2

Bridges M, Heron EA, O'Dushlaine C, Segurado R, Morris D, Corvin A et al (2011) Genetic classification of populations using supervised learning. PLoS One 6(5):e14802. Retrieved from http://dx.doi.org/10.1371%2Fjournal.pone.0014802

Buermans HPJ, den Dunnen JT (2014) Next generation sequencing technology: advances and applications. Biochim Biophys Acta 1842(10):1932–1941. https://doi.org/10.1016/j.bbadis.2014.06.015

Butler JM, Buel E, Crivellente F, McCord BR (2004) Forensic DNA typing by capillary electrophoresis using the ABI Prism 310 and 3100 genetic analyzers for STR analysis. Electrophoresis 25(10–11):1397–1412. https://doi.org/10.1002/elps.200305822

Ku C-S, Roukos DH (2013) From next-generation sequencing to nanopore sequencing technology: paving the way to personalized genomic medicine. Expert Rev Med Dev 10(1):1–6. https://doi.org/10.1586/erd.12.63

Choi SS, Park KH, Nam DE, Kang TH, Chung KW (2017) Y-chromosome haplogrouping for Asians using Y-SNP target sequencing. Forensic Sci Int Genet Suppl Ser 6:e235–e237. https://doi.org/10.1016/J.FSIGSS.2017.09.100

Clarke TH, Gomez A, Singh H, Nelson KE, Brinkac LM (2017) Integrating the microbiome as a resource in the forensics toolkit. Forensic Sci Int Genet 30:141–147. https://doi.org/10.1016/j.fsigen.2017.06.008

Collins FS, Lander ES, Rogers J, Waterston RH, Conso I (2004) Finishing the euchromatic sequence of the human genome. Nature 431:931–945

ddSEQ™ Single-Cell Isolator | Life Science Research | Bio-Rad (n.d.). http://www.bio-rad.com/es-es/product/ddseq-single-cell-isolator?ID=OKNWBSE8Z. Accessed 7 June 2018

van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C (2014) Ten years of next-generation sequencing technology. Trends Genet 30(9):418–426. https://doi.org/10.1016/j.tig.2014.07.001

Elena S, Alessandro A, Ignazio C, Sharon W, Luigi R, Andrea B (2016) Revealing the challenges of low template DNA analysis with the prototype Ion AmpliSeq™ Identity panel v2.3 on the PGM™ Sequencer. Forensic Sci Int Genet 22:25–36. https://doi.org/10.1016/j.fsigen.2015.07.011

Fattorini P, Previderé C, Carboni I, Marrubini G, Sorçaburu-Cigliero S, Grignani P et al (2017) Performance of the ForenSeq TM DNA Signature Prep kit on highly degraded samples. Electrophoresis 38(8):1163–1174. https://doi.org/10.1002/elps.201600290

Fluidigm | Single-Cell Analysis (n.d.). https://www.fluidigm.com/applications/single-cell-analysis. Accessed 7 June 2018

Fogel O, Richard-Miceli C, Tost J (2017) Epigenetic changes in chronic inflammatory diseases. Adv Protein Chem Struct Biol 106:139–189. https://doi.org/10.1016/bs.apcsb.2016.09.003

Gallimore JM, McElhoe JA, Holland MM (2018) Assessing heteroplasmic variant drift in the mtDNA control region of human hairs using an MPS approach. Forensic Sci Int Genet 32:7–17. https://doi.org/10.1016/J.FSIGEN.2017.09.013

Ganschow S, Wiegand P, Tiemann C (2017) toaSTR: a web-based forensic tool for the analysis of short tandem repeats in massively parallel sequencing data. Forensic Sci Int Genet Suppl Ser 6: e119–e121. https://doi.org/10.1016/j.fsigss.2017.09.034

Gao T, Yun L, Zhou D, Lang M, Wang Z, Qian X et al (2017) Next-generation sequencing of 74 Y-SNPs to construct a concise consensus phylogeny tree for Chinese population. Forensic Sci Int Genet Suppl Ser 6:e96–e98. https://doi.org/10.1016/j.fsigss.2017.09.043

Gettings KB, Borsuk LA, Ballard D, Bodner M, Budowle B, Devesse L et al (2017) STRSeq: a catalog of sequence diversity at human identification Short Tandem Repeat loci. Forensic Sci Int Genet 31:111–117. https://doi.org/10.1016/J.FSIGEN.2017.08.017

Gorgé O, Bennett EA, Massilani D, Daligault J, Pruvost M, Geigl E-M, Grange T (2016) Analysis of ancient DNA in microbial ecology. Methods Mol Biol 1399:289–315. https://doi.org/10.1007/978-1-4939-3369-3_17

Guo J, Fu X, Liao H, Hu Z, Long L, Yan W et al (2016) Potential use of bacterial community succession for estimating post-mortem interval as revealed by high-throughput sequencing. Sci Rep 6(1):24197. https://doi.org/10.1038/srep24197

Gut IG (2013) New sequencing technologies. Clin Transl Oncol 15(11):879–881. https://doi.org/10.1007/s12094-013-1073-6

Hampton-Marcell JT, Lopez JV, Gilbert JA (2017) The human microbiome: an emerging tool in forensics. Microb Biotechnol 10(2):228–230. https://doi.org/10.1111/1751-7915.12699

Han J-P, Yang F, Xu C, Wei Y-L, Zhao X-C, Hu L et al (2014) A new strategy for sperm isolation and STR typing from multi-donor sperm mixtures. Forensic Sci Int Genet 13:239–246. https://doi.org/10.1016/j.fsigen.2014.08.012

Hashiyada M, Adachi N, Akane A, Matsumoto T, Yoshimura S (2017) The detailed examination of the human ancient mitochondrial DNA using the Ion PGM™ System. Forensic Sci Int Genet Suppl Ser 6:e156–e158. https://doi.org/10.1016/J.FSIGSS.2017.09.061

He Y, Wu J, Dressman DC, Iacobuzio-Donahue C, Markowitz SD, Velculescu VE et al (2010) Heteroplasmic mitochondrial DNA mutations in normal and tumour cells. Nature 464:610–614. https://doi.org/10.1038/nature08802

Hofreiter M, Paijmans JLA, Goodchild H, Speller CF, Barlow A, Fortes GG et al (2015) The future of ancient DNA: technical advances and conceptual shifts. BioEssays 37(3):284–293. https://doi.org/10.1002/bies.201400160

Holland MM, McQuillan MR, O'Hanlon KA (2011) Second generation sequencing allows for mtDNA mixture deconvolution and high resolution detection of heteroplasmy. Croat Med J 52 (3):299–313. https://doi.org/10.3325/cmj.2011.52.299

Holland MM, Pack ED, McElhoe JA (2017) Evaluation of GeneMarker®HTS for improved alignment of mtDNA MPS data, haplotype determination, and heteroplasmy assessment. Forensic Sci Int Genet 28:90–98. https://doi.org/10.1016/j.fsigen.2017.01.016

Home—10× Genomics (n.d.). https://www.10xgenomics.com/. Accessed 7 June 2018

Hyde ER, Haarmann DP, Lynne AM, Bucheli SR, Petrosino JF (2013) The living dead: bacterial community structure of a cadaver at the onset and end of the bloat stage of decomposition. PLoS One 8(10):e77733. https://doi.org/10.1371/journal.pone.0077733

Iacovacci G, D'Atanasio E, Marini O, Coppa A, Sellitto D, Trombetta B et al (2017) Forensic data and microvariant sequence characterization of 27 Y-STR loci analyzed in four Eastern African countries. Forensic Sci Int Genet 27:123–131. https://doi.org/10.1016/J.FSIGEN.2016.12.015

International Human Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome. Nature 431(7011):931–945. https://doi.org/10.1038/nature03001

Irwin JA, Parson W, Coble MD, Just RS (2011a) mtGenome reference population databases and the future of forensic mtDNA analysis. Forensic Sci Int Genet 5(3):222–225. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-79953867044&partnerID=40&md5=0c14688affdaf9198aad1f6d2a783bc6

Irwin J, Just R, Scheible M, Loreille O (2011b) Assessing the potential of next generation sequencing technologies for missing persons identification efforts. Forensic Sci Int Genet Suppl Ser 3(1):e447–e448

Ishiya K, Ueda S (2017) MitoSuite: a graphical tool for human mitochondrial genome profiling in massive parallel sequencing. PeerJ 5(5):e3406. https://doi.org/10.7717/peerj.3406

Jäger AC, Alvarez ML, Davis CP, Guzmán E, Han Y, Way L et al (2017) Developmental validation of the MiSeq FGx forensic genomics system for targeted next generation sequencing in forensic DNA casework and database laboratories. Forensic Sci Int Genet 28:52–70. https://doi.org/10.1016/j.fsigen.2017.01.011

Javan GT, Finley SJ, Smith T, Miller J, Wilkinson JE (2017) Cadaver thanatomicrobiome signatures: the ubiquitous nature of Clostridium species in human decomposition. Front Microbiol 8. https://doi.org/10.3389/fmicb.2017.02096

Jobling MA, Gill P (2004) Encoded evidence: DNA in forensic analysis. Nat Rev Genet 5 (10):739–751. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-5044225199&partnerID=40

Just RS, Moreno LI, Smerick JB, Irwin JA (2017) Performance and concordance of the ForenSeq™ system for autosomal and Y chromosome short tandem repeat sequencing of reference-type specimens. Forensic Sci Int Genet 28:1–9. https://doi.org/10.1016/j.fsigen.2017.01.001

Kader F, Ghai M (2015) DNA methylation and application in forensic sciences. Forensic Sci Int 249:255. https://doi.org/10.1016/j.forsciint.2015.01.037

Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF (2008) New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. Genome Res 18(5):830–838. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-43149087535&partnerID=40&md5=1b7dc85ee8aeea4dbbc309c07ac69a5d

Kayser M, de Knijff P (2011) Improving human forensics through advances in genetics, genomics and molecular biology. Nat Rev Genet 12(3):179–192

King JL, Sajantila A, Budowle B (2014) MitoSAVE: mitochondrial sequence analysis of variants in excel. Forensic Sci Int Genet 12:122–125. https://doi.org/10.1016/j.fsigen.2014.05.013

Kwon SY, Lee HY, Kim EH, Lee EY, Shin K-J (2016) Investigation into the sequence structure of 23 Y chromosomal STR loci using massively parallel sequencing. Forensic Sci Int Genet 25:132–141. https://doi.org/10.1016/J.FSIGEN.2016.08.010

Larmuseau MHD, Van Geystelen A, Kayser M, van Oven M, Decorte R (2015) Towards a consensus Y-chromosomal phylogeny and Y-SNP set in forensics in the next-generation sequencing era. Forensic Sci Int Genet 15:39–42. https://doi.org/10.1016/j.fsigen.2014.11.012

Larmuseau MHD, Otten GPPL, Decorte R, Van Damme P, Moisse M (2017) Defining Y-SNP variation among the Flemish population (Western Europe) by full genome sequencing. Forensic Sci Int Genet 31:e12–e16. https://doi.org/10.1016/j.fsigen.2017.10.008

Lee RC, Feinbaum RL, Ambros V (1993) The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75(5):843–854. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8252621

Li L, Ge J, Zhang S, Guo J, Zhao S, Li C et al (2012) Maternity exclusion with a very high autosomal STRs kinship index. Int J Legal Med 126(4):645–648. https://doi.org/10.1007/s00414-012-0668-8

Liu L, Li Y, Li S et al (2012) Comparison of next-generation sequencing systems. J Biomed Biotechnol 2012:251364. https://doi.org/10.1155/2012/251364

Li L, Wang Y, Yang S, Xia M, Yang Y, Wang J et al (2017) Genome-wide screening for highly discriminative SNPs for personal identification and their assessment in world populations. Forensic Sci Int Genet 28:118–127. https://doi.org/10.1016/j.fsigen.2017.02.005

Loomis EW, Eid JS, Peluso P, Yin J, Hickey L, Rank D et al (2013) Sequencing the unsequenceable: expanded CGG-repeat alleles of the fragile X gene. Genome Res 23 (1):121–128. https://doi.org/10.1101/gr.141705.112

Lv S, Ji L, Chen B, Liu S, Lei C, Liu X et al (2017) Histone methyltransferase KMT2D sustains prostate carcinogenesis and metastasis via epigenetically activating LIFR and KLF4. Oncogene 37:1–15. https://doi.org/10.1038/s41388-017-0026-x

Marshall C, Sturk-Andreaggi K, Daniels-Higginbotham J, Oliver RS, Barritt-Ross S, McMahon TP (2017) Performance evaluation of a mitogenome capture and Illumina sequencing protocol using non-probative, case-type skeletal samples: implications for the use of a positive control in a next-generation sequencing procedure. Forensic Sci Int Genet 31:198–206. https://doi.org/10.1016/J.FSIGEN.2017.09.001

Massaia A, Xue Y (2017) Human Y chromosome copy number variation in the next generation sequencing era and beyond. Hum Genet 136:591. https://doi.org/10.1007/s00439-017-1788-5

Mayes C, Seashols-Williams S, Hughes-Stamm S (2018) A capillary electrophoresis method for identifying forensically relevant body fluids using miRNAs. Legal Med 30:1. https://doi.org/10.1016/j.legalmed.2017.10.013

McNevin D, Wilson-Wilde L, Robertson J, Kyd J, Lennard C (2005) Short tandem repeat (STR) genotyping of keratinised hair: Part 2. An optimised genomic DNA extraction procedure reveals donor dependence of STR profiles. Forensic Sci Int 153(2–3):247–259. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-24044545493&partnerID=40

Mendez FL, Krahn T, Schrack B, Krahn AM, Veeramah KR, Woerner AE et al (2013) An African American paternal lineage adds an extremely ancient root to the human y chromosome phylogenetic tree. Am J Hum Genet 92(3):454–459. https://doi.org/10.1016/j.ajhg.2013.02.002

Metcalf JL, Xu ZZ, Bouslimani A, Dorrestein P, Carter DO, Knight R (2017) Microbiome tools for forensic science. Trends Biotechnol 35:814. https://doi.org/10.1016/j.tibtech.2017.03.006

Metzker ML (2009) Sequencing technologies—the next generation. Nat Rev Genet 11(1):31–46

Miller AD, Good RT, Coleman RA, Lancaster ML, Weeks AR (2013) Microsatellite loci and the complete mitochondrial DNA sequence characterized through next generation sequencing and de novo genome assembly for the critically endangered orange-bellied parrot, *Neophema chrysogaster*. Mol Biol Rep 40(1):35–42

Montano EA, Bush JM, Garver AM, Larijani MM, Wiechman SM, Baker CH et al (2018) Optimization of the Promega PowerSeq™ Auto/Y system for efficient integration within a forensic DNA laboratory. Forensic Sci Int Genet 32:26–32. https://doi.org/10.1016/J.FSIGEN.2017.10.002

Novroski NMM, King JL, Churchill JD, Seah LH, Budowle B (2016) Characterization of genetic sequence variation of 58 STR loci in four major population groups. Forensic Sci Int Genet 25:214–226. https://doi.org/10.1016/J.FSIGEN.2016.09.007

Ochiai E, Minaguchi K, Nambiar P, Kakimoto Y, Satoh F, Nakatome M et al (2016) Evaluation of Y chromosomal SNP haplogrouping in the HID-Ion AmpliSeq™ Identity Panel. Legal Med 22:58–61. https://doi.org/10.1016/J.LEGALMED.2016.08.001

Parson W, Strobl C, Huber G, Zimmermann B, Gomes SM, Souto L et al (2013) Evaluation of next generation mtGenome sequencing using the Ion Torrent Personal Genome Machine (PGM). Forensic Sci Int Genet 7(5):543–549

Parson W, Ballard D, Budowle B, Butler JM, Gettings KB, Gill P et al (2016) Massively parallel sequencing of forensic STRs: considerations of the DNA commission of the International Society for Forensic Genetics (ISFG) on minimal nomenclature requirements. Forensic Sci Int Genet 22:54–63. https://doi.org/10.1016/j.fsigen.2016.01.009

Pechal JL, Crippen TL, Benbow ME, Tarone AM, Dowd S, Tomberlin JK (2014) The potential use of bacterial community succession in forensics as described by high throughput metagenomic sequencing. Int J Legal Med 128(1):193–205. https://doi.org/10.1007/s00414-013-0872-1

Peng Y, Croce CM (2016) The role of MicroRNAs in human cancer. Signal Transduct Target Ther 1(1):15004. https://doi.org/10.1038/sigtrans.2015.4

Products (n.d.). https://nanoporetech.com/products. Accessed 7 June 2018

Products | Single Cell Analysis System | Cyto-Mine® (n.d.). https://www.spherefluidics.com/products/integrated-systems/cyto-mine/. Accessed 7 June 2018

Rathbun MM, McElhoe JA, Parson W, Holland MM (2017) Considering DNA damage when interpreting mtDNA heteroplasmy in deep sequencing data. Forensic Sci Int Genet 26:1–11. https://doi.org/10.1016/j.fsigen.2016.09.008

Rockenbauer E, Hansen S, Mikkelsen M, Børsting C, Morling N (2014) Characterization of mutations and sequence variants in the D21S11 locus by next generation sequencing. Forensic Sci Int Genet 8(1):68–72

Sanger F et al (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci 74:5463–5467

Sanchez-Cespedes M, Cairns P, Jen J, Sidransky D (1998) Degenerate oligonucleotide-primed PCR (DOP-PCR): evaluation of its reliability for screening of genetic alterations in neoplasia. BioTechniques 25(6):1036–1038. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9863058

Sauer E, Extra A, Cachée P, Courts C (2017) Identification of organ tissue types and skin from forensic samples by microRNA expression analysis. Forensic Sci Int Genet 28:99–110. https://doi.org/10.1016/j.fsigen.2017.02.002

Schmedes SE, Woerner AE, Novroski NMM, Wendt FR, King JL, Stephens KM, Budowle B (2018) Targeted sequencing of clade-specific markers from skin microbiomes for forensic human identification. Forensic Sci Int Genet 32:50–61. https://doi.org/10.1016/j.fsigen.2017.10.004

Seo SB, King JL, Warshauer DH, Davis CP, Ge J, Budowle B (2013) Single nucleotide polymorphism typing with massively parallel sequencing for human identification. Int J Legal Med 127 (6):1079–1086. https://doi.org/10.1007/s00414-013-0879-7

Sequel System—PacBio (n.d.). https://www.pacb.com/products-and-services/pacbio-systems/sequel. Accessed 7 June 2018

Sequencing Platforms | Compare NGS Platforms (Benchtop, Production-Scale) (n.d.). https://emea.illumina.com/systems/sequencing-platforms.html. Accessed 7 June 2018

Silva DSBS, Sawitzki FR, Scheible MKR, Bailey SF, Alho CS, Faith SA (2018) Genetic analysis of Southern Brazil subjects using the PowerSeq™ AUTO/Y system for short tandem repeat sequencing. Forensic Sci Int Genet 33:129–135. https://doi.org/10.1016/j.fsigen.2017.12.008

Sirker M, Fimmers R, Schneider PM, Gomes I (2017) Evaluating the forensic application of 19 target microRNAs as biomarkers in body fluid and tissue identification. Forensic Sci Int Genet 27:41–49. https://doi.org/10.1016/j.fsigen.2016.11.012

Sturk-Andreaggi K, Peck MA, Boysen C, Dekker P, McMahon TP, Marshall CK (2017) AQME: a forensic mitochondrial DNA analysis tool for next-generation sequencing data. Forensic Sci Int Genet 31:189–197. https://doi.org/10.1016/j.fsigen.2017.09.010

Tridico SR, Murray DC, Addison J, Kirkbride KP, Bunce M (2014) Metagenomic analyses of bacteria on human hairs: a qualitative assessment for applications in forensic science. Investig Genet 5(1):16. https://doi.org/10.1186/s13323-014-0016-5

Tsai CW, Yang CH, Chou SL, Cheng SG, Pai CY (2013) Non-CODIS DNA markers could be more effective than CODIS-based STRs in problematic biological relationship cases. Rom J Leg Med 21(3):245–248. https://doi.org/10.4323/rjlm.2013.245

Ursell LK, Metcalf JL, Parfrey LW, Knight R (2012) Defining the human microbiome. Nutr Rev 70 (Suppl 1):S38. https://doi.org/10.1111/j.1753-4887.2012.00493.x

Van Geystelen A, Decorte R, Larmuseau MHD (2013) AMY-tree: an algorithm to use whole genome SNP calling for Y chromosomal phylogenetic applications. BMC Genomics 14(1):101. https://doi.org/10.1186/1471-2164-14-101

Van Neste C, Van Nieuwerburgh F, Van Hoofstat D, Deforce D (2012) Forensic STR analysis using massive parallel sequencing. Forensic Sci Int Genet 6(6):810–818

Vidaki A, Kayser M (2017) From forensic epigenetics to forensic epigenomics: broadening DNA investigative intelligence. Genome Biol 18(1):238. https://doi.org/10.1186/s13059-017-1373-1

Vidaki A, Díez López C, Carnero-Montoro E, Ralf A, Ward K, Spector T et al (2017) Epigenetic discrimination of identical twins from blood under the forensic scenario. Forensic Sci Int Genet 31:67–80. https://doi.org/10.1016/j.fsigen.2017.07.014

Wang Z, Zhang S, Bian Y, Li C (2015) Differentiating between monozygotic twins in forensics through next generation mtGenome sequencing. Forensic Sci Int Genet Suppl Ser 5:e58–e59. https://doi.org/10.1016/J.FSIGSS.2015.09.023

Wang Z, Zhou D, Wang H, Jia Z, Liu J, Qian X et al (2017) Massively parallel sequencing of 32 forensic markers using the Precision ID GlobalFiler™ NGS STR Panel and the Ion PGM™ System. Forensic Sci Int Genet 31:126–134. https://doi.org/10.1016/J.FSIGEN.2017.09.004

Warshauer DH, Davis C, Holt C, Han Y, Walichiewicz P, Richardson T et al (2015a) Massively parallel sequencing of forensically relevant single nucleotide polymorphisms using TruSeq™ forensic amplicon. Int J Legal Med 129(1):31–36. https://doi.org/10.1007/s00414-014-1108-8

Warshauer DH, Churchill JD, Novroski N, King JL, Budowle B (2015b) Novel Y-chromosome short tandem repeat variants detected through the use of massively parallel sequencing. Genomics Proteomics Bioinformatics 13(4):250–257. https://doi.org/10.1016/J.GPB.2015.08.001

Williams GA (2012) Identification and resolution of capability gaps in forensic science. Doctoral thesis, University of Huddersfield. (Deposit: http://eprints.hud.ac.uk/id/eprint/17500/)

Willems T, Gymrek M, Poznik GD, Tyler-Smith C, Erlich Y (2016) Population-scale sequencing data enable precise estimates of Y-STR mutation rates. Am J Hum Genet 98(5):919–933. https://doi.org/10.1016/j.ajhg.2016.04.001

Wong L-P, Ong RT-H, Poh W-T, Liu X, Chen P, Li R et al (2013) Deep whole-genome sequencing of 100 Southeast Asian Malays. Am J Hum Genet 92(1):52–66. https://doi.org/10.1016/j.ajhg.2012.12.005

Xavier C, Parson W (2017) Evaluation of the Illumina ForenSeq™ DNA Signature Prep Kit—MPS forensic application for the MiSeq FGx™ benchtop sequencer. Forensic Sci Int Genet 28:188–194. https://doi.org/10.1016/j.fsigen.2017.02.018

Yang Y, Xie B, Yan J (2014) Application of next-generation sequencing technology in forensic science. Genomics Proteomics Bioinformatics 12(5):190–197. https://doi.org/10.1016/j.gpb.2014.09.001

Zaragoza MV, Fass J, Diegoli M, Lin D, Arbustini E (2010) Mitochondrial DNA variant discovery and evaluation in human cardiomyopathies through next generation sequencing. PLoS One 5(8). https://doi.org/10.1371/journal.pone.0012295

Zeng X, King JL, Stoljarova M, Warshauer DH, LaRue BL, Sajantila A et al (2015) High sensitivity multiplex short tandem repeat loci analyses with massively parallel sequencing. Forensic Sci Int Genet 16:38–47. https://doi.org/10.1016/j.fsigen.2014.11.022

Zhou Y, Guo F, Yu J, Liu F, Zhao J, Shen H et al (2016) Strategies for complete mitochondrial genome sequencing on Ion Torrent PGM™ platform in forensic sciences. Forensic Sci Int Genet 22:11–21. https://doi.org/10.1016/J.FSIGEN.2016.01.004

24 Utility and Possibility of Next-Generation Sequencing in Forensic DNA Typing

Vijay Nema

Abstract

The introduction of DNA typing itself was a game changer for forensic science as it opened the window for knowing what was thought to be nonexisting. The traces that were of no use suddenly became crucial evidence and were able to precisely identify the individuals and their genetic information too. In DNA typing, the biggest limitation came in for bigger DNA fragments or mixed samples, etc., which were tough to deal with. The solutions to this kind of problems along with the introduction of high-throughput sequencing were introduced by next-generation sequencing. With the improvement in technology not only the throughput has increased but there has been a tremendous decrease in cost of assays. This has increased precision along with affordability. This chapter has tried to educate the users about the basics of next-generation sequencing and has taken references from Sanger sequencing to make the reader understand better and in context with what is available and what are the future avenues as shown by new research evidence. A next-generation sequencing is a tool which can be employed for various applications a glimpse of which has been provided in this chapter; however, they are not and cannot be exhaustive as there is a potential for newer and novel applications based on innovative approaches and need of the field.

Keywords

Next-generation sequencing (NGS) · Forensics · DNA typing · Sequencing platforms · Capillary electrophoresis (CE)

V. Nema (✉)
ICMR-National AIDS Research Institute, Pune, India

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_24

24.1 Introduction

It is a fact that every criminal leave behind some clues. However, sometimes the clues are so vague that finding them out with general observation or techniques becomes impossible. Moreover, the specificity of the available evidence is always a big concern so as to avoid punishment to an innocent person. Much work went into an optimization to utilize the miniscule of the samples or evidence from the spot and to improve upon sensitivity and specificity of detection. Biological samples are the best to correlate specifically to individuals. Among biological samples, blood, semen, skin, saliva, hairs, etc. are mostly available, and DNA from this sample is used to specifically identify the individual from his biological leftover. The procedure is basically a linking technique wherein biological samples of the individuals are used to extract DNA which is then analyzed. When processed, the DNA of each individual generates a pattern called "DNA profile." This DNA profile remains unique for each individual except in identical twins. This technique of DNA profiling is commonly used in criminal investigations in different ways. Utilizing this, an unidentified person may be identified, or in case of suspicion, the identity can be confirmed. This comes as a handy solution when confirmation about a person's presence on a crime scene is a concern. The science of DNA typing is also an evolving science wherein techniques of typing have improved a lot, and sensitivity and specificity have gone up by many folds. In this chapter, a few such techniques are introduced, and their details are discussed.

24.2 The History of DNA as Evidence in Forensic Profiling

The first documented use of DNA fingerprinting in forensic science was recorded in 1986 when police in the United Kingdom requested Dr. Jeffreys, of University of Leicester, for cross checking a person's acceptance that he was responsible for two rape-murders which were proven wrong. The first case of confirmation about the person to be convicted utilizing this technique was Robert Melias in 1987 in the same country (Anon 1996). The same year witnessed another case in the USA, wherein the suspect in a rape case was convicted based on DNA evidence. However, there were a few cases wherein the use of technology suffered lacunae and different directives and guidelines were made regarding the right procedure of testing and keeping of the results as obtained in the laboratory. The guidelines also explained how the reports should be explained and how the errors were to be recorded. Compliance with appropriate standards and controls was made mandatory through regulatory agencies. After following all these, the technology was found to be full proof and was being used as a routine technique. Based on these facts, the US National Research Council's (NRC) 1996 report (Anon 1996) on DNA evidence stated, "The state of the profiling technology and the methods for estimating frequencies and related statistics have progressed to the point where the admissibility of properly collected and analyzed DNA data should not be in doubt" (Kaye 1995).

24.3 The Source of DNA

The human body is made up of basic units called cells. Every cell except red blood cells is having a defined structure with various components including nucleus that functions as its command center. This command center hosts the DNA or deoxyribonucleic acid that codes for genetic information responsible for all cellular processes. The DNA being a stable molecule with specific traits when compared between individuals is the best biomolecule for forensic purposes. The nuclear DNA is limited to one set (two copies) per cell, whereas the DNA found in mitochondrion is available in multiple copies ranging up to 100,000 copies. Hence, both of these DNAs are used for various purposes depending on the availability and quantity of the samples recovered from crime scenes.

24.4 The Source of Variation

Human DNA is a large molecule with different regions having coding and non-coding sequences of nucleotides. Coding sequences are responsible for cellular processes as they code for proteins, enzymes, and various structural proteins. However, noncoding regions that were earlier named as junk DNA are very important with respect to regulation of cellular processes. These processes comprise of the transcriptional and translational regulation of protein-coding sequences, regions for the attachment of scaffold, a place for initiation of DNA replication, centromeres, and telomeres along with some sequences which are repetitive. These functions are very important for cellular and biological maintenance while repetitive sequences have found their applications in the identification of individuals. Majority of the noncoding DNA comprises repetitive sequences. The number of repeats and location of these repeat sequences form the basis for their classification. At the same time, the repeat could be of a single nucleotide to a full gene or a several thousand base pairs. Repeats may be found scattered all through the genome or may be localized at a particular site like telomeric regions. As they were initially found as small satellite bands in density-gradient centrifugations, they are also referred as satellite DNA. Mini- or microsatellites in these noncoding regions are the ones which are highly polymorphic. The ones that are localized in a given human chromosome are single-locus satellites, and those that are scattered throughout the genome are called as multi-locus satellite elements or short tandem repeats (STRs). Repetitive DNA has been classified based on characteristics which are not very stringent. However, for convenience, they are divided into five classes: microsatellites and minisatellites, satellites, retroposons, and proretroviral transposons. Further details show that these repeats indicate a pattern of their repetition. Some are direct while others are indirect repeats; a few are a complement, reverse complement, or palindrome in their occurrences (Table 24.1).

Further, recent studies have tried to understand how much sequence variation exists in these STR loci. A study found that three STRs, in particular, D21S11, D2S1338, and D12S391, showed higher allelic diversity from sequencing the repeat

Table 24.1 Type of DNA repeat sequences

Type of repeat	Pattern	Example
Direct	A repeat also called as forward repeat is the recurrence of a pattern on the same strand in the same nucleotide order	ACTG recurs as ACTG
Indirect	A repeat also called as reverse or inverse repeat recurs on the same strand but the order of the nucleotides is reverse	ACTG recurs as GTCA
Complement	A repeat where the nucleotides are complemented according to Watson Crick pairing	ACTG recurs as TGAC
Reverse complement	A repeat recurs on the same strand but the nucleotides are complemented and the order of the nucleotides is reversed	ACTG recurs as CAGT
Palindrome	A combination of two consecutive occurrences in opposite orientations and read the same when reading from left to right or vice-versa	GGATCC repeats as CCTAGG

motifs in all four populations studied (African American, Caucasian, Hispanic, and Chinese). Several loci, CSF1PO, D10S1248, TH01, and TPOX, showed little to no diversity in length vs. sequence-based alleles (Novroski et al. 2016). Additionally, single-nucleotide polymorphisms (SNPs) in the flanking regions of these STR repeat motif are being studied to estimate their utility in resolving complex mixture and kinship analyses. D5S818, D7S820, and D13S317 have multiple polymorphisms within haplotypes, and this has resulted in new alleles, based on variation in these flanking regions, and they are associated with multiple populations and allele sizes. D16S539 shows an SNP which is also well represented in the most common allele sizes and populations, making it useful for new allele determinations (Gettings et al. 2015a).

24.4.1 How Are the Repeat Sequences Used for Identification and What Are the Techniques?

This variability of DNA is often the most useful tool for differentiation and identification of an individual. There are various methods available for identifying individuals in forensic science. STR analysis is one of the widely used ones. These STRs are the repeats equal to or less than six bases and are also named microsatellites. Conventional polymerase chain reaction (PCR) is being used for STR typing wherein the amplicons obtained after the PCR are electrophoresed and visualized for differences or similarities as compared to others or controls. There are manual as well as kit-based automatic methods being used for STR analysis in different setups. For STR characterization and analysis of size, about 10–13 STR markers are in use as standards. All 13 STR regions are analyzed and compared to establish profiles in criminal investigations with a chance of two people having the exact same 13 regions reaching one in a billion. However, advanced methods are employed for the exploration of STR sequence diversity wherein multiple chromosomal loci are analyzed. Commercial kits with up to 20 core loci are currently

Table 24.2 Other techniques available for genetic analysis in forensic science

Name of the analysis	Technique used	Purpose
STR analysis	PCR	To measure the number of repeats at specific loci
Fluorescent STR analysis	PCR with fluorescently labelled primers	Detection of different STR loci
STR analysis for three loci to resolve parentage dispute	PCR	Paternity testing
Single-Nucleotide Polymorphism (SNP)	PCR-RFLP/PCR-Sequencing	Information on physical characteristics, e.g., eye, hair, or skin color
Point mutation detection	PCR with probe	Detection and quantification of the age-related point mutation A189G in the human mitochondrial DNA Detection of rare diseases, etc.
Quantitative PCR	Real-time PCR (with variants like probe-based RT-PCR, Taqman Assays or Molecular Beacons)	Most important as a quantitation tool for DNA in available sample stain identification (blood/semen, etc.)
Microarray-based forensic DNA analysis	Microarray	SNP profiling (even in degraded samples where STR analysis is not feasible)
Fluorescence in situ hybridization (FISH)	RNA-FISH	Identification of body fluids

available (Hares 2015). Table 24.2 highlights a few more techniques that are being used for forensic analysis. These are mainly dependent on the instrumentation available and the type of information required.

24.4.1.1 Use of Capillary Electrophoresis (CE)

Capillary electrophoresis came as a much easier way for analyzing the product of PCR. It was much convenient as compared to the bulky slab gels and increased the speed and automation of DNA-typing procedures. All the three conventional steps, viz. loading, open electrophoresis, and UV visualization and documentation were replaced by injection, separation, and detection, respectively, and have become fully automated. In addition, quantitative information became available in a single step in conjunction with software analyzing the data. The best gain here was the reduction in sample volume required for the accurate analysis. Currently, allele classifications are based on PCR-based fragment sizing by CE following a simple nomenclature convention (Bodner et al. 2016). The alleles are called relative to known-sized allelic ladders, with a number value indicative of the number of complete repeat motifs and any additional nucleotides. A glimpse of few currently available sequence instruments is given in Fig. 24.1.

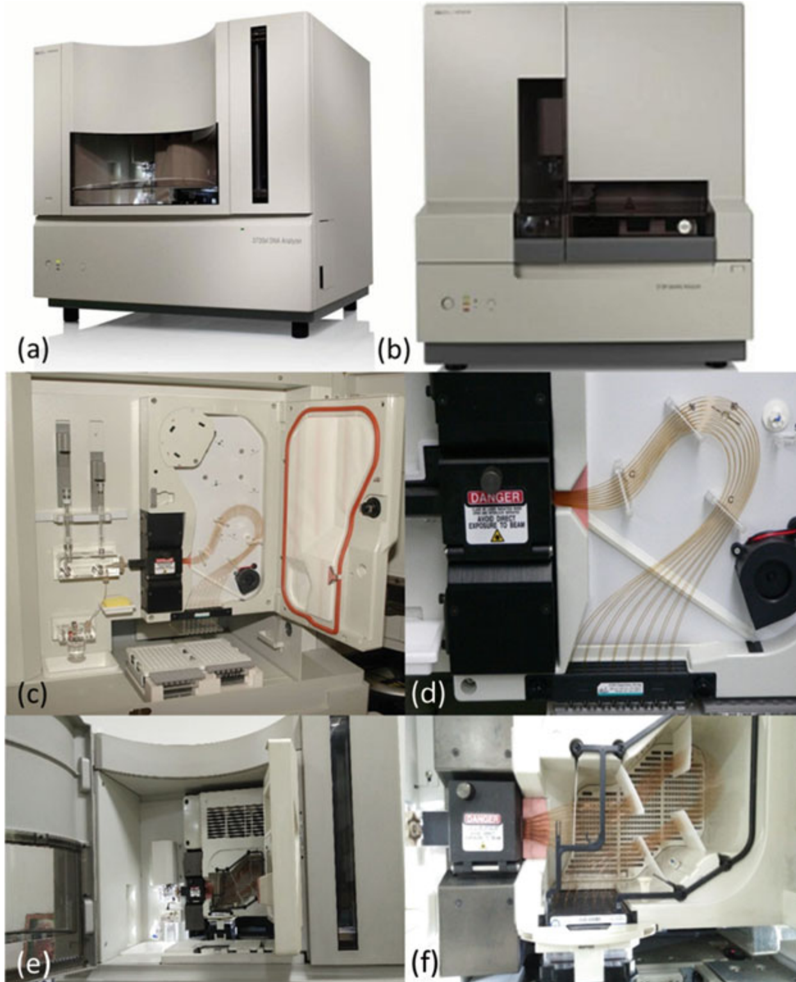


Fig. 24.1 Capillary electrophoresis instruments (Thermo) which can be used for capillary electrophoresis, (**a**, **b**) From outside, (**c**, **d**) inside view of 16 capillary instrument, (**e**, **f**) inside view of 96 capillary instrument

The advantages of capillary electrophoresis for fragment separation are enlisted below:

Multiplexing: Using fluorescence multiplexing different peaks with similar sizes can be detected in a single run by adding different fluorescent dyes to different samples (Fig. 24.2).

Improved sensitivity: With improved chemistry and automation in dispensing, the sensitivity achieved is far beyond the manual methods.

Wide range of sample types: All kind of samples including blood spots or paraffin-embedded samples could be used here.

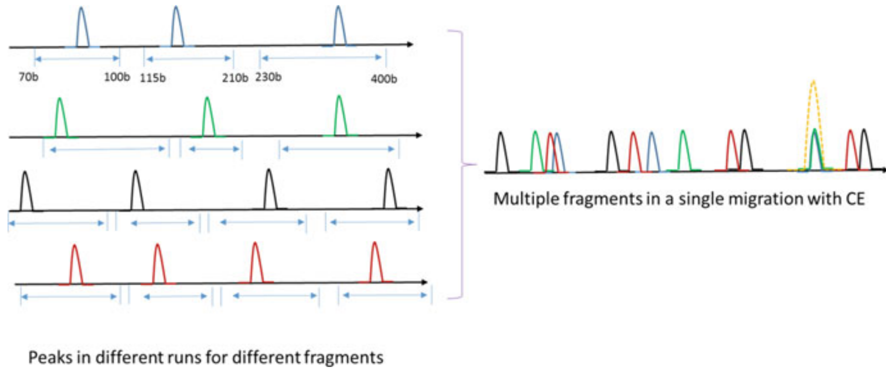


Fig. 24.2 A thematic representation of multiplexing

Simple preparation: Most fragment analysis applications do not need complex sample preparation or clean-up.

Simple data analysis: Determination of peaks remain very simple as it is just based on the sizes.

Large range of fragment analysis applications: Various applications can be run on CE as no prior knowledge of sequences is required.

This also has the scope for allowing relative quantification of peaks, and huge data can be generated once the standard curve is plotted for peak size and fragment length. Applications like paternity testing, animal typing, forensic, disease linkage, microsatellite instability, cancer genotyping, SNP analysis, etc. could be performed using fragment analysis on CE. The technology has become so easy and reliable that most of the forensic science laboratories are switching from biochemical methods to sequencing ones and hence raising the demand of DNA analysis setups with higher throughput capabilities.

24.5 Next-Generation Sequencing: Added Advantages over Sanger Sequencing

DNA analysis came as a precise and rapid method for obtaining test analysis results in forensic science. This has become possible with the analysis of short tandem repeat (STR) markers using PCR and capillary electrophoresis (CE). Analysis of specific regions of mitochondrial DNA (mtDNA) has also become possible with CE-based sequencing. In the current scenario, the load for analysis is growing heavily, but CE-based testing has limited capacity. Hence, a revised interest in developing and utilizing a better analysis method has come up with the demands for the identification of missing persons, kinship testing, ancestry investigations, and other human identification applications. Hence, there was a need for sequencing method which could process samples in parallel and should be free from those artifacts which Sanger sequencing or fragment analysis-based methodology had.

This was well undertaken by a new technology called next-generation sequencing technologies. Using this technology, a huge number of DNA strands can be parallelly sequenced to get higher throughput and reducing the requirement for the fragment-cloning methods that are normally used in Sanger sequencing of genomes (Nature 2020). In forensic science, the concept behind NGS technology and its use is based on CE sequencing only. The DNA polymerase catalyzes the incorporation of deoxyribonucleotide triphosphates (dNTPs) (fluorescently labeled) into a DNA template strand during PCR. In each cycle, the florescence is read when a particular nucleotide is incorporated. The fluorescent dyes normally used needs to have a well-resolved fluorescence emission spectra. A combination of fluorescein amidites-FAM (blue), 4′,5′-dichloro-2′,7′-dimethoxy-6-carboxyfluorescein-JOE (green), 5-carboxytetramethylrhodamine-TAMRA (yellow), and 6-carboxy-X-rhodamine-ROX (red) fits well under this criteria. However, the choice depends on the sequence used and instrument available.

The Sanger method has limitations of utilizing separate steps as discussed earlier. Added to that and compared with NGS, it has lower resolution also. NGS has miraculously solved this problem wherein it is capable of sequencing millions of fragments in parallel fashion. Hence it equals millions of runs from PCR/Sanger sequencing where one template is sequenced in one well with one run.

Figure 24.3 explains the basics of Sanger sequencing which became the basic input for the design of next-generation sequencing with little modification in sequencing chemistry and multiplexing the reactions to get a massive parallel sequencing technology. The PCR mix contains fluorescently labeled chain-terminating ddNTPs generating products of various lengths which are separated according to their size (Fig. 24.3). A gel picture is used in the figure for better understanding of the technology. Same reaction of chain termination takes place in capillary electrophoresis and is run through capillaries having finer path and automated run through a gel like matrix.

The next-generation sequencing (NGS) added a new aspect to the field of forensic human identification by substantiating benefits over CE analysis. It could provide an actual sequence of the repeat motifs in place of just the size of it. It provided an opportunity to multiplex even more markers and allowed the analysis of degraded samples which was not possible with CE analysis, especially for the larger loci (Borsting and Morling 2015; Parson et al. 2015; Eduardoff et al. 2015). Studies have proved that using NGS, STR scan can be typed effectively with the generation of profiles that goes along well with CE analyses, even from challenging crime scene samples (Scheible et al. 2014). The STR profile generated with NGS provides more information as compared to the allele number that is provided by CE analysis. Additionally, it analyzes the nucleotide sequence of the repeat motifs as well as nearby variations in the flanking regions. Evaluations of the results obtained using NGS-based STR typing revealed that this technology may be useful in routine casework as it may improve mixture resolution and increase the PoD by including both more loci, and the extrasequence information (Borsting and Morling 2015; Gettings et al. 2015b).

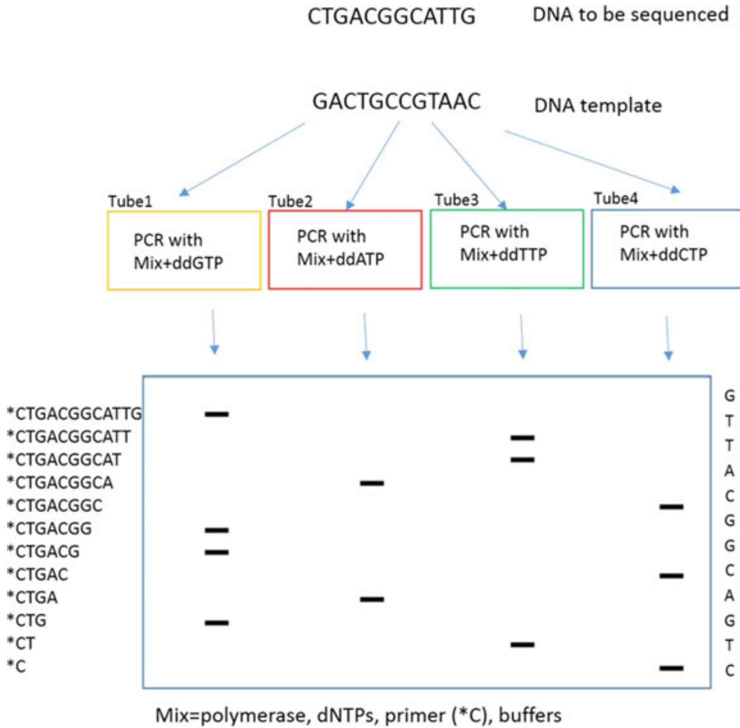


Fig. 24.3 Fundamentals of Sanger sequencing

24.5.1 Principle of Next-Generation Sequencing

As stated earlier, the next-generation sequencing (NGS) is also based on the same principle as like that of Sanger sequencing. The genomic strand serving as a template is broken into pieces, and the nucleotides in each fragment are recognized by the signals generated when the fragments are extended against a template strand (Fig. 24.4). There are arrays which are used in NGS wherein each miniscule well serves as one reaction center having Sanger sequencing ongoing in parallel to millions of other reactions going in other wells of the same array. This increases the throughput as well as reduces the cost to a significant scale. Various NGS systems are available these days with varying throughput and run time. Still, the basic principle in all sequencing technologies is to scan the successive addition of nucleotides to a fixed DAN array. The chemistries used in different technologies are different in generating the templates and the way they scan the sequence extension. Hence, NGS is a technique that uses positional separation rather than using the size separation to arrange the fluorescent molecules. To explain it further, fluorescent technology can be taken as one of the examples. Herein, millions of different template DNA strands sit tightly on segregated and fixed positions on a glass slide or throughout the entire sequencing reaction. With the incorporation of each

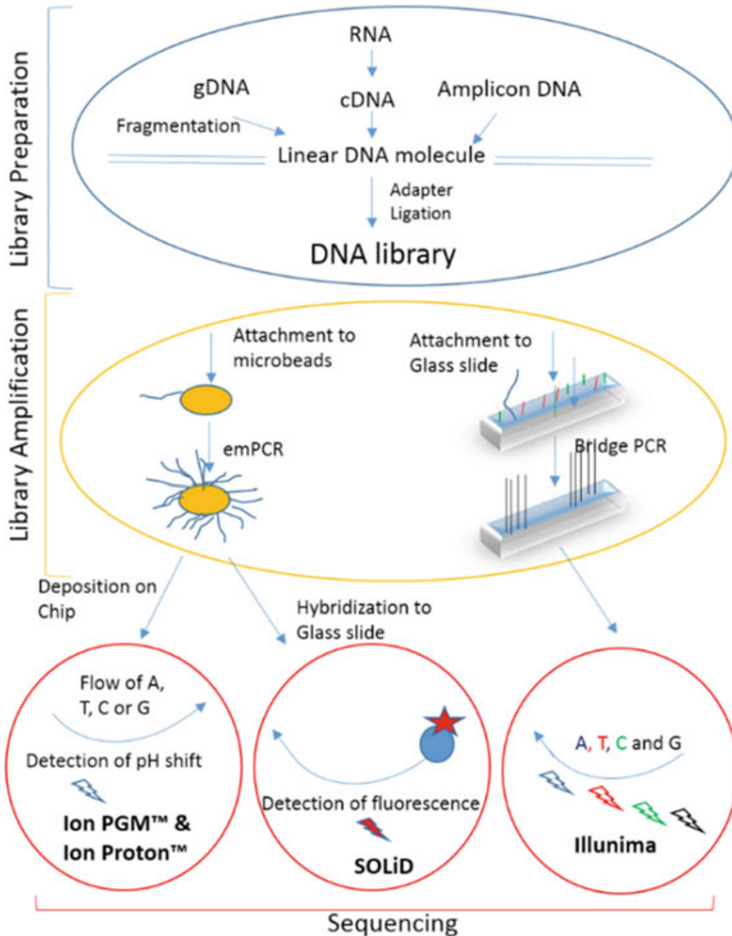


Fig. 24.4 Principle of NGS: generalized using fluorescence detection

modified base, a microscopic scanner records the position of each template on the glass as well as its fluorescent color and intensity. The modified bases are then cleaved to be the regular bases and are used to extend the reaction. After *n* number of scanning cycles, the software assigns all signals to their respective base (i.e., A, T, C, or G) (Muzzey et al. 2015). This is a nontechnical description of the whole process; however, knowing and understanding the technical terms help in imbibing the principle, instruments, and applications.

Read: It is the length of DNA that is actually read by the sequencers. It is to be noted here that the original amplicon or the fragment supplied to the sequencer would be larger than this segment which is read by the sequencer. For example, a particular sequencer and chemistry may read only first 300 bases while the supplied fragment could be in the range of 200–1000 bases. This reading may be

unidirectional (single read, SR) or may be read in a pair reading from both the ends simultaneously (paired end (PE) read).

Read Length and Mean Read Length: The number of bases sequenced are called "read length" which vary from reaction to reaction, and hence, an average of the total length is termed as "mean read length."

Sequencing Depth: The number of reads which align to the reference template in a run is called coverage or the depth of the sequencing. Hence, "It is the average number of times a single base is read during a sequencing run." For example, if a nucleotide or a gene fragment is sequenced 1000 times, the coverage would be called as 1000×. With the increase in coverage, reliability also increases as the errors become minority.

File Formats: Every instrument uses its own software to scan the reads and express them according to the fed algorithm. Hence, the output from the sequencer depends on its software. However, all the output files need to be converted into once common file format that may be read by other analysis software too. Some file formats as originated from various sequencers are given below: "SRF, FASTA, FASTQ, SFF, SCARF, AB1, EMBL" etc.

Q Score: This score indicates the possibility of a given base to be read incorrectly by the sequencer or the analysis software, hence directly related to the quality of the read. The score depends on various processing steps beginning from library preparation to the run and read.

24.5.2 Next-Generation Sequencing: Instrumentation

With the success of Sanger sequencing, different innovative technologies were tried and optimized for keeping the principle same while taking the number of reactions to a million fold throughput and reducing the cost per reaction. Hence, a PCR tube with single reaction has reached to millions of reactions in micro- or nano-sized wells of a chip which itself measures a few millimeters. This was the next-generation sequencing. The best-used technologies in NGS instruments, sequencing-by-synthesis (Ronaghi et al. 1998), sequencing-by-ligation (Shendure et al. 2005), and ion semiconductor sequencing (Rothberg et al. 2011) are the prominent oneswherein sequencing-by-synthesis are mostly used.

The instruments available in the market utilizes minor differences in the technologies to handle the given application. Some instruments and technologies are designed to execute lighter tasks with lowered cost which the others are robust in throughput and are utilized for heavier sequencing like human genome project. They require a heavy cost for their setup and execution. However, the information generated is also significant. Here we discuss about a few commercial instruments, but the list is not comprehensive to cover all the instruments and technologies.

Illumina: A company that is responsible for providing advanced instruments for sequencing and microarray. Providing technologies for the analysis of genetic variation and function is the key area along with all other dimensions of the genomic research and molecular diagnostics. A wide variety of instruments are available for

Fig. 24.5 A NovaSeq™ 6000 instrument (Courtesy of Illumina, Inc.)

NGS from this company. The instruments of NGS are based on the "sequencing by synthesis" principle. The instruments from Illumina ranges from "NovaSeq™, NextSeq®, HiSeq®, HiSeq X™, HiScanSQ™, MiSeq®, MiSeqDx™, MiniSeq™, etc." The MiSeq® and HiSeq® instruments are mostly used. Many of these instruments are can read in single (single reads) as well as in both the directions (paired-end reads), and the read length varies with instrument. NovaSeq™ 6000 (Fig. 24.5 as reproduce with the permission of Illumina Company) is the latest and the most affordable instrument with all the desired applications on board wherein NextSeq segment is the advanced series for very high load and coverage of all the latest applications.

Ion (Thermo Fisher): The sequencers from ion range use optical signals in a few instruments, whereas the release of H^+ ions is captured in others. Ion torrent that make use of optical signals is useful for amplicons, small genomes, or targeting of small regions within a genome. Ion proton with the capacity of larger output captures the release of hydrogen ion which is generated as a result of addition of a dNTP to a DNA polymer. Non-dependency on optics make this technology a bit faster. Ion GeneStudio series of the systems are the advanced ones which are used these days for various applications.

SOLiD (Sequencing by Oligonucleotide Ligation and Detection) by Life Technologies does the sequencing by ligation that takes place after emulsion PCR. In this, the DNA templates are amplified on beads. The system on which the reaction takes place is called the flow cell, and the beads are placed on the solid phase of this

flow cell. The system has a very high accuracy but has limitation of shorter read length.

24.5.2.1 Third-Generation Sequencing

Third-generation technologies of sequencing do not need PCR for amplification of template and also do not break down the DNA: the beauty is that they directly sequence a single DNA molecule and hence avoid errors which are normally introduced by PCR. These recent developments in the field of high-throughput sequencing have given the advanced instrumentation of execution of high-throughput sequencing workflows. The outcome advantages include increment in read length and less load of pre-sequencing preparation (Braslavsky et al. 2003).

Single-Molecule, Real-Time (SMRT) Sequencing (Pacific Biosciences): SMRT utilizes single molecule as template for sequencing and is capable of giving reads of a few kilobase length. With this assembling, complete genomes and sequencing full-length transcripts became feasible and cost-effective. A DNA polymerase which is anchored to the bottom of a Zero Mode Waveguide nanostructure (ZMW) that is a tiny well on the chip is used for "sequencing-by-synthesis." This synthesis is imaged in real time using fluorescence capture from nucleotides (Munroe and Harris 2010; Levene et al. 2003).

FRET Sequencing Platform: This technology is still under evolution and makes use of the principles of fluorescence resonance energy transfer. A DNA polymerase is tagged with a quantum dot that transfers its energy to a new base that is incorporated in the extending strand of DNA against an immobilized DNA template. During incorporation, a light emission phenomenon takes place. This is recorded and translated into a pattern of incorporation reproducing the sequence from the template.

The Oxford Nanopore Sequencing: As the name suggest, nanopore sequencing makes use of nanoscale holes. In place of using labeling, this technology measures the ionic current through the nanopores and measures the change in current whenever a molecule is passed through these nano pores. These nano pores can be created by utilizing biological nanopores as like the holes in membrane created by alpha-hemolysin from bacteria or may be the synthesized material which is a solid state nanopore. "MinION GridION and PromethION" are the names of the models available for this sequencing and are used as per the need and throughput of the sequencing. Further development for creating and working on solid-state nanopore is ongoing as it makes use of synthesized materials like graphene.

24.6 Next-Generation Sequencing: Procedure and Precautions

The conventional next-generation sequencing entails three basic steps. The reagents and technologies differ from these steps in different companies and platforms. Step one gets the sample ready by purifying it and removing all other impurities, and step two links the primers with the small oligonucleotides which are used as identifying bases in the later stages of data analysis. After attachment of the linkers, the fragments are identifiable, and this step is called library preparation. The libraries

are enriched by polymerase chain reaction, and the step is called amplification. The last step of reading or scanning these amplified molecules is to decipher the sequence of nucleotides for which different instruments and technologies use different chemistries as discussed earlier.

24.6.1 Library Preparation

Preparation of library is a step that depends on the purpose of sequencing. For example, if the sequencing is to be done for whole genome, the genome needs to be fragmented using enzyme digestion or by physical means like sonication to create smaller segments which can be ligated with adapters and read with the instrument used. In cases where only a particular segment of DNA is to be sequenced, the segments are enriched from the template DNA using PCR. Adapter-linked primers are used to do so (Fig. 24.6).

24.6.2 Amplification

In most of the cases, the library prepared for sequencing needs to be amplified so as to get signals much above the background, and the threshold limits during the detection by sequencer. The amplification is done using PCR which also have a few limitations like "biasing" and "duplication." Another problem is merging of the two segments to become one segment having a region originating from a little part from both of the two strands. This is called as Chimera, which mostly comes during metagenomic sequencing and is a big concern. Although software takes care of these chimeras in analysis part, these can be avoided during PCR only by using

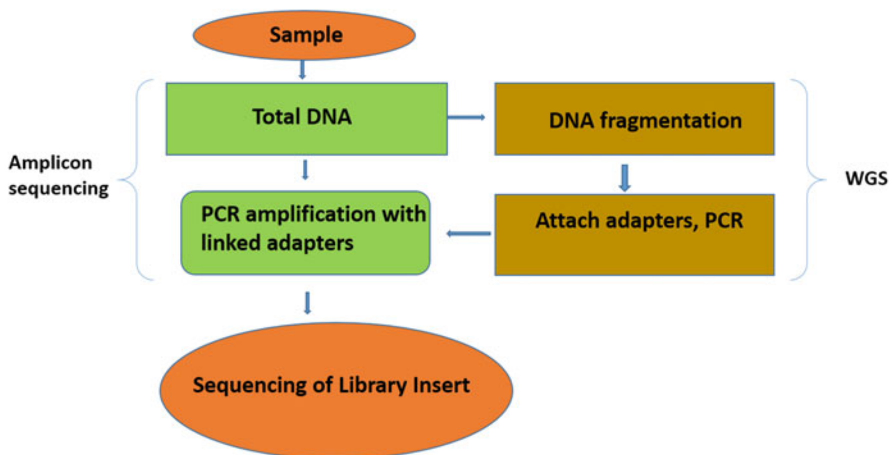


Fig. 24.6 Steps of library preparation for next-generation sequencing

high-fidelity enzymes (Nema 2012). Emulsion PCR and Bridge PCR are the techniques used for amplifying the sequencing material. Emulsion PCR uses emulsion oil, beads, PCR mix, and the library DNA to form a mixture of all in emulsion state wherein these miniscule droplets of emulsion serve as individual center of PCR. Hence, number of reactions take place simultaneously in these micro-droplets. It is believed that one droplet or a micro-well should contain one bead and one strand for a successful reaction. To achieve the purpose of resolution, this is a very good technique. However, having multiple steps like forming of emulsion followed by amplification and enrichment and finally breaking of the same emulsion seems to be a bit cumbersome and time consuming. This also has more chances of introducing errors. If any step goes wrong, the whole setup has to be redone. In a few cases, this may go well but you may just achieve a few good wells having one bead and one strand in it. Another chemistry is bridge PCR wherein complementary oligonucleotides to the primers which are attached to the DNA library fragments are densely coated on flow cell (the glass or silica chip where the sequencing chemistry occurs). The DNA is then allowed to attach to the surface of the cell at random. When the reagents for extension are added, the free ends of the single strands of DNA attach themselves to the surface of the cell via complementary primers, creating bridged structures. This is followed by a normal PCR to form double-stranded structures and is further used as templates for next reactions forming clusters after a few cycles.

24.6.3 Sequencing

As we have discussed earlier, instruments from different companies use different chemistries. "Sequencing by synthesis" is the most widely used one that is utilized by "ion torrent." It uses a template DNA complementary strand, which is synthesized by the addition of one base at a time. A semiconductor chip detects hydrogen ions produced due to the addition of a base during polymerization reaction. When the DNA library fragment is flooded by nucleotide fragments (dNTP), one at a time, the incorporation of compatible dNTP generates a hydrogen ion to be read by pH sensor of the sequencer. This kind of detection is faster as compared to fluorescence and camera scanning. The limitation comes when similar bases come consecutively and the signals merge. Similar steps are followed till bead additions in SOLiD sequencing. After the deposition of the beads, a primer is added to the adapter and a library of 8-mer probes is added to the beads. These probes are tagged with fluorescent dyes at 5′ end and the 3′ end. This probe would hybridize with the target sequence depending on the complementarity. These mers are then joined by the use of DNA ligase. Silver ions are then used to cleave the fluorescent dye from the fragment, resulting in the measurement of the fluorescence. After every round of sequencing, the extension product is melted of and the second round resumes with a primer which is one nucleotide shorter than the earlier one. Hence, after multiple rounds, the target is sequenced.

24.7 Next-Generation Sequencing: Raw Data and Its Interpretation

The wealth of big data generated by these high-throughput sequencing efforts can be used only after analysis and useful interpretations. It is sometimes difficult for normal laboratory workers to understand the complex algorithms used behind the simplification and interpretation of the data as generated by these technologies. Before matching the sequences with the available databases for deriving conclusions, the first step is to ensure if the data generated is of optimum quality or not. The interface between a user language and the machine language has to come as the first step wherein the data generated is converted into a readable format. Also this format has to be compatible with other analysis platforms. As different instruments use different chemistries for reactions and different scanning methods for reading the sequences, the first step is normally provided as embedded software in the instrument itself. The parameters of initial importance are number of successful reads, the read length, and quality of the reads (Q scores). In this step only, the samples that were mixed for multiplexing need to be de-multiplexed by the software. The adapters or the barcodes used to label the reactions initially are read by the software to tease them into their independent identity and to segregate these reads in separate files.

The outcome of the above step is a cleaned and quality data that is called as primary data and is used in secondary analysis as per the user and application requirements. This data is considered as clean with no background noise and errors like chimera. The secondary analysis steps are normally the intensive ones in terms of computational resources required and expertise of the analyst. Depending on the required information, the obtained sequences are either aligned with the standard genomes or the amplicons, and the step is called read mapping or alignment. Different mapping tools are available with different algorithms developed by workers from academics. These are listed in Table 24.3.

The mapping of the FASTQ file with the reference genome generated a sequence alignment map (SAM) or the binary alignment files (BAM). As the name goes, BAM files are the technical version of SAM files with binary codings. This saves space of storage but are readable by machines only. These are the standard formats and are in use universally. This data generated by these files is used in calibrating the quality scores again depending on the purpose of the sequencing. Many open-ended or paid software are available for in-depth analysis of data generated through various platforms of next-generation technology. Moreover, a few platforms have integrated software to take the analysis part also after generating FASTQ and SAM files. However, the integrated ones are provided under commercial licenses and are costly. Segregated tools from various platforms or free sites can also be employed in tandem to get the overall analysis done.

Table 24.3 Secondary sequence analysis tools

Name of the tool	URL/Ref	About the tool
BFAST	Homer et al. (2009)	It can align data produced by any of the current sequencing platforms, at par speed and accuracy, supports paired-end data, parallel and multi-threaded computation on a computer cluster
Bowti	http://bowtie-bio.sourceforge.net/index.shtml	It aligns short DNA sequences (reads) to the human genome at a rate of over 25 million 35-bp reads per hour
BWA	http://bio-bwa.sourceforge.net/	A software package for mapping low-divergent sequences against a large reference genome, such as the human genome
MOSAIK	Lee et al. (2014)	Can align reads generated by all the major sequencing technologies
NovoAlign	http://www.novocraft.com/products/novoalign/	Alignment quality scores using posterior alignment probability, paired-end alignment

24.8 Next-Generation Sequencing: Applications in Forensic Science

This chapter was written with an intent of making the users understand the technology and giving them an indication about its usage possibilities. However, no technology is limited with the applications available in a given piece of time. There is a good scope for innovation in this area. Technically speaking, next-generation sequencing has many applications; a few of which are large whole-genome sequencing of higher eukaryotes like plants, animals, and human, small whole-genome sequencing of microbes, etc., exome and large panel sequencing, amplicon-based or gene panel-based targeted gene sequencing, profiling of single-cell-like scRNAseq, scDNA—Seq, oligo tagging assays, sequencing of transcriptome comprising total RNA, mRNA or gene expression profiling, targeted gene expression profiling, interaction analysis of DNA and protein, modification analysis like methylation, etc., 16S metagenomics sequencing, shotgun metagenomics and metatranscriptomics, cell free sequencing, liquid biopsy analysis, etc. Furthermore, there are many field opening up with the use of NGS technology. Forensic research is one such field that involve construction of DNA database, ancestry and phenotypic inference, monozygotic twin studies, body fluid and species identification, and forensic animal, plant, and microbiological analyses (Yang et al. 2014). It is up to the user and scope of the area of application that anyone or multiple approaches may be employed to get desirable outcomes. A bottleneck in Sanger sequencing for forensic STR analysis/CE typing is the ability to resolve alleles that are identical by size, but different by sequence. NGS has this advantage over STR analysis. At the same time, it has been proposed as a precise method for casework and human identification utilizing applications such as intra-STR SNPs and mtDNA SNPs.

24.8.1 Application of NGS for STR Analysis

Initial phase of NGS with shorter read length was considered non-suitable for STR analysis. However, technological advancements not only made this possible but also allow for the detection of newer STR locus. Other advantages being the discriminating capability of NGS for the resolution of alleles having similar length, differentiation of mixed DNAs in the given samples, and to resolve complex paternity cases (Irwin et al. 2011; Bornman et al. 2012; Warshauer et al. 2013; Van Neste et al. 2014).

24.8.2 Mitochondrial Genome Analysis

Mitochondrial DNA (mtDNA) being from maternal origin in humans serves to provide plenty of materials (number of copies per cell) for analysis of complex cases wherein chromosomal DNA has limitation of having just two versions of a gene. This feature makes it a precious tool in forensic science as most of the time the quantities of the trace evidence is meager. Mitochondria are relatively protected to safeguard the genetic material, whereas nuclear DNA becomes degraded easily. Mitochondrial DNA has been isolated from complex samples like old and deteriorated hairs, bones, and teeth, where finding quality DNA with nuclear origin is either limiting or degraded. Another feature of mitochondrial DNA is heteroplasmy in which a cell can have some mitochondria that have a mutation in the mtDNA and some that do not. This could have forensic application as there are different levels of variations between related and unrelated individuals. Hypervariable sections of an mtDNA sample could be compared with the hypervariable sections of another. With the advent of NGS and its increasing usage in forensic science, it is not feasible to work out the characteristics in detail, for example, heteroplasmy. The whole-genome type of approach deciphering the total mtDNA may provide further details about the known as well as new mutations in this analysis (Tang et al. 2013).

24.8.3 Other Evidence from Research

When a new technology is found useful in the field of forensics, there come various derivatives of the applications which may be used for different novel purposes. Here are some of the examples wherein workers have exploited the immense potential of NGS for harvesting newer information or digging out the information from hard to reach samples. In a study by Xu and coworkers, it is proposed that prediction of age using DNA methylation and "support vector regression model" is possible now. Hence, they state that their strategy is highly useful in forensic practice as far as estimating the individual age is concerned (Xu et al. 2015). Similarly, a study by Vidaki and coworkers have attempted to create a similar age prediction test with better sensitivity and accuracy using a method based on NGS. The methylation

status of the selected 16cCpG sites was assessed and quantified in the method using the Illumina MiSeq® platform. A set of 46 blood samples were taken here to validate the method and DNA standards of known methylation levels served as control. They could achieve MAE = 7.5 years and hope that future optimization of the method will improve both the prediction accuracy and reproducibility (Vidaki et al. 2017).

NGS or massively parallel sequencing (MPS) is not only serving to study complex samples but also coming handy in evaluating different existing technologies for the analysis of complex study material. Similarly, a study to analyze petrosal bone as a source of DNA was conducted and CE- and MPS-based analyses were compared for forensic identification of challenging cranial bones. Cranial bones and other skeletal materials were processed for doing STR typing, and the results were verified using MPS. The outcomes of the analysis supported the potential of MPS for studying degraded samples with an advantage and possibility of getting additional and useful information (Kulstein et al. 2017). Studies are coming up with next-generation sequencing for studying epigenetic markers like analysis of tissue-specific DNA methylation. Workers have distinguished between the epigenetic signals from 5000 pairs of monozygotic twins (Bell and Spector 2011). Identification of body fluid, tissue, and the age accurately and detection of fabricated DNA evidence could be done using the epigenetic markers (Lee et al. 2012; Bocklandt et al. 2011). Along with epigenetic markers, microRNA and small noncoding RNAs are also being used more and more for different forensic applications and could be well studied using next-generation sequencing in a high-throughput and time- and cost-effective manner. As microRNA (miRNA) are smaller in size and are protected by proteins and other matrices, they are proven to be more stable as compared to other RNA species. Forensically useful biological fluids like blood, semen, vaginal fluid, menstrual blood, saliva, urine, feces, and perspiration were used in a study to look for the usefulness of miRNA expression which was read through NGS. A typical finding was that the small RNA from bacterial origin in some fluids (vaginal fluid, saliva, and feces) interfered in the yield of miRNA (Seashols-Williams et al. 2016). At the same time, some other studies have also tried to harness the potential of miRNA and ncRNA (noncoding RNA) sequencing in forensic investigations (Courts and Madea 2011; Wang et al. 2012; Tam et al. 2014; Ghildiyal and Zamore 2009; Courts and Madea 2010).

Identification of insects is an integral part of forensic entomology and is getting great help with NGS. Identification of insects and arthropods has a potential to support medico-criminal investigations like abuse cases, movement of corpse, and the most useful one postmortem interval estimates. It may also help in the detection of biological toxins, determination of the location of the incident, and in finding the presence and time of the infliction of wounds. An interesting study in this area dealt with the sampling of the insect population from the house of a women who died but the cause of death was unknown. The technology could help not only in identifying the insects but also the extraction and analysis of human DNA from the GI tract of the insect. They could individuate the presence of human DNA in the samples and determined the genetic profile (Pilli et al. 2016).

24.8.4 Microbiome-Related Work

With the advent of sensitive and high-throughput technology of DNA extraction from very little amount of samples and using the same for next-generation sequencing has made, it is possible to know about all cultivable and non-cultivable microbes and deciphering if there are some typical species or groups of microbes present as signatures in a particular crime scene exhibit and can be linked with the individuals. Research on the microbial interactions between the individuals and their environment including the material they come in contact with has supported the use of this in forensic investigations (Fierer et al. 2010). Even the microbial signature of the members of a same family could be used to predict their relatedness and differentiation from other members outside of the family. Within that family the individuals can also be differentiated by the microbiome of the surfaces with which these different individuals come in contact with on a regular basis (Lax et al. 2014). A study has shown that microbiome has a systematic and time bound succession after the death of the individual (Metcalf et al. 2013) which can be of importance in finding the chronology and direction of the event. Even there are reports to demonstrate that microbial community structure was determined by both surface type and participant (Lax et al. 2015).

The role of microbial forensics in bioterrorism is an important area of work. A comprehensive list of microbes consisting of around 14,000 strains of bacteria has been prepared to understand the potential health threat to humans (Taylor et al. 2001). However, to identify one or two from these, many number of strains in an unknown spread of pathogens is an impossible task. Hence, a tool like NGS can play an important role here by sequencing all available DNAs in a given sample in a short duration of time and with good precision. This can finally alert the system about a potential threat or can easily identify a hoax. A linked threat is about “agroterrorism” wherein “there is a deliberate misuse of biological agents against agriculture, including crops and cattle.” Agroterrorism contributed to the emergence of agricultural bioforensics, and the role of next-generation sequencing is very handy and significant here also.

In parallel to individual research efforts, commercial efforts are coming up for creating and optimizing dedicated sequencing instruments as per the needs of the forensic workflow. MiSeqFGx™ Forensic Genomics System is one such system which is designed and validated for forensic genomics applications. The system is capable of performing various tasks such as detection of autosomal STR markers, Y- and X-STRs, and SNP marker sets. These SNPs can also be used for identifying the color of eye, hair, etc. (Kidd et al. 2006, 2014; Sanchez et al. 2006; Walsh et al. 2013; Phillips et al. 2009). Another series of instruments from Applied Biosystems™ with 3500 series are also available with specific features for human identification (HID) applications. They work in conjunction with the well-optimized reagents available as “Applied Biosystems Precision ID panels” and enable analysis of up to hundreds of forensically targeted markers such as STRs, SNPs, or the mtDNA genome. Similarly, kits like PowerSeq® Auto System and PowerSeq® Auto/Mito/Y System are available from Promega Corporation for STR analysis and added

Y-chromosome STRs and the control region of the mitochondria, respectively. Being a high-throughput parallel sequencing, NGS is capable of analyzing different classes of polymorphism simultaneously. Other commercial systems are also helping STR analysis and mtDNA analysis (Zeng et al. 2015; Gettings et al. 2016; Kim et al. 2016; Parson et al. 2013; Zhou et al. 2016). Also, the problem of isolating quality DNA from mixed samples, trace samples, and samples contaminated with various inhibitors remained a challenge till date. Addressing this has been started with complementary technologies developed along with NGS systems. An example of same is the development of automated forensic DNA isolation systems like Auto-Mate Express™ forensic DNA extraction system. They help in improving the yield, reproducibility, and integrity of DNA that is obtained from normal as well as the challenging forensic samples. There are specific collection tubes/swabs/buffers developed by different commercial bodies in order to collect the sample without microbial contamination while preserving nucleic acid integrity. They are available from various companies and with various commercial names.

References

Anon A (1996) The evaluation off forensic evidence. Proc Natl Acad Sci 94(11):5498–5500

Bell JT, Spector TD (2011) A twin approach to unraveling epigenetics. Trends Genet 27 (3):116–125

Bocklandt S, Lin W, Sehl ME, Sanchez FJ, Sinsheimer JS, Horvath S, Vilain E (2011) Epigenetic predictor of age. PLoS One 6(6):e14821

Bodner M, Bastisch I, Butler JM, Fimmers R, Gill P, Gusmao L, Morling N, Phillips C, Prinz M, Schneider PM, Parson W (2016) Recommendations of the DNA Commission of the International Society for Forensic Genetics (ISFG) on quality control of autosomal Short Tandem Repeat allele frequency databasing (STRidER). Forensic Sci Int Genet 24:97–102

Bornman DM, Hester ME, Schuetter JM, Kasoji MD, Minard-Smith A, Barden CA, Nelson SC, Godbold GD, Baker CH, Yang B, Walther JE, Tornes IE, Yan PS, Rodriguez B, Bundschuh R, Dickens ML, Young BA, Faith SA (2012) Short-read, high-throughput sequencing technology for STR genotyping. Biotech Rapid Dispatches 2012:1–6

Borsting C, Morling N (2015) Next generation sequencing and its applications in forensic genetics. Forensic Sci Int Genet 18:78–89

Braslavsky I, Hebert B, Kartalov E, Quake SR (2003) Sequence information can be obtained from single DNA molecules. Proc Natl Acad Sci U S A 100(7):3960–3964

Courts C, Madea B (2010) Micro-RNA – a potential for forensic science? Forensic Sci Int 203 (1–3):106–111

Courts C, Madea B (2011) Specific micro-RNA signatures for the detection of saliva and blood in forensic body-fluid identification. J Forensic Sci 56(6):1464–1470

Eduardoff M, Santos C, de la Puente M, Gross TE, Fondevila M, Strobl C, Sobrino B, Ballard D, Schneider PM, Carracedo A, Lareu MV, Parson W, Phillips C (2015) Inter-laboratory evaluation of SNP-based forensic identification by massively parallel sequencing using the ion PGM. Forensic Sci Int Genet 17:110–121

Fierer N, Lauber CL, Zhou N, McDonald D, Costello EK, Knight R (2010) Forensic identification using skin bacterial communities. Proc Natl Acad Sci U S A 107(14):6477–6481

Gettings KB, Aponte RA, Kiesler KM, Vallone PM (2015a) The next dimension in STR sequencing: Polymorphisms in flanking regions and their allelic associations. Forensic Sci Int Genet Suppl Ser 5:e121–e123

Gettings KB, Aponte RA, Vallone PM, Butler JM (2015b) STR allele sequence variation: current knowledge and future issues. Forensic Sci Int Genet 18:118–130

Gettings KB, Kiesler KM, Faith SA, Montano E, Baker CH, Young BA, Guerrieri RA, Vallone PM (2016) Sequence variation of 22 autosomal STR loci detected by next generation sequencing. Forensic Sci Int Genet 21:15–21

Ghildiyal M, Zamore PD (2009) Small silencing RNAs: an expanding universe. Nat Rev Genet 10 (2):94–108

Hares DR (2015) Selection and implementation of expanded CODIS core loci in the United States. Forensic Sci Int Genet 17:33–34

Homer N, Merriman B, Nelson SF (2009) BFAST: an alignment tool for large scale genome resequencing. PLoS One 4(11):e7767

Irwin J, Just R, Scheible M, Loreille O (2011) Assessing the potential of next generation sequencing technologies for missing persons identification efforts. Forensic Sci Int Genet Suppl Ser 3:447–448

Kaye DH (1995) The forensic debut of the NRC's DNA report: population structure, ceiling frequencies and the need for numbers. Genetica 96(1–2):99–105

Kidd KK, Pakstis AJ, Speed WC, Grigorenko EL, Kajuna SL, Karoma NJ, Kungulilo S, Kim JJ, Lu RB, Odunsi A, Okonofua F, Parnas J, Schulz LO, Zhukova OV, Kidd JR (2006) Developing a SNP panel for forensic identification of individuals. Forensic Sci Int 164(1):20–32

Kidd KK, Speed WC, Pakstis AJ, Furtado MR, Fang R, Madbouly A, Maiers M, Middha M, Friedlaender FR, Kidd JR (2014) Progress toward an efficient panel of SNPs for ancestry inference. Forensic Sci Int Genet 10:23–32

Kim EH, Lee HY, Yang IS, Jung SE, Yang WI, Shin KJ (2016) Massively parallel sequencing of 17 commonly used forensic autosomal STRs and amelogenin with small amplicons. Forensic Sci Int Genet 22:1–7

Kulstein G, Hadrys T, Wiegand P (2017) As solid as a rock-comparison of CE- and MPS-based analyses of the petrosal bone as a source of DNA for forensic identification of challenging cranial bones. Int J Legal Med 132(1):13–24

Lax S, Smith DP, Hampton-Marcell J, Owens SM, Handley KM, Scott NM, Gibbons SM, Larsen P, Shogan BD, Weiss S, Metcalf JL, Ursell LK, Vazquez-Baeza Y, Van Treuren W, Hasan NA, Gibson MK, Colwell R, Dantas G, Knight R, Gilbert JA (2014) Longitudinal analysis of microbial interaction between humans and the indoor environment. Science 345 (6200):1048–1052

Lax S, Hampton-Marcell JT, Gibbons SM, Colares GB, Smith D, Eisen JA, Gilbert JA (2015) Forensic analysis of the microbiome of phones and shoes. Microbiome 3:21

Lee HY, Park MJ, Choi A, An JH, Yang WI, Shin KJ (2012) Potential forensic application of DNA methylation profiling to body fluid identification. Int J Legal Med 126(1):55–62

Lee WP, Stromberg MP, Ward A, Stewart C, Garrison EP, Marth GT (2014) MOSAIK: a hash-based algorithm for accurate next-generation sequencing short-read mapping. PLoS One 9(3): e90581

Levene MJ, Korlach J, Turner SW, Foquet M, Craighead HG, Webb WW (2003) Zero-mode waveguides for single-molecule analysis at high concentrations. Science 299(5607):682–686

Metcalf JL, Wegener Parfrey L, Gonzalez A, Lauber CL, Knights D, Ackermann G, Humphrey GC, Gebert MJ, Van Treuren W, Berg-Lyons D, Keepers K, Guo Y, Bullard J, Fierer N, Carter DO, Knight R (2013) A microbial clock provides an accurate estimate of the postmortem interval in a mouse model system. eLife 2:e01104

Munroe DJ, Harris TJ (2010) Third-generation sequencing fireworks at Marco Island. Nat Biotechnol 28(5):426–428

Muzzey D, Evans EA, Lieber C (2015) Understanding the basics of NGS: from mechanism to variant calling. Curr Genet Med Rep 3(4):158–165

Nature (2020). https://www.nature.com/subjects/next-generation-sequencing

Nema V (2012) Rectification of artificial molecular recombination with the use of high fidelity enzyme in the amplification of 16S rDNA sequences from stool sample. All Res J Biol 3:6–9

Novroski NMM, King JL, Churchill JD, Seah LH, Budowle B (2016) Characterization of genetic sequence variation of 58 STR loci in four major population groups. Forensic Sci Int Genet 25:214–226

Parson W, Strobl C, Huber G, Zimmermann B, Gomes SM, Souto L, Fendt L, Delport R, Langit R, Wootton S, Lagace R, Irwin J (2013) Evaluation of next generation mtGenome sequencing using the Ion Torrent Personal Genome Machine (PGM). Forensic Sci Int Genet 7(5):543–549

Parson W, Huber G, Moreno L, Madel MB, Brandhagen MD, Nagl S, Xavier C, Eduardoff M, Callaghan TC, Irwin JA (2015) Massively parallel sequencing of complete mitochondrial genomes from hair shaft samples. Forensic Sci Int Genet 15:8–15

Phillips C, Prieto L, Fondevila M, Salas A, Gomez-Tato A, Alvarez-Dios J, Alonso A, Blanco-Verea A, Brion M, Montesino M, Carracedo A, Lareu MV (2009) Ancestry analysis in the 11-M Madrid bomb attack investigation. PLoS One 4(8):e6583

Pilli E, Agostino A, Vergani D, Salata E, Ciuna I, Berti A, Caramelli D, Lambiase S (2016) Human identification by lice: a next generation sequencing challenge. Forensic Sci Int 266:e71–e78

Ronaghi M, Uhlen M, Nyren P (1998) A sequencing method based on real-time pyrophosphate. Science 281(5375):363–365

Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, Leamon JH, Johnson K, Milgrew MJ, Edwards M, Hoon J, Simons JF, Marran D, Myers JW, Davidson JF, Branting A, Nobile JR, Puc BP, Light D, Clark TA, Huber M, Branciforte JT, Stoner IB, Cawley SE, Lyons M, Fu Y, Homer N, Sedova M, Miao X, Reed B, Sabina J, Feierstein E, Schorn M, Alanjary M, Dimalanta E, Dressman D, Kasinskas R, Sokolsky T, Fidanza JA, Namsaraev E, McKernan KJ, Williams A, Roth GT, Bustillo J (2011) An integrated semiconductor device enabling non-optical genome sequencing. Nature 475(7356):348–352

Sanchez JJ, Phillips C, Borsting C, Balogh K, Bogus M, Fondevila M, Harrison CD, Musgrave-Brown E, Salas A, Syndercombe-Court D, Schneider PM, Carracedo A, Morling N (2006) A multiplex assay with 52 single nucleotide polymorphisms for human identification. Electrophoresis 27(9):1713–1724

Scheible M, Loreille O, Just R, Irwin J (2014) Short tandem repeat typing on the 454 platform: strategies and considerations for targeted sequencing of common forensic markers. Forensic Sci Int Genet 12:107–119

Seashols-Williams S, Lewis C, Calloway C, Peace N, Harrison A, Hayes-Nash C, Fleming S, Wu Q, Zehner ZE (2016) High-throughput miRNA sequencing and identification of biomarkers for forensically relevant biological fluids. Electrophoresis 37(21):2780–2788

Shendure J, Porreca GJ, Reppas NB, Lin X, McCutcheon JP, Rosenbaum AM, Wang MD, Zhang K, Mitra RD, Church GM (2005) Accurate multiplex polony sequencing of an evolved bacterial genome. Science 309(5741):1728–1732

Tam S, de Borja R, Tsao MS, McPherson JD (2014) Robust global microRNA expression profiling using next-generation sequencing technologies. Lab Investig 94(3):350–358

Tang S, Wang J, Zhang VW, Li FY, Landsverk M, Cui H, Truong CK, Wang G, Chen LC, Graham B, Scaglia F, Schmitt ES, Craigen WJ, Wong LJ (2013) Transition to next generation analysis of the whole mitochondrial genome: a summary of molecular defects. Hum Mutat 34 (6):882–893

Taylor LH, Latham SM, Woolhouse ME (2001) Risk factors for human disease emergence. Philos Trans R Soc Lond B Biol Sci 356(1411):983–989

Van Neste C, Vandewoestyne M, Van Criekinge W, Deforce D, Van Nieuwerburgh F (2014) My-Forensic-Loci-queries (MyFLq) framework for analysis of forensic STR data generated by massive parallel sequencing. Forensic Sci Int Genet 9:1–8

Vidaki A, Ballard D, Aliferi A, Miller TH, Barron LP, Syndercombe Court D (2017) DNA methylation-based forensic age prediction using artificial neural networks and next generation sequencing. Forensic Sci Int Genet 28:225–236

Walsh S, Liu F, Wollstein A, Kovatsi L, Ralf A, Kosiniak-Kamysz A, Branicki W, Kayser M (2013) The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA. Forensic Sci Int Genet 7(1):98–115

Wang Z, Luo H, Pan X, Liao M, Hou Y (2012) A model for data analysis of microRNA expression in forensic body fluid identification. Forensic Sci Int Genet 6(3):419–423

Warshauer DH, Lin D, Hari K, Jain R, Davis C, Larue B, King JL, Budowle B (2013) STRait Razor: a length-based forensic STR allele-calling tool for use with second generation sequencing data. Forensic Sci Int Genet 7(4):409–417

Xu C, Qu H, Wang G, Xie B, Shi Y, Yang Y, Zhao Z, Hu L, Fang X, Yan J, Feng L (2015) A novel strategy for forensic age prediction by DNA methylation and support vector regression model. Sci Rep 5:17788

Yang Y, Xie B, Yan J (2014) Application of next-generation sequencing technology in forensic science. Genomics Proteomics Bioinformatics 12(5):190–197

Zeng X, King JL, Stoljarova M, Warshauer DH, LaRue BL, Sajantila A, Patel J, Storts DR, Budowle B (2015) High sensitivity multiplex short tandem repeat loci analyses with massively parallel sequencing. Forensic Sci Int Genet 16:38–47

Zhou Y, Guo F, Yu J, Liu F, Zhao J, Shen H, Zhao B, Jia F, Sun Z, Song H, Jiang X (2016) Strategies for complete mitochondrial genome sequencing on Ion Torrent PGM platform in forensic sciences. Forensic Sci Int Genet 22:11–21

25 Oral Microbes: A Hidden Yet Powerful Evidence for Futuristic Forensic Investigation

Bhargav C. Patel and Mayuri Ganguly Soni

Abstract

Microbiology is an ancient science, but its application in forensic investigation is a novel endeavor. Because of modern sequencing technologies, it is proven that human body harbors millions of microbes at various sites like oral cavity, gut, skin, urogenital tract, etc. Many of the species are unique and specific for the body site and can be exploited for forensic investigation. This chapter digs into details for oral microbes which seem to be an interesting niche of microbial diversity for unique application in some of the forensic investigations. The microbial diversity and dynamics of the oral cavity is explained in detail with the important online databases for the same. The unique and specific microbes for forensic applications are discussed with modern techniques for isolation, detection, and identification of microbes. The human identification, ethnicity, and postmortem interval (PMI) estimation are explained in more depth along with the detailed literature available till date. The chapter unveils the real value of these invisible creatures in human oral cavity for modern BioForensic investigation.

Keywords

Oral microbes · Forensic microbiology · Diversity · Bite mark · Body fluid · Post-mortem interval (PMI) · Ethnicity · Human identification

B. C. Patel (✉)
Laboratory of Forensic Biology and Biotechnology, School of Forensic Science, National Forensic Sciences University, Gandhinagar, Gujarat, India
e-mail: bhargav.patel@gfsu.edu.in

M. G. Soni
Police Hospital, Rourkela, Odisha, India

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_25

25.1 Introduction

Since the earliest days of microbiology, it has been clear that all humans carry many of the same microbial lineages. The microbes living in or on our bodies are now known as the human microbiome (Turnbaugh et al. 2007). A collective international effort, called the Human Microbiome Project (HMP) by the National Institutes of Health (Peterson et al. 2009), is now broadly known as the International Human Microbiome Consortium (IHMC). The consortium is aiming to characterize the microbiome of various parts of human body.

Human body can support several bacterial ecosystems all over the body especially in hair (Tridico et al. 2014), skin, gut, mouth, and urogenital tracts (Giampaoli et al. 2012; Juusola and Ballantyne 2005). Different studies for gut microflora showed extensive diversity of the gut microbiome between healthy humans (Ley et al. 2006a, b; Raman et al. 2010). Studies since past 5 years for skin bacterial diversity showed that in addition to conserved organisms, there is massive diversity in identities and richness (Costello et al. 2009; Grice et al. 2009; Fierer et al. 2008; Gao et al. 2007; Blaser 2010). Among these, the oral cavity is a unique ecological niche because it can provide hard, non-shedding surfaces that are teeth which are durable and accessible for microbial colonization. Our oral cavity harbors a diverse group of microorganisms that include bacteria, fungi, mycoplasma, protozoa, and sometimes may be viruses (Costello et al. 2009; Ventura Spagnolo et al. 2019). Altogether they are known as the oral microbiome. There are billions of bacteria in the oral cavity that forms ecosystems in different sites in the oral cavity. These ecosystems are dorsum of the tongue, buccal epithelium, the supragingival and subgingival tooth surfaces, crevicular surfaces, and tonsils where microflora varies from one site to another (Muhammad et al. 2012). These diverse ecosystems make saliva a rich source of a diverse group of bacteria.

25.2 Diversity of Oral Microflora

Our oral microflora constitutes more than 500 species of bacteria. The unique feature of oral bacteria is that they vary from mouth to mouth, though the majority of them are similar in most of the individuals and differences are created mainly by minority members of the bacterial group (Takeshita et al. 2016). More precisely diversity can be appreciated at species and subspecies level. A study by Bik et al. (2010) revealed how each oral cavity harbors a distinct population of microbial groups, but that such groups appear to be more identical when categorized at genus level. Previously designed ecological methods for larger species, such as co-occurrence evaluation, can significantly promote the study of diverse bacterial populations such as those present in the human body and improve our perception of the role of microbiome in human physiology.

25.2.1 Human Oral Microbiome Database (HOMD)

Human oral microbiome database is the first compiled human-associated microbiome characterization which offers resources for use in studying the microbiome's function in health and disease. HOMD's aim is to provide extensive information to the research community about nearly 700 prokaryote species that are found in the human oral cavity. HOMD is based entirely on a tentative naming scheme focused on the carefully curated 16S rRNA gene. The facility has sequenced more than 600 16S RNA gene libraries over the last 20 years and collected more than 35,000 clone sequences. The samples were collected from healthy individuals and participants with more than a dozen disease states such as caries, periodontal disease, endodontic diseases, and oral cancer. The HOMD ties data from sequences to phenotypic, phylogenetic, scientific, and bibliographic metadata. HOMD data structure, integration, and description may be used as a blueprint for microbiome data from other locations of the human body such as stomach, head, and vagina. Nearly 700 organisms are specified in the HOMD, of which 51% are officially listed, 13% are not mentioned (but cultivated), and 28% are recognized only as uncultivated phylotypes (http://www.homd.org, last accessed on 14 August 2020). There are about 150 genera, 700 species, in the HOMD collection. HOMD currently contains genomes for 400 oral taxa and even more than 1300 species of microorganisms.

Streptococcus, for instance, is a species with greater abundance than several genera (Butler et al. 2017). Streptococcus group has 43 species in the HOMD, 26 of which are named, 9 of which are not named, 7 are dropped, and 2 are lost.

Genomes for 30 oral taxa and 202 Streptococcus strains can be found on the HOMD. The genus *Prevotella* has 53 species in the HOMD, of which the genomes are accessible on HOMD for 32 species and 67 different strains. HOMD is focused on the microorganism cultivation. But constraint is quite a part of HOMD data's oral microorganisms that cannot be cultivated, of which as much as 20–60% is reported to be uncultivable, due to the limitation of growth conditions, microbial interaction, and so on.

25.2.2 Factors Affecting the Diversity in Oral Microflora

Many factors contributing to the reason for this mixed microbial diversity (Muhammad et al. 2012) described are as follows:

1. *Time*: Costello et al. (2009) have on four instances analyzed microbiome at 27 sites among seven to nine healthy persons. Such findings revealed that our microbiome has been customized, continuously varying between body environments and periods. In 300 healthy individuals in 18 body locations, the HMP Consortium has documented the composition and role of the human microbiome in 12–18 months. The populations in the oral cavity have become more variable over the duration of the sampling period (Anukam and Agbakoba 2017). In specimens from stable and vulnerable elderly people using 16S rRNA

Table 25.1 Colonization of bacteria in different age groups

Age group	General characteristics	Bacterial species
Just after birth	The sterile environment gets colonized. Mostly aerobic bacteria	*Streptococcus salivarius* commonly with *Staphylococcus albus*
Infants (by 6 months)	Anaerobic bacteria start colonizing	Veillonella species, Fusobacteria
Tooth eruption stage/Early childhood	They form biofilms over a hard surface	*Streptococcus mutans*; *Streptococcus sanguis*
Adolescence to adulthood	Anaerobes increase more	Increased *Bacteroides* and *Spirochaetes*
Edentulous mouth	Few *Bacteroides* and *Spirochaetes* but more of yeast	*Streptococcus sanguis* and *Streptococcus mutans* disappear
Dentures and another prosthesis	Bacteria that attach to hard surface again reappears	

sequencing analysis, Ogawa et al. (2018) studied microbiome composition of saliva. The research proposed that overall weakness is related to the structure and development of oral microbiota.

2. *Age*: The exciting fact about oral microflora is that it varies with age. Bacterial colonization in the oral cavity starts just a few hours after the birth. Microflora in the oral cavity changes continuously as various phases of teeth eruption and shedding goes on throughout the lifespan of a human because physiological changes take place each time a tooth erupts or sheds which changes the environment of oral cavity, leading to microbial shift (Xu et al. 2014). Table 25.1 illustrates different groups of bacteria colonizing the oral cavity as the age advances. This suggests that a unique group of bacteria can be cultivated or identified from the oral cavity of an individual at a particular age. Anukam and Agbakoba (2017) obtained oral specimens from three randomly chosen women aged 56, 28, and 8 years, isolated DNA, and amplified the 16S rRNA V4 region employing custom-coded primers prior to Illumina MiSeq sequencing. The study found that microbes with fluctuating diversity inhabited oral cavities in females of different ages. As the human life expectancy has increased with progress of medical science, An et al. (2018) researched the physiology of the aging and oral disease processes, anticipating that possible geriatric health studies will be funded and that further clinical work would take care of the significance of age.
3. *Diet*: Lassalle et al. (2018) collected saliva from three pairs of hunter-gatherers and farmers living nearby in the Philippines. The findings indicated that significant dietary changes were selected for different commensal populations and were likely to play a role in the production of modern oral pathogens. Adler et al. (2013) have shown that the transformation from hunter-gatherer to farming has turned the oral microbial environment into a disease-related configuration. Modern oral microbial communities have been substantially less diverse than historical communities and can lead to post-industrial chronic oral diseases. Brito et al.

(2016) correlated mobile genes present in 81 North American metropolises to 172 Fiji agrarians using a fusion of genomics with metagenomics from the single cell. They experienced significant distinctions between Fijian and North American microbiomes in their mobile gene material. These results show that the abundance of certain genes can reflect ecological selection. Many research studies of well-preserved dental calculus were engaged in an ancient oral microbiome to examine genomic innovations, diseases, etc. (Metcalf et al. 2013; Brown et al. 1976; Galvão-Moreira et al. 2018).

4. *Extreme environment*: Brown et al. (1976) measured the preflight and postflight monitoring of changes in microbial populations at various intraoral sites. Microbiologic assessments showed noteworthy elevations in counts of specific anaerobic components of the oral microflora, *Streptococci*, *Neisseria*, *Lactobacilli*, and *Enteric bacilli*, which were believed to be diet related. The relative absence of hazardous intraoral changes to one's wellness is regarded as this study's most important result.
5. *pH of the Oral Cavity*: pH for bacterial growth is one of the most significant considerations. The low pH change of the oral cavity induces a microbial shift owing to whatever explanation. The therapeutic significance of oral cavity pH and associated microflora dynamics is largely documented in the literature. Dental caries is attributed to acid in oral biofilms known as plaque formed by commensal microbes. The pH of the environment may be the consequence of the acidogenicity of bacteria coupled with or without the pH enforced by foreign sources, like food and beverages (Kianoush et al. 2014). The initial "substantial core model" research by Kianoush et al. (2014) establishes a pH-distinctive taxa and shows improvements in bacterial abundance in the acidic to neutral pH ranges. Some of the bacterial species have been shown to be capable of metabolic action under moderately acidic environments in oral cavity, both in proteolytic and in saccharolytic *Prevotella* sp. The microbial dynamics in oral cavities in healthy people and people with dental disorders such as caries became possible thanks to new NGS technologies (Zhou et al. 2017). A change in acid and acid-bacterial consortia, as seen in this research, namely *Lactobacillus vaginalis* and *Streptococcus mitis*, which favors caries lesions (Galvão-Moreira et al. 2018), was indicated to induce lower pH. The structural features of a bacterial salivary population are affected by the pH (Zhou et al. 2017).
6. *Host defense*: Growth of bacteria is mostly influenced by the immune system of the individual. The immune system keeps the growth of bacteria in control to prevent it from causing disease. It generally depends on the genetic makeup of the individual. Every human body reacts differently to various microbes. So some bacteria can grow more preferentially in some individuals than others.
7. *Genetics*: Gomez et al. (2017) illustrated the importance of host genetics in assessing the dynamics of bacteria on tooth surfaces just beside the gums (supragingival plaque) and beyond the gum line. For both saliva and supragingival plaque microbiomes, the more similar the genomes of individuals, the more similar their microbiome composition—correctly pointing to the function of host genetics. They obtained data on microbiota, caries, and sugar intake

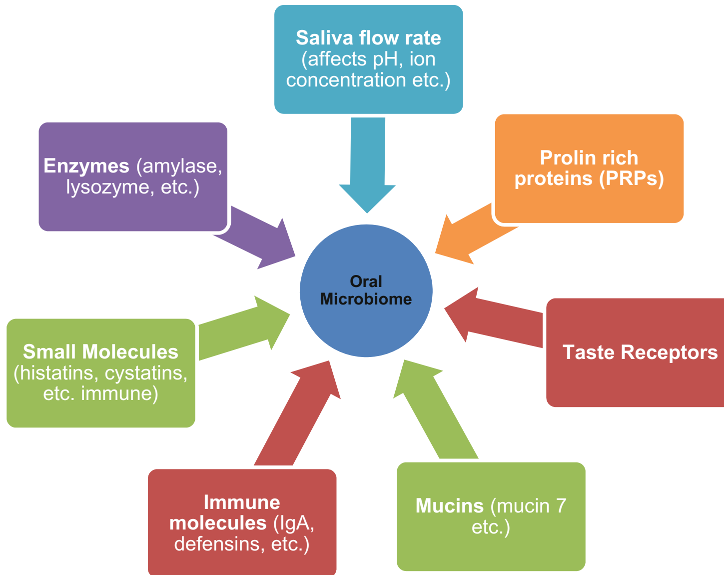


Fig. 25.1 Host secreted biomolecules that can affect oral microbiome heritability (Source: Davenport (2017))

for 485 twins aged 5–11 years. By comparing monozygotic (identical; MZ) and dizygotic (fraternal; DZ) twin pair microbiomes, they estimated heritability (h^2), which is the difference in the association between MZ and DZ twin pairs, for the comparative prevalence of bacterial taxa in at least half of their specimens. Of the 91 common bacteria investigated, nearly half show heritability at least 20%. Among these, several taxa's heritage estimates are relatively small, including *Prevotella pallens* ($h^2 = 0.65$), a *Veillonella* genus ($h^2 = 0.60$), and *Corynebacterium durum* ($h^2 = 0.54$), clearly showing the importance of the host genomes to oral occupancy by these microbes. As shown in Fig. 25.1, saliva and epithelial tissues comprise many compounds of known genetic variation. Proline proteins (PRPs) are significant component of saliva, causing bacterial adherence to tooth surfaces. Mucins function as tethering and energy sources for different microbiota. Immune molecules and other biomolecules (histatins, defensins, etc.) control oral microbial composition and inhibit pathogen invasion. Taste receptors make a contribution to food preferences, and certain enzymes catalyze food, like amylase. All these molecular groups generate specific dietary content for oral bacteria. Eventually, the salivary flow rate determines these molecule concentration and electrolytes.

8. *Other conditions*: In certain oral health-related conditions like diabetes, hypertension, smokers, dental caries, and periodontitis, phylogenetic diversity of oral bacteria is more than in healthy individuals (Takeshita et al. 2016). HMP reported that there were strong associations between whether individuals had been

breastfed as an infant, their gender, and their level of education with their community types at several body sites (Anukam and Agbakoba 2017). Galvão-Moreira et al. (2018) studied 46 female and 24 male patients, aged 18–40 years, and counted both groups' *Streptococcus mutans*. The study suggested that there was a significant difference for *S. mutans* levels in both groups.

All these conditions lead to a various combination of bacteria, and it seems to provide uniqueness to every individual. In general 15–16 genera of bacterial species can be cultivated from saliva (Takeshita et al. 2016; Nasidze et al. 2009a). As mentioned above bacterial community is influenced by many factors. Accordingly, three types of bacterial composition are observed during the analysis of saliva samples:

- *Streptococcus* dominant type
- *Prevotella* dominant type
- Neisseria, Haemophilus/Porphyromonas dominant type

Most of the studies reveal that oral microbiome constitutes *Streptococcus,* as dominant bacterial species along with *Prevotella*, *Veillonella*, *Neisseria*, *Haemophilus*, *Rothia*, *Porphyromonas*, and *Granulicatella* species (Bik et al. 2010) in considerable amount in most of the cases. Although bacterial community remains conserved in most of the individuals, the difference in genus dominance is often there. Even among genera, species differ between individuals. Thus each person's mouth harbors a unique community of bacterial species but more similar at the level of genus (Bik et al. 2010). Saliva samples from different geographic regions presented with a global pattern of diversity in the human oral microbiome. Diversity among individual from the same location was nearly the same as among different locations (Nasidze et al. 2009a).

25.2.3 Oral Microbiome Sequencing Techniques

Given the range of surfaces required for the processing of samples and the abundance of knowledge accessible to scientists and researchers from such tests, the human microbiome is theoretically a very effective forensic analysis tool. Further experiments have also shown that the makeup of the microbiota (taxonomical structures) differ over time (Flores et al. 2014), and the microbial concentrations of the gut and saliva are more stable.

Variety in human microbiome can be defined in many ways. The most recommended practice is the analysis by targeted sequencing of the variable regions of the 16S r RNA gene of the taxonomic distribution of samples. Several online and offline bioinformatics tools/software/databases are available to assemble similar sequences (traditionally 97% 16S rRNA sequence similarities) into groupings, termed Operational Taxonomic Units or OTUs (Schloss et al. 2009; Caporaso

et al. 2010; Arumugam et al. 2011). While it is still an imperfect taxonomic classification system, these OTUs lead ideally to species-level classification of microbial taxa (Goebel and Stackebrandt 1994). By comparing known bacterial species, the programs match OTUs to previously analyzed bacterial taxa and measure the distribution of distinct taxa in a given sample. Although these software differ in the specific algorithms and presumptions used, they all exhibit similar outcomes (Gajer et al. 2012).

As sequencing technologies continue to advance and sequencing costs decrease, whole-genome shotgun (WGS) approaches have begun to be used to test human microbiome rather than targeted 16S sequencing. WGS directly targets all gene content in a given microbial environment and is capable of differentiating microbial and taxa species to a greater extent than those of 16S rRNA amplicons (Schloissnig et al. 2012; Franzosa et al. 2015). Nevertheless, the capacity of microbiome WGS to assist in forensic investigations by linking artifacts and environments with individuals was poorly investigated (Zhernakova et al. 2016), and while it indicates considerable potential, it is currently limited as an forensic tool. The 16S rRNA sequences and their related metadata are the most common data and thus more commonly accessible than in meta-studies carried out by Adams et al. (2015) and Sze and Schloss (2016). The expanded size of data increases the statistical power and thus the efficiency of 16S rRNA information as a forensic tool (Sjödin et al. 2013). The Forensics Microbiome Database (FMD) created by a group that connects publicly accessible 16srRNA-derived taxa data or geolocation (http:/www.fmd.jcvi.org) is an example of using meta-samples for forensic analysis.

25.3 Forensic Applications

Recent advances in genetic data generation have fostered significant progress in microbial forensics (or forensic microbiology) through massive parallel sequencing (MPS). Initial uses in the contexts of biocrime, bioterrorism, and epidemiology are now accompanied by (1) the possibility of using microorganisms as additional evidence in criminal proceedings; (2) explaining causes of death (e.g., drowning, toxicology, hospital-acquired infections sudden infant death, and shaken baby syndromes); (3) assisting in the identification of human beings (skin, hair, and body fluid microbiomes); (4) for geographical diversity (soil microbiome); and (5) to estimate postmortem interval (thanatomicrobiome and epinecrotic microbial community). When contrasted with classical microbiological methods, MPS offers a wide range of privileges and alternative potentials. Prior to its application in the forensic context, however, critical efforts must be made to develop consolidated standards and guidelines through the creation of resilient and exhaustive reference databases (Oliveira and Amorim 2018).

25.3.1 Bite Mark

Human identification is the critical factor in the forensic investigation. It is very crucial to correctly identify the victim either dead or living as well as the perpetrator. There are many ways by which identification is being made. In sexual assault cases, one can come across bite marks which can then be used as evidence to identify the perpetrator. Physical and metric analytical methods are generally used to compare the bite marks with that of suspect's dentition. Saliva sample from bite marks contains oral epithelial cells that can be used to extract DNA and match with that of suspect's DNA sample, but human DNA can get degraded in due course of time due to the presence of various enzymes in saliva (Spradbery 2010). Saliva also contains oral bacteria which also gets transferred to the skin when the bite is inflicted according to Locard's principle. Bacteria can sustain harsh conditions like drying, degradation, and putrefaction, and many more because bacterial DNA is generally enclosed by a cell wall and cell membrane which protects the DNA. So bacterial DNA analysis can pave new way for identification.

Oral microbiome differs from skin microbiome. In bite wounds, bacteria are isolated from normal oral flora rather than skin flora. In a study when forearm was inoculated with tongue bacteria, the transferred bacteria were more similar to tongue bacteria than forearm bacteria. Alpha-hemolytic *Streptococci* are most frequently isolated bacteria from most of the bite wounds. Streptococcus group of bacteria are invariably present in dental plaque as long as the teeth are present in the oral cavity. Bite marks mostly involve teeth marks. So it is possible that this bacterial species might get transferred when the bite has been inflicted. A study by Brown and colleagues from the Department of Oral Biology at the University of Adelaide, South Australia in 1984, had proved that *Streptococcus salivarius* is the major species in saliva, and the rate of loss of bacteria is 45–50% per hour from the site of infliction. That means loss of recoverable bacteria was less as time passes. Prominent amplicons can be found up to 48 h, not beyond that. Bacterial sample recovered from a dead body is viable for more duration of time, but in living individuals, it may decrease over time due to washing the area of bite or bathing or using antiseptics over that area.

Direct amplification of oral streptococcal DNA samples from bite marks and incisor teeth of an individual who has inflicted the bite through PCR followed by denaturing gradient gel electrophoresis has given significant results (DGGE) (Hsu et al. 2012). Arbitrarily primed PCR (AP PCR) can be used to analyze bacterial DNA, and genotypes of bacterial DNA from bite mark and teeth can be compared. Some of the dominant streptococcal genotypes do not change over time. So this method is helpful in cases where suspect tooth sample is not available for a longer period, but a bacterial sample from the bite mark should be kept well preserved (Rahimi et al. 2005). In AP, PCR sampling and culture should be done within 24 h of the bite infliction, but in DGGE, sampling and culture time can be increased more than 24 h but the resolution of the result obtained is less. 16srRNA along with rpoB (RNA polymerase beta subunit) helps in the identification of bacteria to species and subspecies level as many bacterial species contain multiple copies of 16srRNA gene,

and the number of copies can vary between the species, and the hypervariable region of rpoB can bring down the level of identification to species and subspecies level. 16srRNA and rpoB1 reveal that *Streptococcus* is the most commonly detected genus in saliva, and according to rpoB2 gene analysis, *Rothia* is the most abundant genus. Experiments using this technique demonstrate that samples from all the individuals can be differentiated from each other and no two individuals had the same oral microbial profiles (Rahimi et al. 2005; Leake et al. 2016). RFLP technique is the most straightforward method to find out polymorphism in 16srRNA gene in bacteria and compare bite mark DNA to suspect's teeth DNA (Spradbery 2010). Multilocus sequence typing (MLST) scheme which is used to isolate and differentiate strains of microbial species by observing DNA sequence variations of housekeeping genes has been developed for oral *Streptococci* (Do et al. 2010). High-throughput sequencing technique is used to sequence 16srRNA, 16s–23s ITS (intergenic spacer), and rpoB gene region of oral streptococcal DNA. Results obtained had proven that *Streptococci* amplified from bite marks originated from teeth only. rpoB gene gave the most satisfactory results, whereas the discriminatory power of 16srRNA and ITS was less (Kennedy 2011).

A very crucial factor for forensic investigation is the temporal stability of the reference sample which means if the sample is collected and subjected for experiments, it should give the same results. Temporal stability of oral streptococcal DNA has been assessed in various studies. AP PCR, as well as sequencing techniques, revealed that although the oral streptococcal population is dynamic with species number and proportions fluctuating over time, dominant strains are retained for a longer period, and they were able to identify the suspect with the same potential as earlier correctly (Rahimi et al. 2005; Kennedy 2011).

In infected human bites also *Streptococcus* species is the most abundant with *Streptococcus anginosus* as dominant group followed by *Staphylococcus aureus* and *Eikenella corrodens* as aerobic species and *Prevotella*, *Fusobacterium*, and *Veillonella* as anaerobic species. Even in abscess bite mark wounds, *Streptococcus* and *Staphylococcus* species are abundant (Talan et al. 2003).

Thus it is well proven that salivary microflora can be used as a potential tool for identification as bacterial DNA resist degradation more efficiently and can differentiate two individuals along with their lifestyle. Oral flora can even distinguish between twins that means genetics do not have any influence on the oral microbiome. Antibiotics do have a promising effect on the oral microbiome, but it recovers back to pre-antibiotic diversity after a few days of treatment (Leake et al. 2016).

25.3.2 Body Fluid Prediction

Biological evidence in a crime scene usually constitutes various body fluids that too in trace amounts. It is very crucial to identify the type of body fluid to know the nature of the crime. Various presumptive tests are still preferred to identify the body fluids using specific enzymes that are particular for that fluid. The result is obtained

as a color change. However, this test may often give false-positive result due to the presence of that enzyme in a small amount in other body fluids. An alpha-amylase enzyme present in saliva is generally used to identify the saliva, but this enzyme is generally present in small amount in other body fluids like urine and semen that can falsely be recognized as saliva. RNA-based assay targeting saliva-specific gene products have been recently introduced for the detection of saliva (Nakanishi et al. 2009).

An alternative method for the detection of body fluids is taxonomic profiling of microbes. This can be possible for fluids rich in microbes such as saliva, vaginal secretion, feces, and menstrual blood while sterile or nearly sterile body fluids such as blood, semen, and tears are difficult to recognize. Sequencing of microbial 16srRNA gene can be done to discriminate saliva collected from the human body or a crime scene (Hanssen et al. 2017). Oral *Streptococci* such as *S. salivarius* and *S. mutans* are only found in saliva and not in semen, urine, skin, or vaginal fluid. In a study, *S. salivarius* was obtained in all and *S. mutans* in most of the mock forensic samples cigarette butt, cotton gauze wiped licked skin and aged saliva samples. Neither *S. salivarius* nor *S. mutans* can be detected in the saliva of other animals. Hence, oral streptococci are the new marker for detection of saliva (Nakanishi et al. 2009).

25.3.3 Postmortem Interval (PMI) Estimation

One of the most complex microbiomes in the human body is in the oral cavity. This has been shown to be the second most complex microbiome of the body after the gastrointestinal tract. The native microorganisms biodegrade the cadaver upon death. Oral cavity and gastrointestinal tract microbiomes, thereby play a crucial role in biological decomposition of the cadaver. Many studies are reported for the application of different microbial succession for the PMI estimation (Damann et al. 2015; Burcham et al. 2016; Guo et al. 2015; Metcalf 2019; DeBruyn and Hauther 2017; Hyde et al. 2013). It has been well established that skin microbiome can be used to calculate the postmortem interval of the dead body (Aaspõllu et al. 2011). However, a study (Aaspõllu et al. 2011) reported that most variation over decomposition occurred in mouth microbial populations, while group size and composition were most consistent in the rectum. Recently, oral microbes are assessed to calculate the time since death (Adserias-Garriga et al. 2017). *Firmicutes* and *Actinobacteria* are the predominant phyla in the fresh stage of the cadaver. This phylum generally includes *Lactobacilli*, *Staphylococcus aureus*, *Carnobacteria*, *Veillonella*, *Streptococcus*, *Campylobacteria*, *Micrococcus*, *Bifidobacteria*, *Actinomycetes*, and *Corynebacterium*. *Tenericutes* presence corresponds to bloat stage with *Peptostreptococcus* and *Bacteroides* as dominant species in an early stage, and *Clostridiale* in a later stage. *Firmicutes* is the predominant phyla in advanced stage being that *Firmicutes* are different from that of the fresh stage. They are mainly soil representatives. Dry remains mainly habitat by *Bacillus* and *Clostridiale*

(Adserias-Garriga et al. 2017). Thus, careful analysis of oral microflora can help in estimating accurate time since death of the dead victim.

25.3.3.1 Forensic Microbiome Database

The Forensic Microbiome Database (FMD) is a human microbiome analysis tool which, regardless of a sequencing tool or the sequenced region acquired from various body sites, compares publicly accessible 16s rRNA datasets to metadata in terms of forensic analysis.

The Forensic Microbiome Database (FMD) aims to:

- Provide an evidence-based tool for the forensic and scientific community, thoroughly documenting the literature and sequences of microbial samples.
- Maintain a database of results, meta-data, and related analyzes managed by quality.
- Build a website consisting of software that allow users to evaluate their individual data, interpret and quantify the results, and compare the results with the existing data for the public at large.

Machine learning has been used to determine the geographic position from where the sample was collected, using the comparative richness of bacterial taxa in one sample (Edgar 2013; Human and Project 2012). The more detailed and complex the FMD data are, the more precise the forecasts are. Such forecasting could be used to assess the location of a survivor of trafficking in human beings or as evidence in a crime scene to limit a person's investigation range.

Originally, this research involved the selection and management of samples of human microbiome across five countries worldwide, namely Chile, Barbados, Hong Kong, and two South African sites. Oral and stool samples from healthy women aged 18–26 years have been obtained. Immediately frozen samples were sent to J. Sequencing and subsequent research laboratory of Craig Venter Institute (JCVI).

This also gathers data from databases such as NCBI, EBI, and DDBJ. A quick read (src), which stores raw sequencing for thousands of scientific studies (Fig. 25.1), is held in NCBI (National Center for Biotechnology Information) (Fig. 25.1). Scientists deposit their sequences with the SRA to make their experiments reproductive and use them for new insights through other experiments. In order to identify newly added studies to the database, the FMD regularly checks the SRA. It is of concern to only a certain section of the samples contained in the SRA. These are human-associated studies of the microbiome—sequence of readings of bacterial 16s rRNA on a human body site with appropriate metadata. Using 16s rRNA samples, the relative diversity of the population of microbes in that body site can be determined. SRA samples comprise details on each sample—primarily how it was processed and sequenced.

Within NCBI, the BioProject contains several samples of the same study in an umbrella group. This BioProject explains the analysis and permits users to access all SRA samples. To query databases in order to locate these studies, search words such as "human microbiome 16S" are used. If a sample is found that fulfills the

specifications mentioned above, other conditions need to be met, since the intended FMD feature must be further analyzed. Metadata linked with a report must at least include the geographic location of the collection sample and the reported document detailing the report (in the city level). Test inclusion does not allow other details such as the age and gender of the human host to be useful and preserved in the database. Most of the metadata is obtained from the website where the study has been downloaded, but also from other sources such as the manuscripts released. Actually, the public database and its associated research methods contain only safe human samples, but unhealthy specimens are still collected for potential inclusion.

Samples of other sequence databases such as EBI and MG-RAST with a similar organizational structure are obtained in addition to NCBI.

25.3.3.2 The Data Analysis and Interpretation Process

The FMD is a human microbiome analysis resource for 16S rRNA sequence data obtained from multiple body sites to relevant corresponding metadata as it relates to forensics, utilizing several analytic techniques. Following are the major analysis that can be performed on this database.

- Looking at the taxonomic distribution of individual samples from the collected publicly available 16S rRNA sequence data.
- Comparing the taxonomic distribution of multiple samples from the collected publicly available 16S rRNA sequence data.
- Comparing a user-supplied sample with all the publicly available 16S rRNA sequence data and geolocating it through the closest matching taxonomic distributions.

For more information about the 16S rRNA sequence data used in the FMD, one can refer to the statistics page.

25.3.3.3 Where Does This Data Come from?

The FMD collates the accessible 16S rRNA sequence data. Oral and stool samples from healthy women were obtained and analyzed in cooperation with the co-regional investigators of different locations around the world (Hong Kong, Barbados, Chile and two sites in South Africa). Sequence data are also extracted and evaluated from public websites. All FMD sequence data were obtained from different body sites of samples of healthy adults ($\geq$18 years). On the data statistics page, a comprehensive list of available data like tests, counts, and metadata is maintained and updated regularly.

25.3.3.4 How Was the Data Analyzed?

For each sequence, the FMD pre-calculates the taxonomic population distribution from the public dataset using the UPARSE pipeline (Edgar 2013) and then maps these populations to their geographic position (discreetly, not continuously) using machine learning strategies to classify bacterial taxa (at different taxonomic levels)

that better differentiate across different geographic locations. These two measures are outlined in the manual tab available on the FMD website.

25.3.3.5 What Data Should I Upload into the Database?

In addition, the FMD will take 16S rRNA sequence data in FASTQ format (see SOP on website). The website will currently use user-supplied Mothur formatted taxonomy and OTU files as inputs to display the taxonomic structure of the data, equate it with current FMD data, and estimate the geographic origin of the data via the FMD analysis tab.

25.3.3.6 Why Does the FMD Pipeline Use UPARSE, and Not Mothur or Qiime?

Various programs will predict a 16S sample taxonomic structure, like Mothur and Qiime. Because UPARSE utilizes Mothur throughout the taxonomic assignment phase, the software differ except during OTU construction. UPARSE was shown to be faster during benchmarking with less OTUs (Edgar 2013). This is largely attributed to the elimination of singletons by UPARSE (a singleton is a read with a sequence existing entirely once, i.e., is specific among reads) before OTU clustering (or OTU generation). Removing singletons eliminates sequencing errors. Moreover, UPARSE is an agnostic sequencing tool, allowing it impartial to interpret sequence datasets generated separately (i.e., 454 or Illumina). The manual on the website includes a comprehensive explanation of the UPARSE system.

25.3.3.7 What Metadata Is Available on the FMD?

The database includes metadata such as body location, topic age and gender, and geographical details such as region, district (i.e., state or department or province), and area. In certain applications, these parameters are undefined and called "NA." More and more details of the FMD is filled with, the more reliable would be the geolocation predictions.

25.3.4 Human Identification, Ethnicity Prediction, and Geographic Location Determination

Community composition within the human microbiome varies across individuals, but it remains unknown if this variation is sufficient to uniquely identify individuals within large populations or stable enough to identify them over time (Franzosa et al. 2015).

25.3.4.1 Identity

Recent large-scale human microbiome studies have shown considerable variation in body-site-specific community composition and bacterial organism activity among healthy individuals (Human and Project 2012; Qin et al. 2010; Cao et al. 2018). Furthermore, it was shown that human microbiome characteristics may be optimally correlated with individual people over significant phases of time (Schloissnig et al.

2012; Fierer et al. 2010; Poethig et al. 2013). Such findings indicate a special and secure distinction within a population dependent on their native microbiota. There has, however, been no systematic attempt to experimentally assess the viability of microbiome-based detection. To do so, it involves proving (1) that an individual-specific "metagenomic code" can be found in a population sample; (2) that perhaps the code can be strongly redetected at a later stage; (3) that the code is impossible to fit an unknown sample incorrectly; and (4) that such codes can be built for a significant fraction of individuals. Such criteria stress the nature of human microbiome identity with microbiome formation, composition, personalization, and transient stability—foundational topics in microbiome science ecological approaches (Franzosa et al. 2015).

In 2014, proof of concept was documented to distinguish two persons using oral micrbiome (Sarah 2014). Tests from individual and combination studies revealed that samples can be clustered from one individual and isolated from samples from a second individual (Leake et al. 2016). It had shown that salivary microflora displays substantial heterogeneity, and using a PCR-based metagenomic method using the two gene targets, respectively, 16s rRNA gene and rpoB2, it was possible to discriminate between two different individuals. In this analysis, a 58 genera central microbiome was postulated by integrating three targets. This high number of genera comprises about 95% of each individual population, suggesting that most variations come from species/strain level. Results showed that with this type of testing, the minimum number of sequences is 100,000 as this provided an excellent distinction between individuals with all targets.

Recently Neckovic et al. (2020) defined how well a distinguishable microbiome could be passed to another person and substrates, and vice versa. The findings of this pilot study introduce a variety of potential challenges for the use of microbiome sampling for forensic purposes; without the existence of comparison samples from an individual(s) obtained from the correct body location, with a detailed history of all individuals or artifacts closely touching the skin/body site within a currently undefined timeline, it would be different.

25.3.4.2 Geolocation

This high heterogeneity in the human salivary microbiota is not geographically organized. Although there is considerably more variation in bacterial genera relative to different people than from the same population, variability between persons from the same region is about the same as variation between individuals from different places. Generally, the interpopulation portion of the variance in the distribution of the bacterial genera is 13.5%, which means that each organism comprises on average 86.5% of the overall variance found when all organisms are collected. This is strikingly proportionate to the amount of interpopulation variability usually found among humans for neutral genetic markers (Romualdi et al. 2002; Li et al. 2007). Thus, in this context, genera distribution of individuals acts as a neutral genetic marker if one equates individual bacterial composition with the genetic composition of human populations (Nasidze et al. 2009a, b).

25.3.4.3 Ethnicity

Mason et al. (2013) found that the subgingival microbial fingerprint can successfully discriminate between the four ethnicities. To do this, a Random Forest machine-learning classifier was trained to develop an educated classification algorithm using subgingival microbial signatures, which was then applied to a test dataset to examine the accuracy, sensitivity, and specificity of the prediction. The subgingival microbial community was able to predict an individual's ethnicity with a 62% accuracy, 58% sensitivity, and 86% specificity.

25.4 Concluding Remarks

The human microbiota contains fungi, bacteria, and viruses, which live in and around the body. As evidence-based tool for the correlation or exclusion of people of concern linked to illegal activity, microbiomes have intangible value in forensics. Work has shown that the microbiome is isolated and specific signatures from textures such as tablets, shoes, and textiles can be retrieved. In order to explore the effectiveness and possible shortcomings of the micronutrient profiling, further research is necessary before the human microbiome in general and salivary microbiome in particular are used as a research tool. This involves the detection range, risks involved with, or the application of microbial profiling for forensic applications, of microbial transfers among humans or objects.

References

Aaspõllu A, Lillsaar T, Tummeleht L, Simm J, Metsis M (2011) Can microbes on skin help linking persons and crimes? Forensic Sci Int Genet Suppl Ser 3:269–270. https://doi.org/10.1016/j.fsigss.2011.08.131

Adams RI, Bateman AC, Bik HM, Meadow JF (2015) Microbiota of the indoor environment: a meta-analysis. Microbiome 3:49. https://doi.org/10.1186/s40168-015-0108-3

Adler CJ, Dobney K, Weyrich LS, Kaidonis J, Walker AW, Haak W, Bradshaw CJA, Townsend G, Sołtysiak A, Alt KW, Parkhill J, Cooper A (2013) Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions. Nat Genet 45:450–455. https://doi.org/10.1038/ng.2536

Adserias-Garriga J, Quijada NM, Hernandez M, Rodríguez Lázaro D, Steadman D, Garcia-Gil LJ (2017) Dynamics of the oral microbiota as a tool to estimate time since death. Mol Oral Microbiol 32:511–516. https://doi.org/10.1111/omi.12191

An JY, Darveau R, Kaeberlein M (2018) Oral health in geroscience: animal models and the aging oral cavity. Geroscience 40:1–10. https://doi.org/10.1007/s11357-017-0004-9

Anukam K, Agbakoba N (2017) A comparative study of the oral microbiome compositions of healthy postmenopausal, premenopausal, and prepubertal Nigerian females, using 16s rrna metagenomics methods. Niger J Clin Pract 20:1250–1258. https://doi.org/10.4103/njcp.njcp_32_17

Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, Fernandes GR, Tap J, Bruls T, Batto JM, Bertalan M, Borruel N, Casellas F, Fernandez L, Gautier L, Hansen T, Hattori M, Hayashi T, Kleerebezem M, Kurokawa K, Leclerc M, Levenez F, Manichanh C, Nielsen HB, Nielsen T, Pons N, Poulain J, Qin J, Sicheritz-Ponten T, Tims S, Torrents D,

Ugarte E, Zoetendal EG, Wang J, Guarner F, Pedersen O, De Vos WM, Brunak S, Doré J, Weissenbach J, Ehrlich SD, Bork P (2011) Enterotypes of the human gut microbiome. Nature 473:174–180. https://doi.org/10.1038/nature09944

Bik EM, Long CD, Armitage GC, Loomer P, Emerson J, Mongodin EF, Nelson KE, Gill SR, Fraser-Liggett CM, Relman DA (2010) Bacterial diversity in the oral cavity of 10 healthy individuals. ISME J 4:962–974. https://doi.org/10.1038/ismej.2010.30

Blaser MJ (2010) Harnessing the power of the human microbiome. Proc Natl Acad Sci U S A 107:6125–6126. https://doi.org/10.1073/pnas.1002112107

Brito IL, Yilmaz S, Huang K, Xu L, Jupiter SD, Jenkins AP, Naisilisili W, Tamminen M, Smillie CS, Wortman JR, Birren BW, Xavier RJ, Blainey PC, Singh AK, Gevers D, Alm EJ (2016) Mobile genes in the human microbiome are structured from global to individual scales. Nature 535:435–439. https://doi.org/10.1038/nature18927

Brown LR, Fromme WJ, Handler SF, Wheatcroft MG, Johnston DA (1976) Effect of Skylab missions on clinical and microbiologic aspects of oral health. J Am Dent Assoc 93:357–363. https://doi.org/10.14219/jada.archive.1976.0502

Burcham ZM, Hood JA, Pechal JL, Krausz KL, Bose JL, Schmidt CJ, Benbow ME, Jordan HR (2016) Fluorescently labeled bacteria provide insight on post-mortem microbial transmigration. Forensic Sci Int 264:63–69. https://doi.org/10.1016/j.forsciint.2016.03.019

Butler JM (2011) Forensic DNA testing. Cold Spring Harb Protoc 6(12):1438–1450. https://doi.org/10.1101/pdb.top066928

Cao S, Bai P, Zhu W, Chen D, Wang H, Jin B, Zhang L, Liang W (2018) Genetic portrait of 27 Y-STR loci in the Tibetan ethnic population of the Qinghai province of China. Forensic Sci Int Genet 34:e18–e19. https://doi.org/10.1016/j.fsigen.2018.02.005

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, Mcdonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R (2010) QIIME allows analysis of high- throughput community sequencing data. Intensity normalization improves color calling in SOLiD sequencing. Nat Publ Gr 7:335–336. https://doi.org/10.1038/nmeth0510-335

Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R (2009) Bacterial community variation in human body habitats across space and time. Science 326:1694–1697. https://doi.org/10.1126/science.1177486

Damann FE, Williams DE, Layton AC (2015) Potential use of bacterial community succession in decaying human bone for estimating postmortem interval. J Forensic Sci 60:844–850. https://doi.org/10.1111/1556-4029.12744

Davenport ER (2017) Tooth be told, genetics influences oral microbiome. Cell Host Microbe 22:251–253. https://doi.org/10.1016/j.chom.2017.08.018

DeBruyn JM, Hauther KA (2017) Postmortem succession of gut microbial communities in deceased human subjects. PeerJ 2017:1–14. https://doi.org/10.7717/peerj.3437

Do T, Gilbert SC, Clark D, Ali F, Fatturi Parolo CC, Maltz M, Russell RR, Holbrook P, Wade WG, Beighton D (2010) Generation of diversity in streptococcus mutans genes demonstrated by MLST. PLoS One 5. https://doi.org/10.1371/journal.pone.0009073

Edgar RC (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat Methods 10:996–998. https://doi.org/10.1038/nmeth.2604

Fierer N, Hamady M, Lauber CL, Knight R (2008) The influence of sex, handedness, and washing on the diversity of hand surface bacteria. Proc Natl Acad Sci U S A 105:17994–17999. https://doi.org/10.1073/pnas.0807920105

Fierer N, Lauber CL, Zhou N, McDonald D, Costello EK, Knight R (2010) Forensic identification using skin bacterial communities. Proc Natl Acad Sci U S A 107:6477–6481. https://doi.org/10.1073/pnas.1000162107

Flores GE, Caporaso JG, Henley JB, Rideout JR, Domogala D, Chase J, Leff JW, Vázquez-Baeza Y, Gonzalez A, Knight R, Dunn RR, Fierer N (2014) Temporal variability is a

personalized feature of the human microbiome. Genome Biol 15:531. https://doi.org/10.1186/s13059-014-0531-y

Franzosa EA, Huang K, Meadow JF, Gevers D, Lemon KP, Bohannan BJM, Huttenhower C (2015) Identifying personal microbiomes using metagenomic codes. Proc Natl Acad Sci U S A 112: E2930–E2938. https://doi.org/10.1073/pnas.1423854112

Gajer P, Brotman RM, Bai G, Sakamoto J, Schütte UME, Zhong X, Koenig SSK, Fu L, Ma Z, Zhou X, Abdo Z, Forney LJ, Ravel J (2012) Temporal dynamics of the human vaginal microbiota. Sci Transl Med 4. https://doi.org/10.1126/scitranslmed.3003605

Galvão-Moreira LV, de Andrade CM, de Oliveira JFF, Bomfim MRQ, Figueiredo P d MS, Branco-de-Almeida LS (2018) Sex differences in salivary parameters of caries susceptibility in healthy individuals. Oral Heal Prev Dent 16:71–77. https://doi.org/10.3290/j.ohpd.a39684

Gao Z, Tseng CH, Pei Z, Blaser MJ (2007) Molecular analysis of human forearm superficial skin bacterial biota. Proc Natl Acad Sci U S A 104:2927–2932. https://doi.org/10.1073/pnas.0607077104

Giampaoli S, Berti A, Valeriani F, Gianfranceschi G, Piccolella A, Buggiotti L, Rapone C, Valentini A, Ripani L, Romano Spica V (2012) Molecular identification of vaginal fluid by microbial signature. Forensic Sci Int Genet 6:559–564. https://doi.org/10.1016/j.fsigen.2012.01.005

Goebel BM, Stackebrandt E (1994) Cultural and phylogenetic analysis of mixed microbial populations found in natural and commercial bioleaching environments. Appl Environ Microbiol 60:1614–1621. https://doi.org/10.1128/aem.60.5.1614-1621.1994

Gomez A, Espinoza JL, Harkins DM, Leong P, Saffery R, Bockmann M, Torralba M, Kuelbs C, Kodukula R, Inman J, Hughes T, Craig JM, Highlander SK, Jones MB, Dupont CL, Nelson KE (2017) Host genetic control of the oral microbiome in health and disease. Cell Host Microbe 22:269–278.e3. https://doi.org/10.1016/j.chom.2017.08.013

Grice EA, Kong HH, Conlan S, Deming CB, Davis J, Young AC, Bouffard GG, Blakesley RW, Murray PR, Green ED, Turner ML, Segre JA (2009) Topographical and temporal diversity of the human skin microbiome. Science 324:1190–1192. https://doi.org/10.1126/science.1171700

Guo JJ, Liao HD, Fu XL, Zha L, Liu JS, Cai JF (2015) Bacterial community succession analysis by next generation sequencing in Changsha city, China. Forensic Sci Int Genet Suppl Ser 5:e107–e108. https://doi.org/10.1016/j.fsigss.2015.09.043

Hanssen EN, Avershina E, Rudi K, Gill P, Snipen L (2017) Body fluid prediction from microbial patterns for forensic application. Forensic Sci Int Genet 30:10–17. https://doi.org/10.1016/j.fsigen.2017.05.009

Hsu L, Power D, Upritchard J, Burton J, Friedlander R, Horswell J, MacDonald C, Kieser J, Tompkins G (2012) Amplification of oral streptococcal DNA from human incisors and bite marks. Curr Microbiol 65:207–211. https://doi.org/10.1007/s00284-012-0148-x

http://www.homd.org. Last accessed 14 Aug 2020

Human T, Project M (2012) Structure, function and diversity of the healthy human microbiome. Nature 486:207–214. https://doi.org/10.1038/nature11234

Hyde ER, Haarmann DP, Lynne AM, Bucheli SR, Petrosino JF (2013) The living dead: bacterial community structure of a cadaver at the onset and end of the bloat stage of decomposition. PLoS One 8. https://doi.org/10.1371/journal.pone.0077733

Juusola J, Ballantyne J (2005) Multiplex mRNA profiling for the identification of body fluids. Forensic Sci Int 152:1–12. https://doi.org/10.1016/j.forsciint.2005.02.020

Kennedy D (2011) Forensic dentistry and microbial analysis of bite marks. APJ:6–15

Kianoush N, Adler CJ, Nguyen KAT, Browne GV, Simonian M, Hunter N (2014) Bacterial profile of dentine caries and the impact of pH on bacterial population diversity. PLoS One 9. https://doi.org/10.1371/journal.pone.0092940

Lassalle F, Spagnoletti M, Fumagalli M, Shaw L, Dyble M, Walker C, Thomas MG, Bamberg Migliano A, Balloux F (2018) Oral microbiomes from hunter-gatherers and traditional farmers reveal shifts in commensal balance and pathogen load linked to diet. Mol Ecol 27:182–195. https://doi.org/10.1111/mec.14435

Leake SL, Pagni M, Falquet L, Taroni F, Greub G (2016) The salivary microbiome for differentiating individuals: proof of principle. Microbes Infect 18:399–405. https://doi.org/10.1016/j.micinf.2016.03.011
Ley RE, Turnbaugh PJ, Klein S, Gordon JI (2006a) Human gut microbes associated with obesity. Nature 444:1021–1022. https://doi.org/10.1038/nature4441022a
Ley RE, Peterson DA, Gordon JI (2006b) Ecological and evolutionary forces shaping microbial diversity in the human intestine. Cell 124:837–848. https://doi.org/10.1016/j.cell.2006.02.017
Li Y, Ismail AI, Ge Y, Tellez M, Sohn W (2007) Similarity of bacterial populations in saliva from African-American mother-child dyads. J Clin Microbiol 45:3082–3085. https://doi.org/10.1128/JCM.00771-07
Mason MR, Nagaraja HN, Camerlengo T, Joshi V, Kumar PS (2013) Deep sequencing identifies ethnicity-specific bacterial signatures in the oral microbiome. PLoS One 8:1–7. https://doi.org/10.1371/journal.pone.0077287
Metcalf JL (2019) Estimating the postmortem interval using microbes: knowledge gaps and a path to technology adoption. Forensic Sci Int Genet 38:211–218. https://doi.org/10.1016/j.fsigen.2018.11.004
Metcalf JL, Wegener Parfrey L, Gonzalez A, Lauber CL, Knights D, Ackermann G, Humphrey GC, Gebert MJ, Van Treuren W, Berg-Lyons D, Keepers K, Guo Y, Bullard J, Fierer N, Carter DO, Knight R (2013) A microbial clock provides an accurate estimate of the postmortem interval in a mouse model system. elife 2:1–19. https://doi.org/10.7554/elife.01104
Muhammad S, Ali F, Tanwir F (2012) Oral microbial habitat a dynamic entity. J Oral Biol Craniofacial Res 2:181–187. https://doi.org/10.1016/j.jobcr.2012.07.001
Nakanishi H, Kido A, Ohmori T, Takada A, Hara M, Adachi N, Saito K (2009) A novel method for the identification of saliva by detecting oral streptococci using PCR. Forensic Sci Int 183:20–23. https://doi.org/10.1016/j.forsciint.2008.10.003
Nasidze I, Quinque D, Li J, Li M, Tang K, Stoneking M (2009a) Comparative analysis of human saliva microbiome diversity by barcoded pyrosequencing and cloning approaches. Anal Biochem 391:64–68. https://doi.org/10.1016/j.ab.2009.04.034
Nasidze I, Li J, Quinque D, Tang K, Stoneking M (2009b) Global diversity in the human salivary microbiome. Genome Res 19:636–643. https://doi.org/10.1101/gr.084616.108
Neckovic A, van Oorschot RAH, Szkuta B, Durdle A (2020) Investigation of direct and indirect transfer of microbiomes between individuals. Forensic Sci Int Genet 45:102212. https://doi.org/10.1016/j.fsigen.2019.102212
Ogawa T, Hirose Y, Honda-Ogawa M, Sugimoto M, Sasaki S, Kibi M, Kawabata S, Ikebe K, Maeda Y (2018) Composition of salivary microbiota in elderly subjects. Sci Rep 8:1–8. https://doi.org/10.1038/s41598-017-18677-0
Oliveira M, Amorim A (2018) Microbial forensics: new breakthroughs and future prospects. Appl Microbiol Biotechnol 102:10377–10391. https://doi.org/10.1007/s00253-018-9414-6
Peterson J, Garges S, Giovanni M, McInnes P, Wang L, Schloss JA, Bonazzi V, McEwen JE, Wetterstrand KA, Deal C, Baker CC, Di Francesco V, Howcroft TK, Karp RW, Lunsford RD, Wellington CR, Belachew T, Wright M, Giblin C, David H, Mills M, Salomon R, Mullins C, Akolkar B, Begg L, Davis C, Grandison L, Humble M, Khalsa J, Roger Little A, Peavy H, Pontzer C, Portnoy M, Sayre MH, Starke-Reed P, Zakhari S, Read J, Watson B, Guyer M (2009) The NIH human microbiome project. Genome Res 19:2317–2323. https://doi.org/10.1101/gr.096651.109
Poethig RS, Faith JJ, Guruge JL, Charbonneau M, Subramanian S, Seedorf H, Goodman AL, Clemente JC, Knight R, Andrew C, Leibel RL, Rosenbaum M, Gordon JI (2013) The long-term stability of the human gut microbiota. Science 341:1–22. https://doi.org/10.1126/science.1237439.The
Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T, Mende DR, Li J, Xu J, Li S, Li D, Cao J, Wang B, Liang H, Zheng H, Xie Y, Tap J, Lepage P, Bertalan M, Batto JM, Hansen T, Le Paslier D, Linneberg A, Nielsen HB, Pelletier E, Renault P, Sicheritz-Ponten T, Turner K, Zhu H, Yu C, Li S, Jian M, Zhou Y, Li Y, Zhang X,

Li S, Qin N, Yang H, Wang J, Brunak S, Doré J, Guarner F, Kristiansen K, Pedersen O, Parkhill J, Weissenbach J, Bork P, Ehrlich SD, Wang J, Antolin M, Artiguenave F, Blottiere H, Borruel N, Bruls T, Casellas F, Chervaux C, Cultrone A, Delorme C, Denariaz G, Dervyn R, Forte M, Friss C, Van De Guchte M, Guedon E, Haimet F, Jamet A, Juste C, Kaci G, Kleerebezem M, Knol J, Kristensen M, Layec S, Le Roux K, Leclerc M, Maguin E, Minardi RM, Oozeer R, Rescigno M, Sanchez N, Tims S, Torrejon T, Varela E, De Vos W, Winogradsky Y, Zoetendal E (2010) A human gut microbial gene catalogue established by metagenomic sequencing. Nature 464:59–65. https://doi.org/10.1038/nature08821

Rahimi M, Heng NCK, Kieser JA, Tompkins GR (2005) Genotypic comparison of bacteria recovered from human bite marks and teeth using arbitrarily primed PCR. J Appl Microbiol 99:1265–1270. https://doi.org/10.1111/j.1365-2672.2005.02703.x

Raman R, Thomas RG, Weiner MW, Jack CR, Ernstrom K, Aisen PS, Tariot PN, Quinn JF (2010) Diversity of the human intestinal microbial flora. Science 23:333–336. https://doi.org/10.1097/WAD.0b013e3181aba588.MRI

Romualdi C, Balding D, Nasidze IS, Risch G, Robichaux M, Sherry ST, Stoneking M, Batzer MA, Barbujani G (2002) Patterns of human diversity, within and among continents, inferred from biallelic DNA polymorphisms. Genome Res 12:602–612. https://doi.org/10.1101/gr.214902

Sarah L (2014) Human identification through analysis of the salivary microbiome: proof of principle. Universite de Lausanne, Lausanne

Schloissnig S, Arumugam M, Sunagawa S, Mitreva M, Tap J, Zhu A (2012) Genomic variation landscape of the human gut microbiome. Nature 493:45. https://doi.org/10.1038/nature11711

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75:7537–7541. https://doi.org/10.1128/AEM.01541-09

Sjödin A, Broman T, Melefors Ö, Andersson G, Rasmusson B, Knutsson R, Forsman M (2013) The need for high-quality whole-genome sequence databases in microbial forensics. Biosecur Bioterror 11. https://doi.org/10.1089/bsp.2013.0007

Spradbery P (2010) Restriction fragment length polymorphisms of mutans streptococci in forensic odontological analysis. Biosci Horizons 3:166–178. https://doi.org/10.1093/biohorizons/hzq021

Sze MA, Schloss PD (2016) Looking for a signal in the noise: revisiting obesity and the microbiome. MBio 7:1–10. https://doi.org/10.1128/mBio.01018-16

Takeshita T, Kageyama S, Furuta M, Tsuboi H, Takeuchi K, Shibata Y, Shimazaki Y, Akifusa S, Ninomiya T, Kiyohara Y, Yamashita Y (2016) Bacterial diversity in saliva and oral health-related conditions: the Hisayama study. Sci Rep 6:1–11. https://doi.org/10.1038/srep22164

Talan DA, Abrahamian FM, Moran GJ, Citron DM, Tan JO, Goldstein EJC (2003) Clinical presentation and bacteriologic analysis of infected human bites in patients presenting to emergency departments. Clin Infect Dis 37:1481–1489. https://doi.org/10.1086/379331

Tridico SR, Murray DC, Addison J, Kirkbride KP, Bunce M (2014) Metagenomic analyses of bacteria on human hairs: a qualitative assessment for applications in forensic science. Investig Genet 5. https://doi.org/10.1186/s13323-014-0016-5

Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI (2007) The human microbiome project. Nature 449:804–810. https://doi.org/10.1038/nature06244

Ventura Spagnolo E, Stassi C, Mondello C, Zerbo S, Milone L, Argo A (2019) Forensic microbiology applications: a systematic review. Legal Med 36:73–80. https://doi.org/10.1016/j.legalmed.2018.11.002

Xu H, Hao W, Zhou Q, Wang W, Xia Z, Liu C, Chen X, Qin M, Chen F (2014) Plaque bacterial microbiome diversity in children younger than 30 months with or without caries prior to eruption of second primary molars. PLoS One 9:1–10. https://doi.org/10.1371/journal.pone.0089269

Zhernakova A, Kurilshikov A, Bonder MJ, Tigchelaar EF, Schirmer M, Vatanen T, Mujagic Z, Vila AV, Falony G, Vieira-Silva S, Wang J, Imhann F, Brandsma E, Jankipersadsing SA, Joossens M, Cenit MC, Deelen P, Swertz MA, Weersma RK, Feskens EJM, Netea MG, Gevers D, Jonkers D, Franke L, Aulchenko YS, Huttenhower C, Raes J, Hofker MH, Xavier RJ, Wijmenga C, Fu J (2016) Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. Science 352:565–569. https://doi.org/10.1126/science.aad3369

Zhou J, Jiang N, Wang Z, Li L, Zhang J, Ma R, Nie H, Li Z (2017) Influences of pH and iron concentration on the salivary microbiome in individual humans with and without caries. Appl Environ Microbiol 83:1–12. https://doi.org/10.1128/AEM.02412-16

MALDITOF the Fourth Generation Techniques Still at Its Infancy to Identify Forensically Important Insects

26

Vandana Vinayak and Anshuman Rai

Abstract

Forensic entomology plays an important role in death investigations. The morphological identification of forensically important insects (FII) recovered from the scene of crime is important to help estimate time since death. It is seen that many times insects are in their immature stages which are indistinguishable. The molecular techniques come here as golden standards to barcode the genetic code of such FII and identify them at species level targeting mainly on *NADH dehydrogenase 1* to *5 (ND 1–5)*, *cytochrome b* and *c oxidase* (*COI*) mitochondrial gene. However, at times species are so closely related that dissimilar species are grouped together. This often ends with their incorrect identification both morphologically and at molecular level. Alternatively, Matrix assisted laser desorption/ionization time of flight mass spectrophotometry (MALDI TOF MS) techniques based on the proteins content in a given species, help in discriminating closely related species along with deciphering their age. Studies have shown that MALDI TOF MS discriminates morphologically, biochemically, and genetically similar species at higher degree of confidence. Not the least, they are economical, quick and do not require prior reference data, knowledge or expertise in taxon identification. The technique is although not new but it is emerging a little slowly to identify forensically important insects for correct estimation of post mortem interval.

V. Vinayak (✉)
Diatom Nano Engineering and Metabolism (DNM) Lab., Department of Criminology and Forensic Science, School of Applied Sciences, Dr. Harisingh Gour Central University, Sagar, Madhya Pradesh, India
e-mail: vvinayak@dhsgsu.co.in

A. Rai
DNA Division, State Forensic Science Laboratory, Madhuban, Haryana, India

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_26

Keywords

Forensically important insects (FII) · Matrix assisted laser desorption/ionization time of flight mass spectrophotometry (MALDI TOF MS) · Proteins · Identification of species and age · Minimum post mortem interval (PMI_{min})

Abbreviations

(Q-TOF) MS	Quadrupole orthogonal acceleration time of flight
AmTRP	*Apis mellifera* tachykinin-related peptide
bp	Base pair
CBOL	Consortium for barcode of life
CHC	Cuticular hydrocarbons
COI,II,III	Cytochrome C Oxidase Subunit I,II,III
cytb	Cytochrome b
DNA	Deoxyribonucleic acid
Ds	Dice similarity
FII	Forensically important insects
GAM	Generalized additive model
Ile	Isoleucine
ITS	Internal Transcribed Spacer
kDa	Kilo Daltons
Leu	Leucine
LSV	Log score value
MALDI-TOF-MS	Matrix assisted laser desorption/ionization time of flight mass spectrophotometry
mt COI	Mitochondrial Cytochrome C Oxidase Subunit I
Mt-DNA	Mitochondrial Deoxyribonucleic acid DNA
NCBI	National centre of Biotechnology and Information
NU-DNA	Nuclear deoxyribonucleic acid
PCR-RFLP	Polymerase chain reaction restriction fragment length polymorphism
PMI_{min}	Minimum post mortem interval
PVK/CAP2b	Periviscerokinin/cardioacceleratory peptide 2b
rDNA	Ribosomal deoxyribonucleic acid
Real-time PCR	Real time polymerase chain reaction
rRNA	Ribosomal ribonucleic acid

26.1 Introduction

Forensically important insects (FII) belong to families Calliphoridae, Sarcophagidae, Muscidae, Sepsidae, Sphaeroceridae, Piophilidae, and Phoridae (Payne 1965; Goff 1993; Tantawi et al. 1996). They are not just limited to find minimum post mortem

interval (PMI_{min}); but also play a very important role in determining; cause of death, geographical location of death, movement or storage of remains after death, time of decapitation and dismemberment, submersion interval, specific site of injury on the body, toxic drugs in body, etc. (Amendt et al. 2004). Soon after the death, decomposition of the body starts which attracts various FII at the open wounds and natural orifices. These sites serve as the spectacular breeding place for these necrophageous insects (Amendt et al. 2011; Gennard 2012). Among FII, flies belonging to families Calliphoridae (the blow flies) and Sarcophagidae (the flesh flies) are among the first ones to arrive on a corpse within a few minutes (Payne 1965; Tullis and Goff 1987; Anderson and VanLaerhoven 1996; Dadour et al. 2001; Wells et al. 2001; Greenberg and Kunich 2002; Byrd and Castner 2009) and even few seconds (De Jong 1995). These first cadaver colonizers are attracted by volatile organic chemicals (VOC) released by the cadavers (Ashworth and Wall 1994; Morris et al. 1998; Wells and LaMotte 2001; Hart and Whitaker 2006; Statheropoulos et al. 2007; Dekeirsschieter et al. 2009).The arrival of necrophageous insects on the cadaver starts by laying of their eggs and starting their reproductive life cycle. Since every insect has a life cycle period during which it develops into a complete adult from the egg, taking few days to weeks. The collection of insects at their different stages of development by forensic scientist and knowing their life cycle helps in estimating the time since death of a particular body. Identification of these insects on or around the cadaver is therefore of paramount importance. Although recovery of these insects at their various stages is helpful they show variation in their development which in turn is highly dependent upon the environmental conditions (Berg and Benbow 2013). The bodies found in water, plains, summers, winters, ice, soil, coffin, indoors, caskets, closed vehicles, drugs, rural and urban areas have different impact on the colonization of different insects and their stages (Campobasso et al. 2001; Pohjoismäki et al. 2010; Griffiths et al. 2020). Also if carrion is toxicated with some drugs like lead, cocaine, heroin, morphine, paracetamol it effects not only the type of insect which arrive but also show variation sin their developments (Carvalho et al. 2001; Aneyo et al. 2020). Of the total ten cases studied by (Al-Khalifa et al. 2020), it was found that *Dermestes masculatus* was identified on human cadavers which were in advanced stage of decomposition (Al-Qahtni et al. 2020). *Dermes frischii* was observed only in indoor skeletonised human cadaver. A study showed that blow fly larvae develop much slowly in indoor cadavers as compared to outdoor ones. This definitely results in approaching PMI_{min} much earlier than the actual (Hofer et al. 2017). However, the identification of some insect's species, especially in their immature stages may be complicated by many factors (Ingrisch 1995; Cywinska et al. 2006; Rindi et al. 2008; Packer et al. 2009), even for experienced taxonomists (Nekola and Barthel 2002; Gutiérrez-Gutiérrez et al. 2013). It is thus very necessary to establish the identity of the particular species to determine post mortem interval.

Although every colonizer on a cadaver has a different morphological life cycle. However, it's difficult to distinguish members of closely related species. Figure 26.1 shows similarities in eggs of four different species of Calliphoridae (*Lucilia sericata*, *Phormia regina*, *Lucilia coeruleiviridis* and *Calliphora vicina*) (Giffen et al. 2017). The figure depict that although there is much difference in the morphological look of

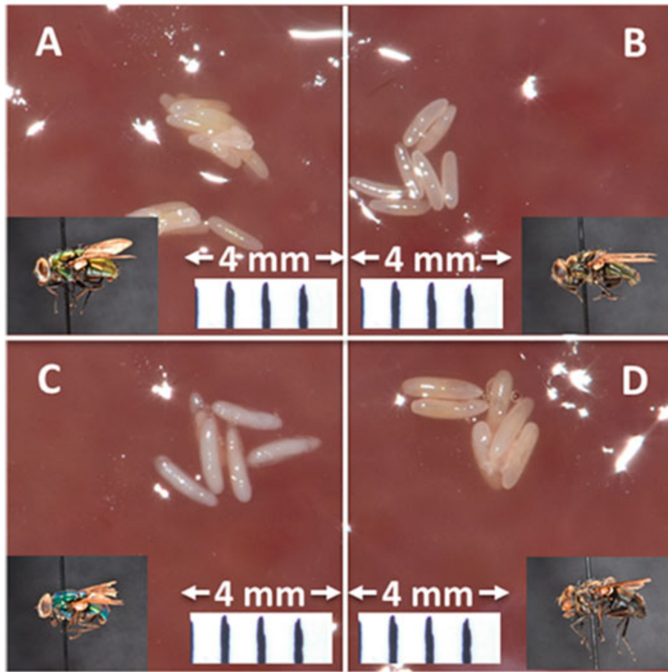


Fig. 26.1 Images of eggs derived from four blow fly species from the Calliphoridae family, with the inset in each case showing a picture of the adult fly species corresponding to the eggs displayed in each panel. (**a**) *Lucilia sericata* eggs and adult fly; (**b**) *Phormia regina* eggs and adult fly; (**c**) *L. coeruleiviridis* eggs and adult fly; (**d**) *Calliphora vicina* eggs and adult fly. The scale provided refers to egg images only and does not reflect adult size. (Reproduced with permission from Giffen et al. (2017))

adult flies their eggs look quite similar. Thus it becomes necessary to rear the eggs back in lab and to be careful in transporting them in their live stage. This is also very important for forensic laboratories to have an expertise, lab set up and conditions for rearing and identifying insects. However, most of the times insects collected at different stages by the scene of crime and police unit fix the stages with ethanol or boiling water for their easy transport. This not only destroys their true morphology but also alters their actual weight and size which creates hindrance to their identification. Besides this the environmental and physical conditions of the corpse may alter the morphological characters in the larvae or egg invading it. Even though with the use of state of art microscopy like Scanning electron microscopy it becomes simple to identify different stages but in cases of closely related species it becomes challenging.

On the other hand, due to enormous number of species and diversity present in the invertebrates DNA based methods are considered to be most accurate and authentic (Hebert et al. 2003, 2004; Gupta and Setia 2004). The molecular barcoding to identify insects targets on specific mitochondrial cytochrome c oxidase (*COI-III*),

Cytochrome b oxidase (Cytb) and nitrogenous dehydrogenase (*ND1,2,4,5*) and DNA markers (*16S rDNA*) (Metcalf et al. 2013; Donaldson and Lamont 2014). DNA barcoding using mitochondrial gene is most common method of identifying the taxon (Wells and Stevens 2008; Amendt et al. 2011). Since mitochondrial DNA barcoding is beneficial over nuclear DNA over two three reasons. First COI gene is easy to amplify and second they are present in all cells and approximating 3/4th the size than that of nuclear DNA Third they have high rate of evolution and hence high rate of resolution (Ratnasingham and Hebert 2007). In addition because of its strictly maternal inheritance and no genetic recombination, mt DNA haplotype is a good candidate for evolutionary and population genetics studies (Sperling et al. 1994; Benecke and Wells 2001; Xinghua et al. 2010). The 600 bp locus of cytochrome oxidase gene is best known for targeting any tissue for taxon identification. The technique is definitely superior to morphological identification as it can help in identifying an insect at a stage which maybe deteriorated or even decomposed.

The consortium for the barcode of life-CBOL (www.barcodeoflife.org) and national centre for biotechnology and information (www.ncbi.nlm.nih.gov) encodes DNA sequences of insects like fruit flies, mosquitoes, bees (Order-Diptera) and that of Trichoptera and Lepidoptera which serves as reference spectra for blind tests (Federhen 2011; Boykin et al. 2012). The research of molecular identification of dipterans and coleopterons confirm the usefulness of molecular markers in differentiating and identifying the insects of forensic importance. This represents one more step toward a detailed investigation for the molecular identification of Dipterans and Coleopterons. However, it is possible only when nucleotide sequence information is available in the existing databases (Mizrachi 2007). Molecular methods are indeed golden standards for species identification due to their reproducibility and sensitivity (Ratnasingham and Hebert 2007).

However, it is seen that there exists variations in the mitochondrial DNA (mt-DNA) profiling obtained due to inclusion of bacteria associated with decomposition either of insect stage or the cadaver in which it's manifested (Whitworth et al. 2007). This therefore results in very low species identification. It is seen that bacteria *Wolbachia cosegregate* with mitochondria results into many bacterial species reading same DNA framework as that of the targeted insect species. This would therefore require huge DNA barcode to identify the species of interest. *Wolbachia* can cause mitochondrial introgression in closely related species for which barcoding was not possible (Whitworth et al. 2007). It happened between closely related members of *Drosophila* (Rousset and Solignac 1995; Ballard et al. 2002, 2007), among two species of *Acraea encedon* and *Acraea encedana* (Jiggins 2002) and between yellow and brown type of *Eurema hecabe* (Narita et al. 2006). Whitworth et al. (2007) investigated 12 species of blow fly *Protocalliphora* infected by *Wolbachia*. They found it was impossible to barcode 60% of the unknown species and it was impossible to identify about 75% of the new species in the genus. The low species identification along with similar barcodes among the four different species was challenging at the species level if not at the genus level.

In other special case it is seen that in conditions like myiasis where insects larva live in living or dead host tissue or body fluid or ingested food, the exact time of arrival of insects in a carrion is a matter of concern (Anderson 2001). It is seen that in cases of myiasis the morphological identification of insects is difficult (Anderson 2001; Noël et al. 2004). This is mainly due to the reason that the vectors are in their immature, incomplete or destructive stage of their life cycle. Molecular sequencing to identify the myiasis infested insect species is done targeting genes like *12SDNA* (Norris et al. 1996), *18SrDNA* (Mangold et al. 1997) and *16S DNA* (Norris et al. 1996). To add that DNA barcoding of the insects requires specific primers, intensive labour, is time consuming and a huge database library is required (Wells and Sperling 2001).

On the other hand protein profiling using the fourth generation matrix assisted laser desorption/ionization time of flight mass spectrophotometery (MALDI-TOF-MS) is an important tool for identification and establishing the phylogeny of a species (Sauer et al. 2008; Freiwald and Sauer 2009; Seng et al. 2010). The evaluation of MALDITOF MS analysed bacterial species showed that the results were in consistent to that obtained when using 16S rRNA sequencing (Mellmann et al. 2008). The technique serves as a species classification tool comparable to DNA sequencing and has successfully tested to identify Eukarya, Archaea (Dridi et al. 2012) and some giant viruses (La Scola et al. 2010).

The current review however, emphasizes on adapting modern and economical techniques like matrix assisted laser desorption/ionization time of flight mass spectrophotometery (MALDI-TOF-MS) for correct identification of insect species up to subspecies level at rapid and economical rates. Importantly to identify the unknown species by MALDI TOF MS no taxon reference data is required since the protein profile pattern need not to be established for unknown species (Singhal et al. 2015).

In fact when compared to PCR or other antibody based assays it does not require genetic information of the species sample under study. However, it is necessary to make database of such reference samples (Campbell 2005). The recent studies have shown increase in number of reports where combination of MALDI time of flight (TOF) MS has been widely used to identify the microorganism at genus, species and subspecies level (Singhal et al. 2015).

26.2 Matrix Assisted Laser Desorption/Ionization Time of Flight Mass Spectrophotometry (MALDI-TOF-MS)

Protein typing by matrix assisted laser desorption/ionization time of flight mass spectrophotometry (MALDI-TOF MS) is the most advanced tool for species identification and with time researchers are using more of the proteomic than of the genomic data to distinguish microorganisms (Claydon et al. 1996; Holland et al. 1996; Krishnamurthy et al. 1996). The technique involves subjecting the organism or tissue to acid based protein extraction followed by mixing with aromatic acid (Singhal et al. 2015). The samples are spotted onto target plate and upon drying and

crystallization of proteins and peptides spectra, the spectra is recorded by laser desorption ionization and time of flight MS (Feltens et al. 2010). The desorbtion and ionization of protein sample spotted on the target metal or silicon plate generates protonated ions. These ions are then move with certain velocity at some fixed potential splitting as per their mass to charge ratio. The time of flight analyzers measures these charged analytes before they reach the detector to produce a protein spectra. Studies have shown the possibility of using MS based methods to identify metazoan species such as insects as illustrated below.

26.2.1 Insects

The mass fingerprinting approach has differentiated insects of order Mantophasmatodea and those of Diptera but it is still emerging for FII (Feltens et al. 2010). Very little work has been done on identification of forensically important insects (Li et al. 2017). Table 26.1 tabulates the list of insects identified by MALDI TOF MS. Scientist have identified neuropeptides from central nervous system (CNS) specifically brain of flies like *Lucila caprina* (Rahman et al. 2013) and *Protophormia terraenovae* (Inosaki et al. 2010). In insects neuropeptides are messengers located in the neurohemal organs of CNS, mainly *corpora cardiac* and *corpora allata* (Rahman et al. 2013). The peptidome of these neurohemal organs is tagma specific as it differs from head, thorax and abdomen (Predel et al. 2005). However, the neuropepetides have less variability among different genera/species. It was seen that peptidome of CNS in *Lucila* was partially investigated and found to be remarkably similar with that of neuropeptides of other flies. The detection of neuropeptides does not give unique identification of a sub group or species. The analysis of cuticular hydrocarbons (CHC) in *Reticulitermes lucifugus*, *Reticulitermes (lucifugus) grassei*, *Reticulitermes santonensis*, *Reticulitermes virginicus*, *Coptotermes formosanus*, *Neotermes castaneus*, *Neotermes cubanus*, *Cryptotermes*

Table 26.1 Publication on analysis of MALDITOF MS on insects

Insects	Mass range (kDa)	Reference
Drosophila melanogaster	2–40	Campbell (2005)
Mantophasmatodea	0.8–3	Predel et al. (2005)
Aphid spp.	3–25	Perera et al. (2005)
Drosophila spp.	1.8–15	Feltens et al. (2010)
Culicoides spp.	2–30	Kaufmann et al. (2011)
Culicoides spp.	2–20	Steinmann et al. (2013)
Glossina spp.	2–20	Hoppenheit et al. (2013)
Phlebotomus spp.	2–25	Dvorak et al. (2014)
Flea spp.	2–20	Yssouf et al. (2015)
Mosquito species	2–20	Yssouf et al. (2014)
Culicoides spp.	2–20	Sambou et al. (2015)

Reproduced with permission from Ulrich et al. (2017)

declivis, *Neobellieria bullata*, *Periplaneta americana*, *Acromyrmex echinatior Atta colombica* by MALDI TOF MS showed distinctive patterns which on comparison with GC/MS confirmed that MALDITOF MS analyzed high molecular weight and some unknown hydrocarbons (Cvačka et al. 2006). The question related to identification of these insects by chemotaxonomy is still a challenge and hence though MALDITOF MS can do CHC analysis of individual insect but failed to explain their phylogenetic relationships. The technique was first time applied for the study of Periviscerokinin/cardioacceleratory peptide 2b (PVK/CAP2b) class of neuropeptides in *Musca domestica* and *Neobellieria bullata* (Nachman et al. 2005). PVK peptides are typical of neurosecretory neurons in abdominal ganglia until they are released (Nachman et al. 2005). The earlier MS analysis of the neuropeptides of these flies could not distinguish between isomers of same molecular masses, e.g. Isoleucine (Ile) and leucine (Leu) (Audsley and Weaver 2003; Clynen et al. 2003; Predel et al. 2004). This is essentially due to the reason that low energy of fragmentation for ion of interest resulting into poor cleavage of side chains to distinguish Ile and Leu. However, with the advent of MALDITOF MS not only the sample requirement is less but it has allowed collision induced fragmentation of primary ions of peptides that distinguished side chains of Leu and Ile (Medzihradszky et al. 2000; Macht et al. 2004; Nachman et al. 2005).

26.2.1.1 *Drosophila*

Among insects the MALDI TOF MS technique was first used to identify *Drosophila* species and is a advanced tool to identify inter-specific variations (Feltens et al. 2010). Campbell (2005) studied the MALDITOF MS spectra of three strains of *Drosophila melanogaster (Oregon R*, *Canton S* and *Harare)* and its three species viz*; D. simulans*, *D. mauritiana* and *D. yakuba*. He found that overall peaks of *Drosophila melanogaster* and that for three species of *Drosophila* were similar among the three species than between species as can be seen in Fig. 26.2a. To prove it he took *Drosophila* strains from different localities but from same species of

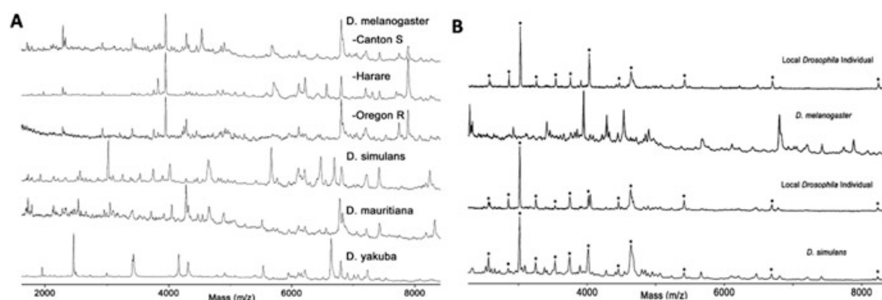


Fig. 26.2 (**a**) Spectra from five male *Drosophila* flies, three for *D. melanogaster* strains and one strain from *D. simulans*, *D. mauritiana* and *D. yakuba*. The vertical scale is expanded above *m*/*z* 5500 for clarity. (**b**) Identification of unknown species using reference spectra. The asterisks show the common spectral peaks shared by individuals of unknown *Drosophila* species and *D. simulans*. (Reproduced with permission from Campbell (2005))

D. melanogaster showed similarities and thus could be identified from closely related species (Campbell 2005). Campbell studied the MALDI TOF MS spectra of unknown *Drosophila* individuals and reference spectra of *D. melanogaster* and *D. simulans* to show that local *Drosophila* individual's spectra matched with that of MS spectra of *D. simulans* with least similarity with *D. melanogaster* (Campbell 2005). The concept was proved by crossing the males of unknown *Drosophila* individuals whose spectra was closely similar to *D. simulans* with females of *D. simulans* and *D. melanogaster* as seen in Fig. 26.2b. The progenies developed from the cross between unknown *Drosophila* individuals and *D. simulans* survived whereas no progenies were developed with *D. melanogaster* (Sawamura 2000). MALDI-TOF MS thus differentiated different *Drosophila* species by detecting single amino acid substitutions in neuropeptides (Powell and DeSalle 1995; Predel et al. 2005; Wegener and Gorbashov 2008).

Another work by Feltens et al. (2010) discriminated 128 *Drosophila* genera from 13 different species via MALDI TOF, and detected about 236 spectrum of proteins. The total protein profile spectra by MALDI TOF MS measurements show that among 128, three did not show linearity and were excluded. The phylogenetic relationship and a 7 cluster analysis of 125 spectra revealed further that 8 species of *Drosophila* (*D. novamexicana*, *D. virilis*, *D. hydei*, *D. erecta*, *D. melanogaster*, *D. yakuba* and *D. ananassae*) were clustered together and two species (*D. pseudobscura and D. miranda)* were excluded as seen in Fig. 26.3a. *D. funebris* however clustered with all spectra of *D. lummei* clearly differentiating these two species. Another species *D. teissieri* spectra was clustered with all spectra of *D. mauritiana*. The species *D. mauritiana* was found to be having highest spectral heterogeneity even after removal of two spectra and together with *D. ere* were found unsuitable for reference spectral study due to their high dissimilarity in the complete spectra generated. This dissimilarity index for these two species is measured by average Dice similarity coefficient (D_s) of 0.43, whereas D_s within groups of spectra were between 0.49 (*D. yak*) and 0.63 (*D. vir*). This low dissimilarity can be seen in the heat map where the white outlined box shows pair wise similarity indices in *D. mauritiana.* This is the first study demonstrating complete protein spectra from a complete multicellular organism much economically and quickly. Thus without any reference spectra Feltens et al. (2010) carried out cluster analysis of 125 species of *Drosophila*. Identification of *D. po* and *D. mir* at species level and at subspecies level for most of them was achieved. The spectra from *D. mau* clustered with two related but separate groups, one of which clustered with *D. tei*. This heterogeneity in the bootstrap is also visible in the heat map where the spectra from *D. mau* are outlined in poorly defined white square (Fig. 26.3b).

The MALDI-TOF-MS studies would thus serve a reliable and high throughput tool not only for species identification in *Drosophila* but also for forensically important insects to help determine time since death and post mortem interval. Besides *Drosophila* (Campbell 2005; Feltens et al. 2010), MALDI TOF MS has helped in identifying different arthropod groups like *Culicoides* (Kaufmann et al. 2012a), Ixodidae (Yssouf et al. 2013) and Culicidae species (Yssouf et al. 2013) using either whole organism or different body parts.

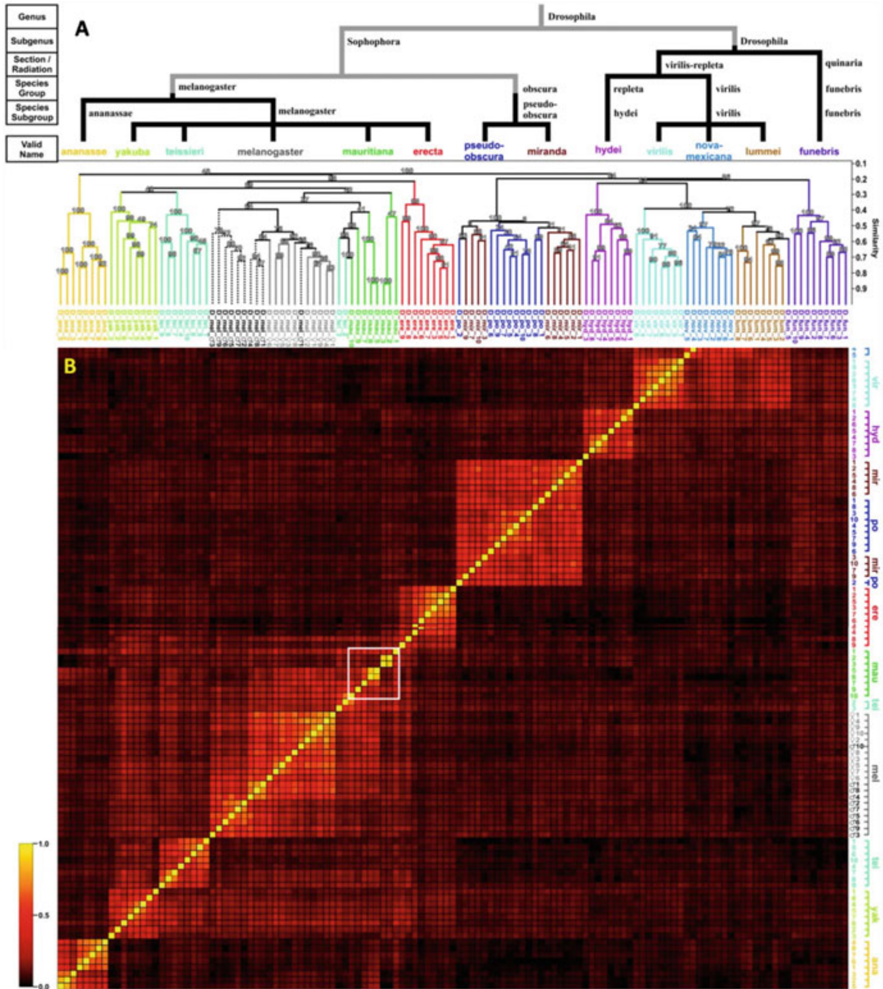


Fig. 26.3 (**a**) Dendrogram indicating the phylogeny of the different species is shown. Showing clustered species. Percentages of bootstrapping replicates supporting the location of individual nodes are indicated, Cluster analysis of 125 spectra from flies of 13 different *s*pecies clustered using Dice coefficients. (**b**) A heatmap representation of the pairwise similarity (Dice) coefficients in the lower part of the map shows White Square outlining the cluster of *D. mauritiana* spectra. (Copyright Feltens et al. (2010))

26.2.1.2 Aedes

The second major insect in which intensive protein profiling via MALDI TOF MS has been employed is group Aedes. The advent of increased mosquito related diseases has caused the technique like MALDI TOF to rapidly identify the mosquito species. Even though mosquitoes are not colonizers on a cadaver, but in order to understand and promote this technique in forensic science for species identification

it's necessary to understand the work done by researchers on these insects. It is important to know which body parts of these insects should be used in order to get consistent protein spectra via this technique. Kaufmann et al. (2011) showed that MALDI TOF MS spectra from body parts except abdomen in two species of *Culicoides* gave consistent results irrespective of their gender and age. Similarly in another study Kaufmann et al. (2012a) showed that protein profiles from legs, head and wings of the insect of 14 fresh specimen of *Culicoides* established their identification. It is necessary to standardise whether whole organism or its parts would be better to yield MALDI TOF protein spectra for phylogeny identification. Different researchers have used different body parts for the protein extraction to do MALDI TOF, however it is essential to know if decomposed or least amount of sample provides consistency in the results. In 2013 Yssouf et al. investigated the identification of some 129 mosquito samples and created a reference database of for some 20 species of which 4 were *Aede*s spp., 9 *Anopheles* spp., 4 *Culex* spp., *Lutzia tigripes*, *Orthopodomyia reunionensis* and *Mansonia uniformis* (Yssouf et al. 2013). Among these MALDI TOF MS was done from leg of about 95 mosquitoes to create a reference database. The analysis showed that protein spectral peaks fell in the range of 2–20 kDa and was consistent in all species which were tested as shown in Fig. 26.4. Further the flexi analysis and Clin Pro tools confirmed that the protein spectra in all mosquito species were identical within the species.

According to Yssouf et al. (2013) the threshold log score value (LSV) of 1.8 was relevant for identification. The 74 specimen submitted to MALDITOF MS comprised majority of *Culex pipiens* having LSV of 1.628 and 2.378. However, spectra of two *Culiseta longiareolata* were added to the reference spectra as their LSV was comparatively low (1.315–1.351). Another study of MALDI TOF MS was obtained for 13 mosquito species with 5 specimens for each. The spectra obtained by Flexi analysis for 13 *Anopheles* mosquitoes are consistent in all the tested species. Thus the data basing of mosquitoes from different parts of world would help control mosquito borne diseases.

In yet another study Yssouf et al. (2014) identified 11 mosquito species from leg protein of 74 specimen collected from different region of European areas. The protein profiles have been generated for larvae of holometabolous *Culicoides* and Culicidae (Dieme et al. 2014) and for eggs of nine aedine species (Schaffner et al. 2014). The technique assisted to correctly identify a sample of 1200 collected specimens of *Culicoides*.

26.2.1.3 Phlebotomine Sand Flies

Mathis et al. (2015) identified 20 adult phlebotomine sand flies yielding high quality spectra with specificity of 94.7% (means 100 minus percentage of wrong morphological identifications). Among the 312 specimens of 20 morphologically identified sand flies there was a concordant match of 297 species. One species morphologically identified as *Ph. galilaeus* yielded a MS spectrum and was identified to be as *Ph. pernicious* by MALDITOF MS. Another specimen of sand fly morphologically identified as *Ph. perfiliewi* yielded a novel spectrum. This species is known to exist as complex species explaining its variability in mass spectrum (Mathis et al.

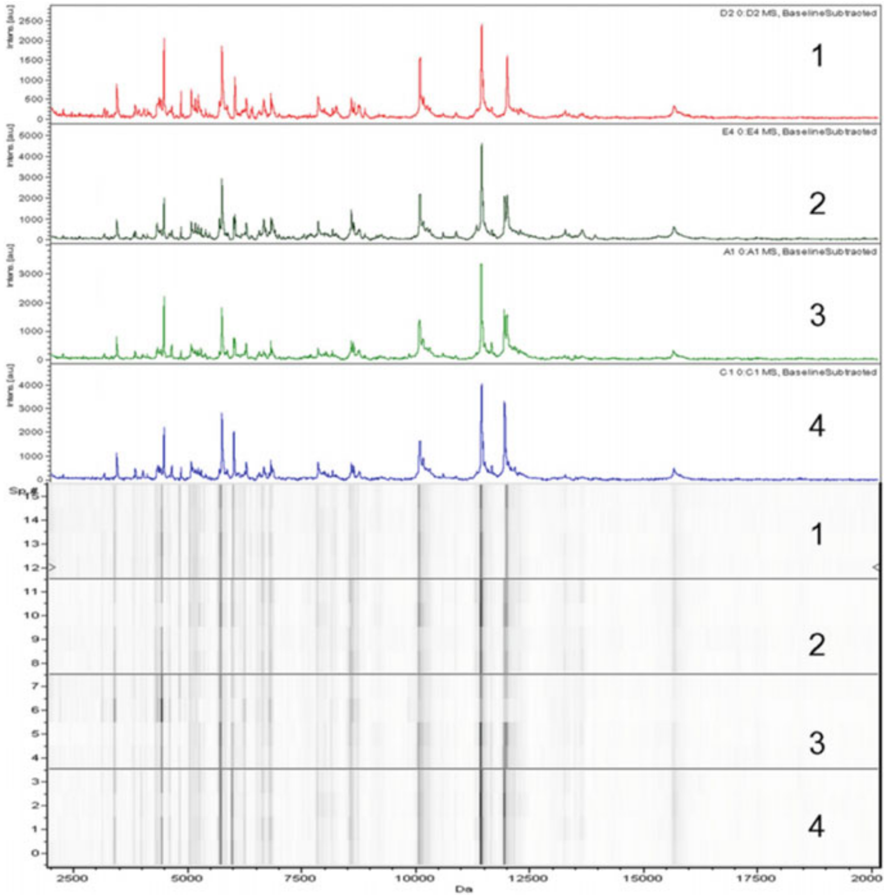


Fig. 26.4 MALDITOF MS from leg protein extraction of four *An. gambiae* form M specimens its spectra by Flexi analysis using Flex analysis. (Copyright Yssouf et al. (2013))

2015). Similarly all morphologically identified *Ph. barguesae* showed genetic difference at barcode locus corresponding to their Gene Bank entries. The super spectra, i.e. the identical biomarker mass spectra for fresh insects and 55/59 for insects frozen in liquid nitrogen at −80 °C for 25 years as seen in Fig. 26.5. However, only 36/52 specimens stored in ethanol were identified stored at room temperature or −20 °C for few years. The overall sensitivity was 87% and specificity of 100% for stored specimens compared to sensitivity for 312 specimens from 8 sand fly species giving high quality spectra of 98.3% and specificity of 100%. Thus specimens stored at different conditions and collected from different geographical locations yield different spectral yield. This also validated that storage of sample and not the period affected the mass spectrum data. It is clear from Fig. 26.5 that freezing is superior for storing the insect samples than 70% ethanol. Further 70% is superior to any higher concentrations of ethanol. To add high quality (98.9%) mass spectra

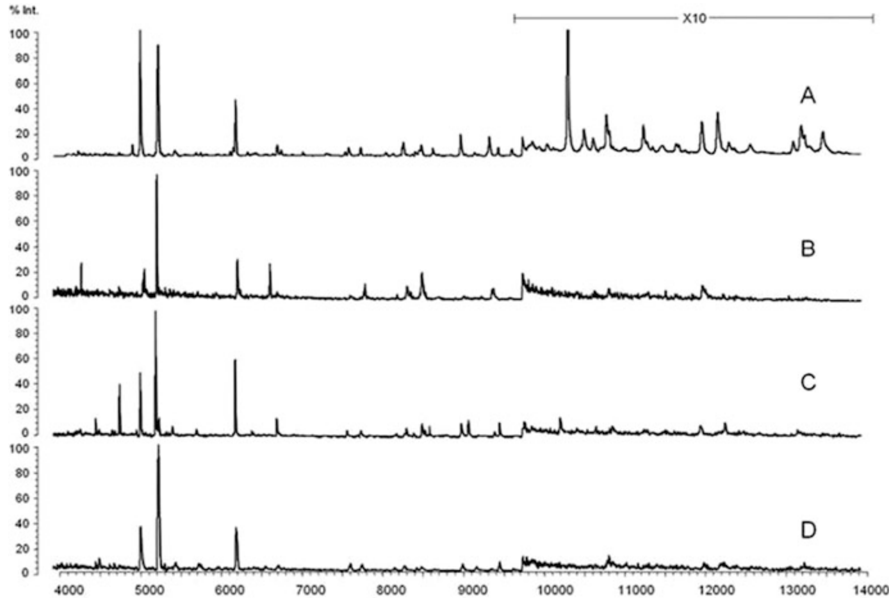


Fig. 26.5 MALDITOF mass spectra of *Phlebotomus perniciosus* (**a**) Fresh specimen from colony (**b**) field specimen stored in 70% ethanol, −20 °C for 3 years. (**c**) Field specimen stored in liquid N_2 for −80 °C since 1991. (**d**) Field specimen stored in liquid N_2 for −80 °C since 1989. Scale: 10× magnification for mass range 10–20 kDa. (Copyright Mathis et al. (2015))

was obtained for 1200 *Culicoides* stored in 70% ethanol at 4 °C for 1 year (Kaufmann et al. 2012b).

26.2.1.4 Tick Vectors

MALDI-TOF-MS was first time used to identify tick vectors however; the technique is yet to test its efficiency and reproducibility for FII. The rapid identification of ticks by these techniques may help medical officers to identify a tick and decide whether post exposure prophylactic antibiotic exposure is required or not. Yssouf et al. (2015) used MALDI-TOF-MS to correctly identify tick vectors from legs of six vectors: *Amblyomma variegatum*, *Rhicephalus sanguineus*, *Hyalomma marginatum rufipes*, *Ixodes ricinus*, *Dermacentor marginatus* and *D. reticulates*. Ticks removed from wild or from patients were identified by MALDITOF MS only if their reference spectra were available. MALDI-TOF-MS is thus an effective tool to identify tick vectors without any expertise if their reference spectra are present. In year 2011 a blind test on 111 specimens compared to *Culicoides* species showed that MALDI-TOF-MS can differentiate species of *Culicoides* collected in the fields. Karger et al. (2012) reported study of tick vectors with spectral data generated from whole body parts excluding legs. The alignment of MALDI TOF MS spectra in several specimen of same species showed reproducible results in both sexes of *Hyalomma marginatum rufipes* and *Rhipicephalus sanguineus* with major identified peaks

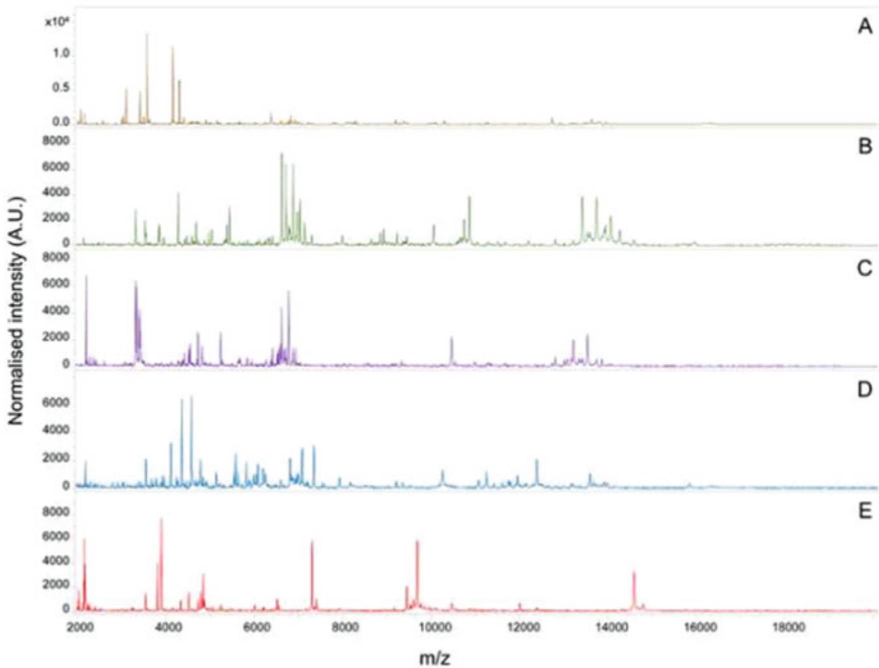


Fig. 26.6 MALDI-TOF MS spectra from leg hemolymph protein extract of (**a**): non infected *Rh. Sanguineus*, (**b**) *Rh. sanguineus* infected by *R. c. conorii* and (**c**) purified *R. c. conorii* strain. (Reproduced with permission from Yssouf et al. (2015))

being same in each specimen of same species (Yssouf et al. 2015). Although the study of Fotso et al. (2014) has shown that MALDI TOF MS of leg proteins has helped in identification of tick species and tick borne pathogens in Borrelia fever infecting Ornithodoros ticks. Among the challenges faced in tick identification using proteomic study by obtaining their protein peaks by MALDI TOF MS, the technique was applied to identify *Rickettsia. c. conorii* or *Rickettsia slovaca* pathogens in *Rhipicephalus sanguineus* and *Dermacentor marginatus* using tick legs (Yssouf et al. 2015). As the circulatory system of ticks contains circulating fluid known as hemolymph, it is the residing place for many bacteria including Rickettsia (Burgdorfer and Mavros 1970). The presence of pathogen in the haemolymph altered the protein expression profiles of these ticks (Antunes et al. 2012). The haemolymph MS profile from five such ticks were recorded in a database to identify new tick species by blind tests. This was shown by work of Yssouf et al. (2015). They showed that MS protein spectra of five tick species (*Rhipicephalus sanguineus*, *Rhipicephalus bursa*, *Dermacentor marginatus*, *Hyalomma marginatum rufipes* and *Amblyomma variegatum*) infected by *Rickettsia africae*.

The MALDITOF MS spectra of a protein mixture of a infected tick *Rh. sanguineus* by *R. c. conorii* shows a change in the MS spectra from that of a non infected according to the infection status as seen in Fig. 26.6. A total of

17 discriminating peaks, 6 were specific to uninfected *Rh. sanguineus* and 11 were specific to that infected with *R. c. conorii*. A total of 4 peaks from infected *Rh. sanguineus* were shared with MS spectra of pure *R. c. conorii* cells. The work done by Yssouf et al. (2015) can help detecting tick species infected or uninfected with *Rickettsia* and also their infection status.

26.2.1.5 Forensically Important Insects (FII)

Giffen et al. (2017) for the first time studied amino acid profile of necrophageous insects by direct analysis in real time high-resolution mass spectrometry (DART-HRMS). They raised eggs derived from four species of Calliphoroidae (*Calliphora vicina*, *Lucilia sericata*, *Lucilia coeruleiviridis*, *Phormia regina*) and members of Phoridae and Sarcophagidae families from liver of fresh Pork. The DART HRMS of the egg samples suspended in ethanol is shown in Fig. 26.7. The chemical spectra showed distinguished patterns for each species, even though they share many common peaks. *Lucilia coeruleiviridis* had highest number of peaks while *Sarcophagidae* had the least. The *m*/*z* 76 in Fig. 26.7a, b and e was observed in *C. vicina*, *L. coeruleiviridis* and *P. regina* only; and *m*/*z* 246 in *C. vicina*. The nominal *m*/*z* 271 in *L. sericata* and *P. regina* (Fig. 26.7c, d) and nominal *m*/*z* 311 in *C. vicina*, *L. coeruleiviridis*, *L. sericata* and *P. regina* (Fig. 26.7a, c, d). The MALDI TOF HRMS consistent with amino acids was also observed and it was found that presence of alanine and glycine have molecular weight below 100 Da which could not be confirmed by MALDI MS/MS. Importantly cysteine was not found in any of these species ethanol extracts for unknown reasons. The DART HRMS mass spectral fingerprints thus could help differentiate intraspecies similarities and inter species differences. The analysis showed how a direct ethanol suspension of eggs

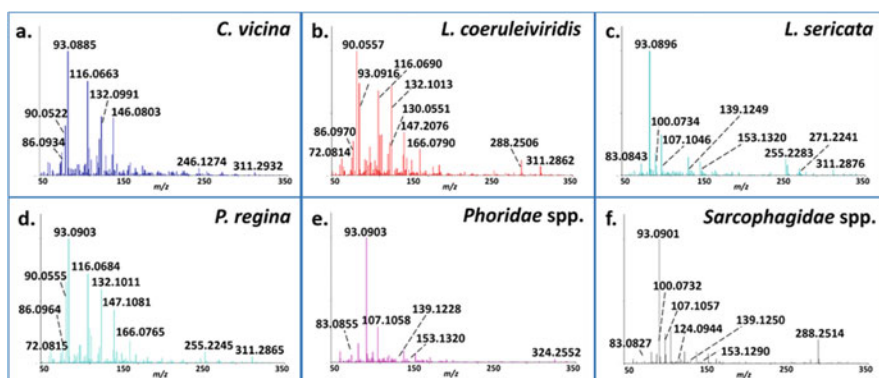


Fig. 26.7 Results of DART-HRMS analyses of aqueous ethanol suspensions of six species of fly eggs: (**a**) *C. vicina*, (**b**) *L. coeruleiviridis*, (**c**) *L. sericata*, (**d**) *P. regina*, (**e**) *Phoridae* spp., and (**f**) *Sarcophagidae* spp. All analyses were performed in positive ion mode at 350 °C. Each spectrum represents an average of three analyses. The spectra generated for each species are unique and can be used as fingerprints for the identification of individual species. (Reproduced with permission from Giffen et al. (2017))

generated important DART HRMS data for species identification without involving any pre treatment.

26.3 Age of Insect

26.3.1 Gene Expression Studies

In addition to species identification gene expression studies help in knowing the age of the insects and their stages which is very important to estimate PMI_{min}. As an organism develops, proteins are required for its development. The identification of age of an insect at different developmental stages is essential in myiasis where actual time of infestation of eggs/larva is not known. The most significant myriad flies included those belonging to group Calliphoridae and Sarcophagidae which are important forensic markers to determine time since death.

Taron and Foran profiled 11 genes in larva and pupa of *L. sericata* at different time periods of the development. The expression results helped in predicting blow fly age using generalized additive models (GAMs) (Tarone and Foran 2011). Among a population of 958 individuals; 48 first instar larvae, 79 second instar larvae, 135 third instar larvae, 334 postfeeding third instar larvae had 362 genes which were up and down regulated between feeding and post feeding stages. Four genes were not regulated at some point, gene *rop-I* was off at earliest stage of development and not included in the binary data. The expression of some genes (*cs*, *ecr*, *hhsp60*, *hsp90*, *rop-I* and *usp*) was effected by temperature with *hsp60* showing weak whereas *hsp90* and *ace* showing strong response. The nine genes evaluated in *L. sericata* (*ace*, *cs*, *eccr*, *hsp60*, *hsp90*, *rop-I*, *sll*, *usp* and *w*) showed differential regulation during developmental stages. Seven among them (*cs*, *ecr*, *hsp60*, *hsp90*, *rop-I*, *usp* and *w*) helped to differentiate feeding and post feeding third instars. Studies in *Drosophila* further shows that hormone ecdysone stimulates switching of third instar from feeding to post feeding stage (Mirth et al. 2009). Therefore the *ecr* and *usp* gene which encodes heterodimeric protein triggers ecdysone mediated changes during the development. The current study of age estimation targeting nine loci in blowfly resulted in 3–8% decrease in age overestimation and would improve accurate PMI_{min} estimate using entomological data. Gene Expression studies are done to determine the amount of mRNA through RNA extraction, reverse transcription into complementary DNA (cDNA) and quantitation of this cDNA using real time polymerase chain reaction (real-time PCR) (Campobasso and Introna 2001; Amendt et al. 2011). The microarray studies in *Drosophila melanogaster* have shown expression of 2–9% genome changes with age. The transcriptional profiles of *D. melanogaster* (Arbeitman et al. 2002) has shown gene expression changes in adult female life stages (Cook et al. 2006). Evans and Wheeler used gene expression studies for studying queen and worker specific gene expression in the development of honeybee (Evans and Wheeler 1999). A clone sequence from gene Locus *1CB6* in queen honeybee showed expression of proteins hexamerins and arylphorins which are expressed more in queens than in workers as seen in Fig. 26.8. The clone

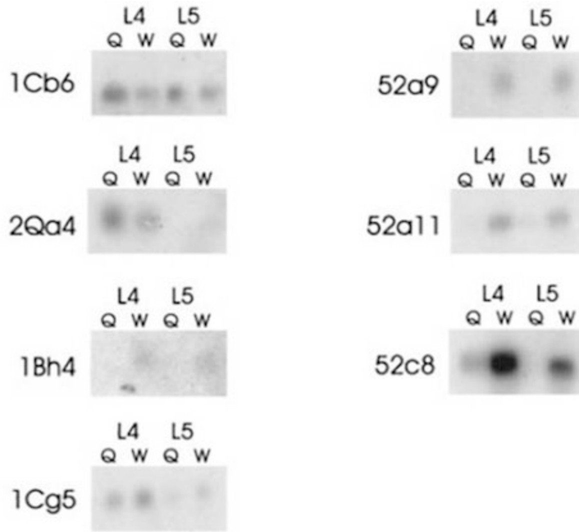


Fig. 26.8 Autoradiographs showing Northern blot results for seven differentially expressed loci. Loci 1CB6 and 2QA5 were expressed more strongly in queen larvae than in worker larvae. Locus 1CG5 was expressed at quantitatively higher levels in workers. The remaining four loci were expressed exclusively by workers. (Reproduced with permissions from Evans and Wheeler (1999))

expressed in queen shows alignment with 760–800 amino acids from larval storage protein 1 of *D. melanogaster.* The workers on the other hand show expression at locus *1CG5*. The amino acid sequence for this locus showed similarity with protein described from lens tissue in mammals. Another clone sequence *2QA4* was also expressed at higher level in fourth instar Queen than in workers. The locus expressed by *1CG5* was however expressed more in worker than in Queen. There were four more additional clone sequences which were strongly expressed by the worker bees than by the queens. The clone 1BH4 which is related to retinoic acid binding proteins is widespread in vertebrates; the clone is aligned with 80% of fatty acid binding proteins in tobacco hornworm *Manduca sexta* (Evans and Wheeler 1999). Another worker expressed clone *52C8*, encoded hexamerin storage protein. The *52A9* clone sequence of worker bees on the other hand showed similarity over 200 amino acid region of oxido-reductase enzyme family. There was great resemblance between clone *52A11* and *ELK-3* between amino acids 210 and 300 of *ELK-3* locus (Evans and Wheeler 1999). Tarone et al. (2007) investigated age of eggs of *L. sericata* by gene expression studies profiled expression of three genes (*bicoid*, *slalom*, *chitin synthatase*). Since gene expression studies are extremely important to investigate age of the larva and pupa at different stages when morphological identification fails specially in cases of myiasis.

However, the gene expression method to estimate age in metamorphosing flies is dependent on the storage conditions (Bala and Sharma 2016). The quantification by realtime-PCR is yields accurate on fresh samples. Even though FII needs to be explored on the possibility of accuracy, extensive study has been done on *Drosophila melangoster* (Diptera) (Beckstead et al. 2005). On the other hand MALDI TOF MS can be done for samples stored, frozen, fixatives, ethanol (Ellison and Hampton

1982). They are more accurate to determine chronological age of young insects of both sexes and for any field material fresh or stored (Bala and Sharma 2016).

26.3.2 MALDI TOF MS

Takeuchi et al. (2003) combined MALDI TOF MS, gene expression and on line capillary reverse-phase HPLC/quadrupole orthogonal acceleration time of flight (Q-TOF) MS to estimate age of different stages of honey bee (Yasuda-Kamatani and Yasuda 2000). They reported that a tachykinin related neuropeptide (TRPs) content is regulated via varied expression of *preproAmTRP* gene in different sex and at different stages of honey bee. The matrix is mixed with brain sample and mounted on the metal plate towards the laser beam. The identified peptide was named as *Apis mellifera* tachykinin-related peptide (AmTRP). After doing cloning of complementary DNA (cDNA) coding AmTRP gene Northern blotting was done to check the expression of genes. It was found that AmTRP is related with female like behaviours in honey bees. On the other hand the gene expression varied with age and showed a intense band in forager bee RNA and less dense in nurse bee RNA (Takeuchi et al. 2003). This study demonstrated that AmTRP neuropeptides forms a molecular basis for sexual behavior in honey bees.

26.3.3 Cuticular Hydrocarbons (CHC) by Gas Chromatography Mass Spectroscopy (GC MS)

Detection of age at different stages by change in hydrocarbon profiles is also possible using cuticular hydrocarbons studies. The hydrocarbons generally vary with age. The Gas chromatography mass spectroscopy (GC MS) results for characterization of larval stages in *Aldrichina graham*, showed that low molecular weight hydrocarbons decreased with larval development and high molecular weight >25 hydrocarbons increased with age (Xu et al. 2014). This is illustrated in Fig. 26.9, the 3 day larva showed 2-methyl-C24 pentacosenes (a and b); n-C25, 2-methyl-C26 heptacosenes (a and b) and less of n-C29. As the age of larvae progressed to fourth day most of the peak pattern was same as that of third day old, however, n-C29 peak expressed effectively along with n-C27. Also n-C25 was most abundantly found in third day larvae ($13.25 \pm 1.54\%$) whereas n-C29 in seventh day larvae ($28.19 \pm 4.88\%$).

The cuticular hydrocarbons (CHC) pattern in insects show variations in insects at different stages for species morphologically undistinguishable. This helps in identifying not only the species and also its approximate age (Zhu et al. 2007; Pechal et al. 2014). These hydrocarbons range from C23 to C33 and are generally composed of n-alkanes, n-alkenes and methyl branched alkanes (Pechal et al. 2014). The CHC profile of North American female native blow fly (*Cochliomyia macellaria*), the primary colonizer and invasive species *Chrysomya rufifacies*, the secondary colonizer at 1, 5, 10, 20 and 30 days was studied to demonstrate age associated change in CHC profile as shown in Fig. 26.10. It was seen that both the

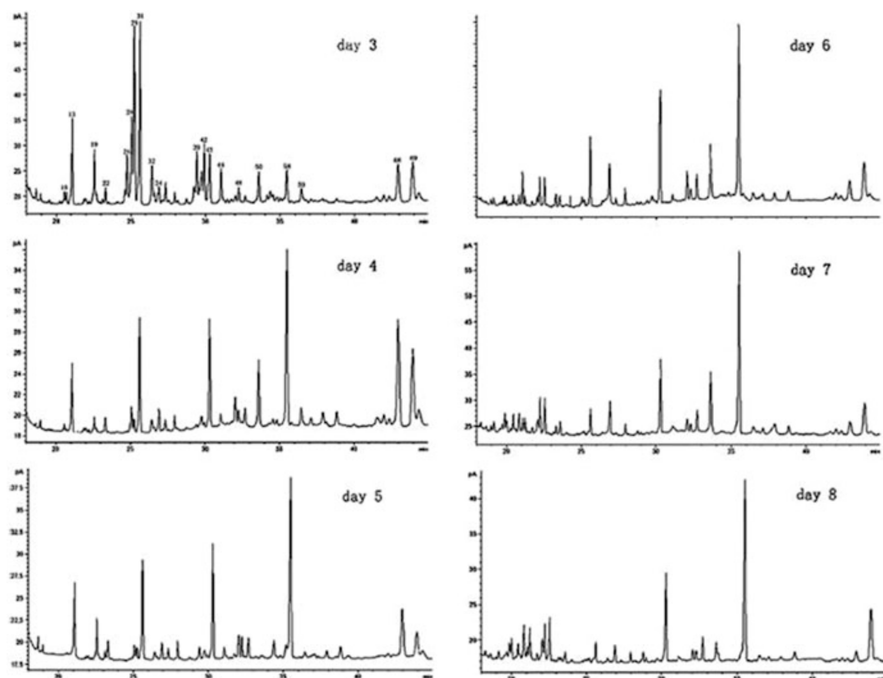


Fig. 26.9 Gas chromatographic profile of cuticular hydrocarbons extracted from larvae of *Aldrichina graham* at various stages of oviposition. (Reproduced with permission from Xu et al. (2014))

genera had 40% of similar hydrocarbons, in both it was also noted that hydrocarbon profile detected on day one of their growth was remarkably different from that found on rest of the days. Further as the adult *Cochliomyia macellaria* matured the hydrocarbon richness increased and reached the plateau on fifth day of post emergence. However, *Chrysomya rufifacies* hydrocarbon profiles decreased as the flies aged. This also proves the fact that adult flies have high molecular weight hydrocarbons, since they have to tolerate unfavorable environmental conditions which require impermeability on their cuticular surface (Toolson and Kuper-Simbron 1989; Roux et al. 2008). This clearly demonstrates that CHC profile generated by GC gives unique chemical fingerprints for native and invasive blow flies which changed constantly with age until reaching a plateau at its adult stage as seen in Fig. 26.10.

26.4 Future of MALDI TOF MS in FII

Li et al. (2017) for the first time used MALDI TOF MS based MALDI imaging mass spectrometry (MALDI IMS) and principal component analysis (PCA) to estimate PMI_{min} from the rat muscle tissue. They visualized the proteins from the slice of a

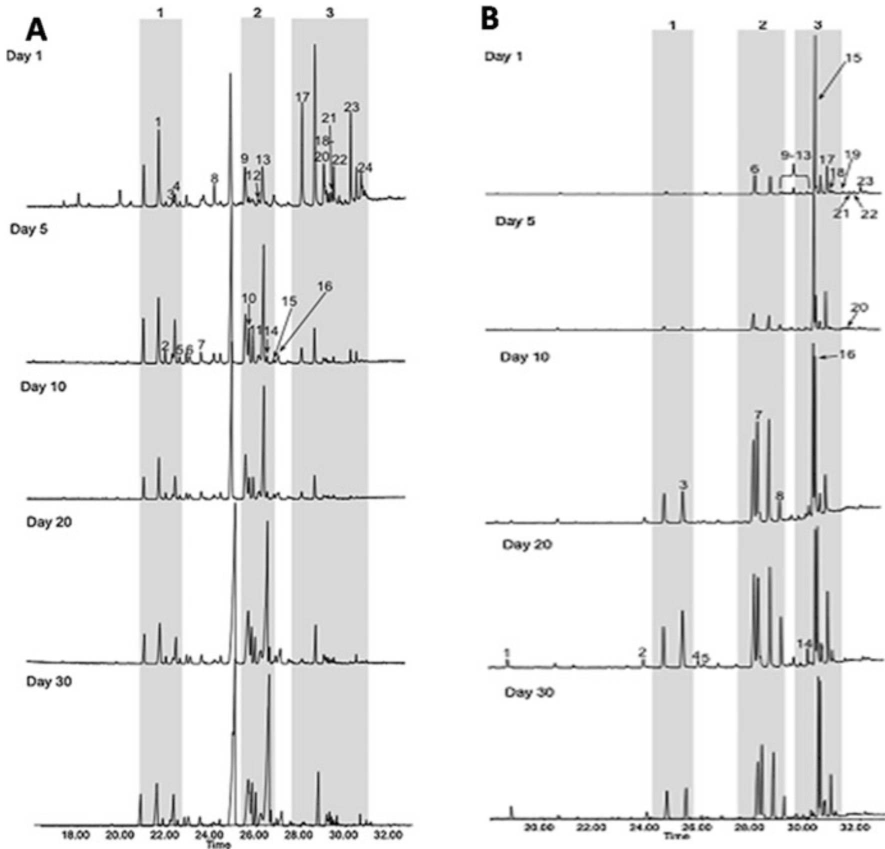


Fig. 26.10 GC chromatograms on day 1, 5, 10, 20 and 30 post emergence in adult female of (**a**) *Cochliomyia macellaria* and (**b**) *Chrysomya rufifacies*. Shaded bars within the chromatograms illustrate distinctive changes over adult age indicating specific areas of interest. (Reproduced with permissions from Pechal et al. (2014))

frozen rat muscle tissue. It was observed that the mass spectra value decreased with the increase in PMI_{min}. Their results were coherent with the histological sections of rat muscle tissue at the respective PMI's. Thus such non molecular techniques may help forensic scientist to determine PMI_{min} by not only estimating the age of the fly but also its identification. This certainly will be of potential application if used in the forensic laboratories which are still in their developing stage as far as entomology and its importance in the court of law is concerned.

26.5 Conclusion

DNA barcoding of forensically important insects does not require any taxonomist for identification process but they do require reference nucleotide sequence available in online molecular biology tools like NCBI and COBL. Though many researchers have sequenced different FII taxas's of different order at gene and species level a more needs to be done from different geographical areas to create database. The database of mitogenomic analysis is efficient but still a cost effective procedure. An alternative technique which is fast and cheap is MALDI TOF-MS which is currently utilized to identify insects which are non forensic in origin like *Drosophila* and Sand flies. However, much progress is required to use this technique to generate profile spectra of FII. The mass fingerprinting is altogether advantageous as it does not need any reference spectra and expertise of taxonomy for species identification. Identification of FII is very important since it gives reliable truth for determining PMI_{min} interval since there are certain limitations in identification of FII both morphologically and at molecular level in cases where insects have pre-infested the carrion. On the other hand the larval and pupal stages show variation in their life cycle. Gene expression studies though help in determination of different developmental stages of insects altered by change in amino acid sequences and protein biomolecules at different phases of growth. However, this too is applicable for fresh instar stages, however cuticular hydrocarbon profile by MALDI TOF MS is more accurate to determine chorological age of instars at different stages of metamorphosis. The research data of experiment based study with animal carcass and as well as human dead bodies must therefore be used for betterment of this sphere**.** We must be riveted on the molecular recognition of insect and must prepare a database of a particular habitat with international interactions. Medico legal entomology has reached an exciting height in its evolution as testimony based on the interpretation if insect evidence is now routinely provided in court of expert witness in some countries. The increasing acceptance and recognition of medico legal entomology as a forensic discipline coupled with the increased reliance of courts on biological evidence shall continue to present increased opportunities for qualified forensic entomologist. The application of MALDI-TOF MS studies for species identification at subspecies level and their comparison with DNA sequences might therefore help to explain taxonomic relationships at species and subspecies level and should be utilized to identify FII by forensic entomologist and molecular biologist.

26.6 Ethics Approval and Consent to Participate

Not applicable.

Acknowledgements V.V. is thankful to University Grant Commission Basic Science Research start up project No.F.30-12/2014(BSR) sanctioned to her.

References

Al-Khalifa MS, Mashaly AM et al (2020) Insect species colonized indoor and outdoor human corpses in Riyadh, Saudi Arabia. J King Saud Univ Sci 32(3):1812–1817

Al-Qahtni AH, Al-Khalifa MS et al (2020) Two human cases associated with forensic insects in Riyadh, Saudi Arabia. Saudi J Biol Sci 27(3):881–886

Amendt J, Krettek R et al (2004) Forensic entomology. Naturwissenschaften 91(2):51–65

Amendt J, Richards C et al (2011) Forensic entomology: applications and limitations. Forensic Sci Med Pathol 7(4):379–392

Anderson GS (2001) Insect succession on carrion and its relationship to determining time of death. In: Forensic entomology: the utility of arthropods in legal investigations. CRC Press, Boca Raton, FL, pp 143–176

Anderson GS, VanLaerhoven SL (1996) Initial studies on insect succession on carrion in southwestern British Columbia. J Forensic Sci 41(4):617–625

Aneyo I, Alafia O et al (2020) Aerobic microbe community and necrophagous insects associated with decomposition of pig carrion poisoned with lead. Legal Med 42:101638

Antunes S, Galindo RC et al (2012) Functional genomics studies of *Rhipicephalus* (Boophilus) *annulatus* ticks in response to infection with the cattle protozoan parasite, *Babesia bigemina*. Int J Parasitol 42(2):187–195

Arbeitman MN, Furlong EE et al (2002) Gene expression during the life cycle of *Drosophila melanogaster*. Science 297(5590):2270–2275

Ashworth JR, Wall R (1994) Responses of the sheep blowflies *Lucilia sericata* and *L. cuprina* to odour and the development of semiochemical baits. Med Vet Entomol 8(4):303–309

Audsley N, Weaver RJ (2003) Identification of neuropeptides from brains of larval *Manduca sexta* and *Lacanobia oleracea* using MALDI-TOF mass spectrometry and post-source decay. Peptides 24(10):1465–1474

Bala M, Sharma A (2016) Review of some recent techniques of age determination of blow flies having forensic implications. Egypt J Forensic Sci 6(3):203–208

Ballard JWO, Chernoff B et al (2002) Divergence of mitochondrial DNA is not corroborated by nuclear DNA, morphology, or behavior in *Drosophila simulans*. Evolution 56(3):527–545

Ballard JWO, Melvin RG et al (2007) Mitochondrial DNA variation is associated with measurable differences in life-history traits and mitochondrial metabolism in *Drosophila simulans*. Evolution 61(7):1735–1747

Beckstead RB, Lam G et al (2005) The genomic response to 20-hydroxyecdysone at the onset of *Drosophila metamorphosis*. Genome Biol 6(12):R99

Benecke M, Wells JD (2001) DNA techniques for forensic entomology. In: Forensic entomology: the utility of arthropods in legal investigations. CRC Press, Boca Raton, FL, pp 341–352

Berg MC, Benbow ME (2013) Environmental factors associated with *Phormia regina* (Diptera: Calliphoridae) oviposition. J Med Entomol 50(2):451–457

Boykin LM, Armstrong K et al (2012) DNA barcoding invasive insects: database roadblocks. Invertebr Syst 26(6):506–514

Burgdorfer W, Mavros AJ (1970) Susceptibility of various species of rodents to the relapsing fever spirochete, *Borrelia hermsii*. Infect Immun 2(3):256–259

Byrd JH, Castner JL (2009) Forensic entomology: the utility of arthropods in legal investigations. CRC Press, Boca Raton, FL

Campbell PM (2005) Species differentiation of insects and other multicellular organisms using matrix-assisted laser desorption/ionization time of flight mass spectrometry protein profiling. Syst Entomol 30(2):186–190

Campobasso CP, Introna F (2001) The forensic entomologist in the context of the forensic pathologist's role. Forensic Sci Int 120(1–2):132–139

Campobasso CP, Di Vella G et al (2001) Factors affecting decomposition and Diptera colonization. Forensic Sci Int 120(1–2):18–27

Carvalho LM, Linhares ACX et al (2001) Determination of drug levels and the effect of diazepam on the growth of necrophagous flies of forensic importance in southeastern Brazil. Forensic Sci Int 120(1–2):140–144

Claydon MA, Davey SN et al (1996) The rapid identification of intact microorganisms using mass spectrometry. Nat Biotechnol 14(11):1584–1586

Clynen E, Baggerman G et al (2003) Peptidomics of the locust corpora allata: identification of novel pyrokinins (-FXPRLamides). Peptides 24(10):1493–1500

Cook PE, Hugo LE et al (2006) The use of transcriptional profiles to predict adult mosquito age under field conditions. Proc Natl Acad Sci U S A 103(48):18060–18065

Cvačka J, Jiroš P et al (2006) Analysis of insect cuticular hydrocarbons using matrix-assisted laser desorption/ionization mass spectrometry. J Chem Ecol 32(2):409–434

Cywinska A, Hunter F et al (2006) Identifying Canadian mosquito species through DNA barcodes. Med Vet Entomol 20(4):413–424

Dadour IR, Cook DF et al (2001) Forensic entomology: application, education and research in Western Australia. Forensic Sci Int 120(1):48–52

De Jong G (1995) Report of *Chrysomya megacephala* (Diptera: Calliphoridae) in northern New Mexico. Entomol News 106(4):192

Dekeirsschieter J, Verheggen F et al (2009) Cadaveric volatile organic compounds released by decaying pig carcasses (*Sus domesticus* L.) in different biotopes. Forensic Sci Int 189 (1–3):46–53

Dieme C, Yssouf A et al (2014) Accurate identification of Culicidae at aquatic developmental stages by MALDI-TOF MS profiling. Parasit Vectors 7(1):544

Donaldson AE, Lamont IL (2014) Estimation of post-mortem interval using biochemical markers. Aust J Forensic Sci 46(1):8–26

Dridi B, Raoult D et al (2012) Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry identification of Archaea: towards the universal identification of living organisms. APMIS 120(2):85–91

Dvorak V, Halada P et al (2014) Identification of phlebotomine sand flies (Diptera: Psychodidae) by matrix-assisted laser desorption/ionization time of flight mass spectrometry. Parasit Vectors 7 (1):21

Ellison JR, Hampton EN (1982) Age determination using the apodeme structure in adult screw-worm flies (*Cochliomyia hominivorax*). J Insect Physiol 28(9):731–736

Evans JD, Wheeler DE (1999) Differential gene expression between developing queens and workers in the honey bee, *Apis mellifera*. Proc Natl Acad Sci U S A 96(10):5575–5580

Federhen S (2011) The NCBI taxonomy database. Nucleic Acids Res 40(D1):D136–D143

Feltens R, Görner R et al (2010) Discrimination of different species from the genus *Drosophila* by intact protein profiling using matrix-assisted laser desorption ionization mass spectrometry. BMC Evol Biol 10(1):95

Fotso AF, Mediannikov O et al (2014) MALDI-TOF mass spectrometry detection of pathogens in vectors: the *Borrelia crocidurae*/*Ornithodoros sonrai* paradigm. PLoS Negl Trop Dis 8(7): e2984

Freiwald A, Sauer S (2009) Phylogenetic classification and identification of bacteria by mass spectrometry. Nat Protoc 4(5):732

Gennard D (2012) Forensic entomology: an introduction. John Wiley & Sons, Hoboken, NJ

Giffen JE, Rosati JY et al (2017) Species identification of necrophagous insect eggs based on amino acid profile differences revealed by direct analysis in real time-high resolution mass spectrometry. Anal Chem 89(14):7719–7726

Goff M (1993) Estimation of postmortem interval using arthropod development and successional patterns. Forensic Sci Rev 5:81–81

Greenberg B, Kunich JC (2002) Entomology and the law: flies as forensic indicators. Cambridge University Press, Cambridge

Griffiths K, Krosch MN et al (2020) Variation in decomposition stages and carrion insect succession in a dry tropical climate and its effect on estimating postmortem interval. Forensic Sci Res:1–9. https://doi.org/10.1080/20961790.2020.1733830

Gupta A, Setia P (2004) Forensic entomology—past, present and future. Anil Aggrawals Internet J Forensic Med Toxicol 5(1):50–53

Gutiérrez-Gutiérrez C, Cantalapiedra-Navarrete C et al (2013) Molecular phylogeny of the nematode genus *Longidorus* (Nematoda: Longidoridae) with description of three new species. Zool J Linnean Soc 167(4):473–500

Hart AJ, Whitaker AP (2006) Forensic entomology: insect activity and its role in the decomposition of human cadavers. Antenna 30(4):159–164

Hebert PD, Cywinska A et al (2003) Biological identifications through DNA barcodes. Proc R Soc Lond B Biol Sci 270(1512):313–321

Hebert PD, Stoeckle MY et al (2004) Identification of birds through DNA barcodes. PLoS Biol 2 (10):e312

Hofer IM, Hart AJ et al (2017) Optimising crime scene temperature collection for forensic entomology casework. Forensic Sci Int 270:129–138

Holland R, Wilkes J et al (1996) Rapid identification of intact whole bacteria based on spectral patterns using matrix-assisted laser desorption/ionization with time-of-flight mass spectrometry. Rapid Commun Mass Spectrom 10(10):1227–1232

Hoppenheit A, Murugaiyan J et al (2013) Identification of Tsetse (*Glossina* spp.) using matrix-assisted laser desorption/ionisation time of flight mass spectrometry. PLoS Negl Trop Dis 7(7): e2305

Ingrisch S (1995) Evolution of the *Chorthippus biguttulus* group (Orthoptera, Acriditae) in the Alps, based on morphology and stridulation. Rev Suisse Zool 102(2):475–535

Inosaki A, Yasuda A et al (2010) Mass spectrometric analysis of peptides in brain neurosecretory cells and neurohemal organs in the adult blowfly, *Protophormia terraenovae*. Comp Biochem Physiol A Mol Integr Physiol 155(2):190–199

Jiggins FM (2002) The rate of recombination in *Wolbachia* bacteria. Mol Biol Evol 19 (9):1640–1643

Karger A, Kampen H et al (2012) Species determination and characterization of developmental stages of ticks by whole-animal matrix-assisted laser desorption/ionization mass spectrometry. Ticks Tick Borne Dis 3(2):78–89

Kaufmann C, Ziegler D et al (2011) Evaluation of matrix-assisted laser desorption/ionization time of flight mass spectrometry for characterization of *Culicoides nubeculosus* biting midges. Med Vet Entomol 25(1):32–38

Kaufmann C, Schaffner F et al (2012a) Identification of field-caught Culicoides biting midges using matrix-assisted laser desorption/ionization time of flight mass spectrometry. Parasitology 139 (2):248–258

Kaufmann C, Steinmann IC et al (2012b) Spatio-temporal occurrence of Culicoides biting midges in the climatic regions of Switzerland, along with large scale species identification by MALDI-TOF mass spectrometry. Parasit Vectors 5(1):246

Krishnamurthy T, Rajamani U et al (1996) Detection of pathogenic and non-pathogenic bacteria by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Rapid Commun Mass Spectrom 10(8):883–888

La Scola B, Campocasso A et al (2010) Tentative characterization of new environmental giant viruses by MALDI-TOF mass spectrometry. Intervirology 53(5):344–353

Li C, Ma D et al (2017) Application of MALDI-TOF MS for estimating the postmortem interval in rat muscle samples. J Forensic Sci 62(5):1345–1350

Macht M, Asperger A et al (2004) Comparison of laser-induced dissociation and high-energy collision-induced dissociation using matrix-assisted laser desorption/ionization tandem time-of-flight (MALDI-TOF/TOF) for peptide and protein identification. Rapid Commun Mass Spectrom 18(18):2093–2105

Mangold A, Bargues M et al (1997) 18S rRNA gene sequences and phylogenetic relationships of European hard-tick species (Acari: Ixodidae). Parasitol Res 84(1):31–37

Mathis A, Depaquit J et al (2015) Identification of phlebotomine sand flies using one MALDI-TOF MS reference database and two mass spectrometer systems. Parasit Vectors 8(1):266

Medzihradszky KF, Campbell JM et al (2000) The characteristics of peptide collision-induced dissociation using a high-performance MALDI-TOF/TOF tandem mass spectrometer. Anal Chem 72(3):552–558

Mellmann A, Cloud J et al (2008) Evaluation of matrix-assisted laser desorption ionization-time-of-flight mass spectrometry in comparison to 16S rRNA gene sequencing for species identification of nonfermenting bacteria. J Clin Microbiol 46(6):1946–1954

Metcalf JL, Parfrey LW et al (2013) A microbial clock provides an accurate estimate of the postmortem interval in a mouse model system. Elife 2:e01104

Mirth CK, Truman JW et al (2009) The ecdysone receptor controls the post-critical weight switch to nutrition-independent differentiation in *Drosophila* wing imaginal discs. Development 136 (14):2345–2353

Mizrachi I (2007) GenBank: the nucleotide sequence database. The NCBI Handbook [Internet], updated 22

Morris M, Woolhouse A et al (1998) Orientation stimulants from substances attractive to *Lucilia cuprina* (Diptera, Calliphoridae). Aust J Exp Agric 38(5):461–468

Nachman RJ, Russell WK et al (2005) Mass spectrometric assignment of Leu/Ile in neuropeptides from single neurohemal organ preparations of insects. Peptides 26(11):2151–2156

Narita S, Nomura M et al (2006) Genetic structure of sibling butterfly species affected by Wolbachia infection sweep: evolutionary and biogeographical implications. Mol Ecol 15 (4):1095–1108

Nekola JC, Barthel M (2002) Morphometric analysis of the genus *Carychium* in the Great Lakes region of North America. J Conchol 37(5):515–532

Noël S, Tessier N et al (2004) Molecular identification of two species of myiasis-causing *Cuterebra* by multiplex PCR and RFLP. Med Vet Entomol 18(2):161–166

Norris DE, Klompen J et al (1996) Population genetics of *Ixodes scapularis* (Acari: Ixodidae) based on mitochondrial 16S and 12S genes. J Med Entomol 33(1):78–89

Packer L, Gibbs J et al (2009) DNA barcoding and the mediocrity of morphology. Mol Ecol Resour 9:42–50

Payne JA (1965) A summer carrion study of the baby pig *Sus scrofa* Linnaeus. Ecology 46 (5):592–602

Pechal JL, Moore H et al (2014) Hydrocarbon profiles throughout adult Calliphoridae aging: a promising tool for forensic entomology. Forensic Sci Int 245:65–71

Perera MR, Vargas RDF et al (2005) Identification of aphid species using protein profiling and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Entomol Exp Appl 117(3):243–247

Pohjoismäki JL, Karhunen PJ et al (2010) Indoors forensic entomology: colonization of human remains in closed environments by specific species of sarcosaprophagous flies. Forensic Sci Int 199(1–3):38–42

Powell JR, DeSalle R (1995) Drosophila molecular phylogenies and their uses. In: Evolutionary biology. Springer, New York, pp 87–138

Predel R, Wegener C et al (2004) Peptidomics of CNS-associated neurohemal systems of adult *Drosophila melanogaster*: a mass spectrometric survey of peptides from individual flies. J Comp Neurol 474(3):379–392

Predel R, Roth S et al (2005) New insect order Mantophasmatodea: species differentiation by mass fingerprints of peptide hormones? J Zool Syst Evol Res 43(2):149–156

Rahman MM, Neupert S et al (2013) Neuropeptidomics of the Australian sheep blowfly *Lucilia cuprina* (Wiedemann) and related Diptera. Peptides 41:31–37

Ratnasingham S, Hebert PD (2007) BOLD: the barcode of life data system. Mol Ecol Notes 7 (3):355–364. http://www.barcodinglife.org

Rindi F, Guiry MD et al (2008) Distribution, morphology, and phylogeny of Klebsormidium (Klebsormidiales, Charophyceae) in urban environments in Europe 1. J Phycol 44 (6):1529–1540

Rousset F, Solignac M (1995) Evolution of single and double Wolbachia symbioses during speciation in the *Drosophila simulans* complex. Proc Natl Acad Sci U S A 92(14):6389–6393

Roux O, Gers C et al (2008) Ontogenetic study of three Calliphoridae of forensic importance through cuticular hydrocarbon analysis. Med Vet Entomol 22(4):309–317

Sambou M, Aubadie-Ladrix M et al (2015) Comparison of matrix-assisted laser desorption ionization–time of flight mass spectrometry and molecular biology techniques for identification of Culicoides (Diptera: Ceratopogonidae) biting midges in Senegal. J Clin Microbiol 53 (2):410–418

Sauer S, Freiwald A et al (2008) Classification and identification of bacteria by mass spectrometry and computational analysis. PLoS One 3(7):e2843

Sawamura K (2000) Genetics of hybrid inviability and sterility in *Drosophila*: the *Drosophila melanogaster–Drosophila simulans* case. Plant Species Biol 15(3):237–247

Schaffner F, Kaufmann C et al (2014) Rapid protein profiling facilitates surveillance of invasive mosquito species. Parasit Vectors 7(1):142

Seng P, Rolain J-M et al (2010) MALDI-TOF-mass spectrometry applications in clinical microbiology. Future Microbiol 5(11):1733–1754

Singhal N, Kumar M et al (2015) MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. Front Microbiol 6:791

Sperling FA, Anderson GS et al (1994) A DNA-based approach to the identification of insect species used for postmortem interval estimation. J Forensic Sci 39(2):418–427

Statheropoulos M, Agapiou A et al (2007) Environmental aspects of VOCs evolved in the early stages of human decomposition. Sci Total Environ 385(1–3):221–227

Steinmann IC, Pflüger V et al (2013) Evaluation of matrix-assisted laser desorption/ionization time of flight mass spectrometry for the identification of ceratopogonid and culicid larvae. Parasitology 140(3):318–327

Takeuchi H, Yasuda A et al (2003) Identification of a tachykinin-related neuropeptide from the honeybee brain using direct MALDI-TOF MS and its gene expression in worker, queen and drone heads. Insect Mol Biol 12(3):291–298

Tantawi TI, El-Kady EM et al (1996) Arthropod succession on exposed rabbit carrion in Alexandria, Egypt. J Med Entomol 33(4):566–580

Tarone AM, Foran DR (2011) Gene expression during blow fly development: improving the precision of age estimates in forensic entomology. J Forensic Sci 56:S112–S122

Tarone AM, Jennings KC et al (2007) Aging blow fly eggs using gene expression: a feasibility study. J Forensic Sci 52(6):1350–1354

Toolson EC, Kuper-Simbron R (1989) Laboratory evolution of epicuticular hydrocarbon composition and cuticular permeability in *Drosophila pseudoobscura*: effects on sexual dimorphism and thermal-acclimation ability. Evolution 43(2):468–473

Tullis K, Goff ML (1987) Arthropod succession in exposed carrion in a tropical rainforest on O'ahu Island, Hawai'i. J Med Entomol 24(3):332–339

Ulrich S, Kühn U et al (2017) Direct identification of edible insects by MALDI-TOF mass spectrometry. Food Control 76:96–101

Wegener C, Gorbashov A (2008) Molecular evolution of neuropeptides in the genus *Drosophila*. Genome Biol 9(8):R131

Wells J, LaMotte L (2001) Estimating the postmortem interval. In: Byrd JH, Castner JL (eds) Forensic entomology: the utility of arthropods in legal investigations. CRC Press, Boca Raton, FL

Wells JD, Sperling FA (2001) DNA-based identification of forensically important Chrysomyinae (Diptera: Calliphoridae). Forensic Sci Int 120(1–2):110–115

Wells JD, Stevens JR (2008) Application of DNA-based methods in forensic entomology. Annu Rev Entomol 53:103–120

Wells JD, Pape T et al (2001) DNA-based identification and molecular systematics of forensically important Sarcophagidae (Diptera). J Forensic Sci 46(5):1098–1102
Whitworth T, Dawson R et al (2007) DNA barcoding cannot reliably identify species of the blowfly genus *Protocalliphora* (Diptera: Calliphoridae). Proc R Soc B Biol Sci 274(1619):1731–1739
Xinghua W, Jifeng C et al (2010) The availability of 16SrDNA gene for identifying forensically important blowflies in China. Rom J Leg Med 1:43–50
Xu H, Ye G-Y et al (2014) Age-dependent changes in cuticular hydrocarbons of larvae in *Aldrichina grahami* (Aldrich) (Diptera: Calliphoridae). Forensic Sci Int 242:236–241
Yasuda-Kamatani Y, Yasuda A (2000) Identification of orcokinin gene-related peptides in the brain of the crayfish *Procambarus clarkii* by the combination of MALDI-TOF and on-line capillary HPLC/Q-TOF mass spectrometries and molecular cloning. Gen Comp Endocrinol 118 (1):161–172
Yssouf A, Socolovschi C et al (2013) Matrix-assisted laser desorption ionization-time of flight mass spectrometry: an emerging tool for the rapid identification of mosquito vectors. PLoS One 8(8): e72380
Yssouf A, Parola P et al (2014) Identification of European mosquito species by MALDI-TOF MS. Parasitol Res 113(6):2375–2378
Yssouf A, Almeras L et al (2015) Identification of tick species and disseminate pathogen using hemolymph by MALDI-TOF MS. Ticks Tick Borne Dis 6(5):579–586
Zhu GH, Xu XH et al (2007) Puparial case hydrocarbons of *Chrysomya megacephala* as an indicator of the postmortem interval. Forensic Sci Int 169(1):1–5

Forensic DNA Phenotyping

27

Dolly Mahna, Meenu Dutt, Sanjeev Puri, Tanzeer Kaur, and Shweta Sharma

Abstract

Forensic DNA Phenotyping (FDP) is evolving as a new technology in forensic science which can make predictions on the *externally visible characteristics* (EVCs) of an individual based on the information derived from DNA left at a crime scene. FDP as a forensic tool is gaining tremendous importance for successfully predicting the phenotype from biological samples such as blood stains, hair strands or body parts especially in the cases where no witnesses or suspects are accessible. The technique is rapidly progressing from science fiction to science fact and is already been able to predict a number of EVCs e.g. gender, height, male baldness pattern, colour of iris, hair and skin and facial features. Although, many countries are making use of FDP in solving real life criminal cases but certain legal and ethical burdens are still attached to it. These dilemmas are substantial but not invincible and can be overcome with suitable regulatory protocols. There is an urgent need to take into account several ethical and legal issues along with extensive work which is to be done on scientific research and technological advancement to make the technique readily available for forensic investigations and judicial authorities.

D. Mahna · M. Dutt · S. Sharma (✉)
Institute of Forensic Science & Criminology, Panjab University, Chandigarh, India
e-mail: 25shweta@pu.ac.in

S. Puri
Biotechnology Branch, University Institute of Engineering & Technology, Panjab University, Chandigarh, India

T. Kaur
Department of Biophysics, Panjab University, Chandigarh, India

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_27

Keywords

Forensic DNA phenotyping · Characteristics · Criminal investigations · Short tandem repeat · Single nucleotide polymorphisms

27.1 Introduction

The phenotype of an organism is an observable characteristic or traits (genetic) including its biochemical aspects, physiological features and morphological or behavioural characters. Phenotype changes could be altered due to genetic modifications or environmental factors. Phenomics is another study that can be used to distinguish individuals on the grounds of modified genomics owing to different health problems (illness) or the developmental suitability. Forensic DNA Phenotyping (FDP) implicates investigation of externally visible as well as non-visible characteristics. It is an application in the area of genetics where information is dug up from human phenotypes from biological samples through molecular analysis. The information extracted from the DNA could be linked to the crime scene and can further help in a forensic investigation such as disease traits, behavioural characters etc. The biological trace sample found at a crime scene is further analysed based on bio BN markers found in DNA, RNA, or mitochondria. There are two methods to estimate human phenotypes from DNA including direct and indirect method (Seo et al. 2017). In direct methods the genetic variants are associated with expression of the phenotypes, measured with appropriate statistical methodologies to conclude a trait value. While in the indirect methods, the genetic component is compared with the phenotype informative markers such as ancestry markers, which are evaluated using appropriate statistical methodologies to summarize the significance of trait values (Hill 2010).

DNA phenotyping denoted as a "biological witness" (Liu et al. 2019) acts as an eye-witness that describes the appearance of a person of interest. The direct method is always considered as a better option because genetic architecture of the phenotype, is not always possible. Forensic DNA phenotyping tests are used to predict the genetic information, but it is challenging to predict the external characteristic likelihood. The phenotypic traits can be very helpful as they deal with a multifaceted interaction between genetic markers and the environmental effects (Hill 2010). Some of the phenotypic features that may assist in creating justice helpful in the inquiry or elevated opportunities are appearance (externally visible characteristics [EVCs]), such as hair, skin and eye color), biological age and bio-geographic background (Kayser 2015). Short tandem repeat (STR) and single nucleotide polymorphisms (SNP) are considered to be the golden standards in forensic science. Both STR and SNP are used in the identification of known and unknown profile matching. However, one of the major limitations of this comparative approach of DNA identification fails to identify persons whose STR or SNP profile is not already known to the investigators (Hussing et al. 2018). There can be many reasons to get diverted in the investigation as persons may be unavailable or have escaped successfully from the

police investigations and thus avoid becoming a known suspect. So the current approach to this problem is getting its shape by collecting database which is screened by investigators, in routine. Sometimes cases may take a long time due to lack of evidence and/or DNA database for the investigation in unknown forensic sample donors (Kayser and de Knijff 2011).

27.2 Phenotypic Variations

Mutation, gene flow, genetic drift and natural selection provide the overview of genetic diversity development and maintenance with the growth of human population. Unlike most animals, individual gene frequencies have been shaped by human migration from one environment to another in an extremely wide range of environmental ecosystems, genetic adaptation and genetic drift (Morjan and Rieseberg 2004).

Skin tone; darker skin is a trait of communities lately residing in equatorial areas, e.g. South Indians, Sub Saharan Africans, natives of the Australian continent because melanin absorbs damaging ultraviolet radiation before it can make way to the epidermis nuclei and harm DNA. Whereas, Caucasian population exposed to low-UV environment acquires lighter skin tone (Kowalczyk et al. 2018; Norton et al. 2005).

In determining patterns of overt genetic variation various factors such as natural selection, natural variation or genetic drift, recombination, mutation and gene conversion, sexual selection come into the scenario playing a significant part. These factors are known to interact with statistical sampling which is induced by the influence of the founders, isolation, migration and demographic bottlenecks, certainly establishing phenotype diversity which even holds a substantial existence in the human family tree in modern times (Phillips et al. 2018). A few apparent phenotypic distinctions among the different population of the world include:

1. Pigmentation: skin, iris, and hair colour
2. Metabolism: xenobiotics, lactose tolerance, "thrifty" 'genotypes
3. Somatic morphology: stature, mass, body proportions, hirsuteness
4. Cranial and facial morphology: epicanthic eye folds
5. Others: facial characteristics, craniofacial architecture, dental and cranial size (Sturm et al. 1998)

The anthropometric characteristic is the content of skin melanin and the Body Mass Index (BMI). BMI differs significantly with ancestry through the rule of Bergmann asserting that BMI improves with respect to distance from the equator because of the cold climatic conditions of northern parts and association of volume-to-surface ratio with heat exchange effectiveness. Bergmann's principle is also able to explain the difference in the bodily appearances of the populations belonging to a certain environment e.g. population native to tropical environments (e.g. Kenyans, Ethiopians, etc.) is big whereas locals residing in Arctic are thin. It also justifies the

presence of shorter limbs in colder climates which by maximizing the volume-to-surface ratio, allow better heat retention thus causing variation in the morphology of species (Ruff 1994).

Variations such as the length of time in the original place of a population and microclimate impacts are to be taken into account while designing an explanatory model for geographic patterns based phenotypic variance. High skin melanin content is a polyphyletic feature, which suggests that the feature is displayed among populations sharing common types of human phylogenetic lines (i.e. indigenous population of New Guinea, West African natives, Australian Aboriginals, Indians of South Asia). Traits owing their lineage to the strengths of natural extracts, e.g. xenobiotic metabolism characteristics, hair colour and iris colour, are often discovered in various distinct ancestries (Chaitanya et al. 2018). It is quite difficult to conclude that convergent evolution is developed from comparable phenotypes across divergent populations. Additionally, the genetic intersection cannot be closed from phenotype convergence.

Eye colour may be used to represent the viability of accurate genetic prediction of complex human phenotypes. Although eye colour is commonly modelled as a simple, Mendelian trait, extended research and observation has shown that eye colour does not come after the classical ways of inheritance (Parson 2018). Eye colour phenotypes demonstrate both epistasis and incomplete control. Although eye colour is liable for some 16 distinct genes, it is mostly attributed to two adjacent genes on chromosome 15, hect domain, and domain-like RCC1 protein 2 (HERC2) and ocular albinism (i.e. oculocutaneous albinism II (OCA2)). A HERC2 intron includes the OCA2 promoter region that affects its expression (White and Rabago-Smith 2011). Consequently, single-nucleotide polymorphisms have a large role in an individual eye colour in either of these two variables. In addition, anomaly happens with all genetic expression. Some individuals may display two phenotypes—one in each eye or a complete absence of pigmentation i.e. albinism in the eye.

Facial features are interesting aspects for phenotyping as individuals are acknowledged primarily through their craniofacial morphology. Facial morphology can be a sort of variation that can be assessed using indirect molecular photography techniques (based on a knowledge of genomic ancestry) (Williams and Wienroth 2017). There are clear signs from family observations that there is a big genetic element to the expression of facial characteristics that is likely to be developed as a concrete phylogeographical function into a concrete.

27.3 Methods of Forensic DNA Phenotyping

DNA polymorphisms (short tandem repeats [STRs]) based human identification is considered to be the benchmark among forensic science techniques. STRs, which are basically polymorphisms created by copies of tiny DNA segments usually range from 2 to 6 base pairs. Many informative alleles have already been identified in humans and there are some estimations which suggest that about one million STRs exist in a distributed form among the human genome (MacLean and Lamparello

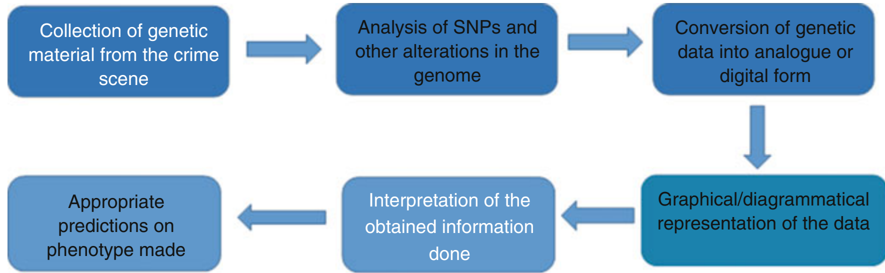


Fig. 27.1 Step by step depiction of forensic DNA phenotyping

2014). A variety of other variations e.g. insertion/deletion and single nucleotide polymorphisms (SNPs) present in DNA coding regions having the capability to cause significant changes in the functional properties of translated proteins and amino acid substitutions are generally studied (Branicki et al. 2007). These alterations which can be expressed in particular phenotypes as visible characteristics of an individual are analysed and converted to digital/ mathematical form to predict phenotypes.

The acquisition of genetic material is the first step in the commencement of forensic DNA phenotyping (MacLean and Lamparello 2014). The material is collected from the crime scene by sampling or swabbing. Analysis of genetic sample is carried out afterwards which generates genetic information in digital form or analogue. Interpretation of raw genetic information is done which is further refined into data for forensic phenotyping or comparison purposes. This data is then represented diagrammatically, graphically or in the form of predictive statements (Walsh 2004) (Fig. 27.1).

27.4 Challenges of Forensic DNA Phenotyping

Processing of biological samples collected from crime scene is done to acquire a DNA profile which is then compared to the profile of the suspect which aids the investigation by either building an association between the culprit and the crime scene or sometimes even eliminating the suspects (Budowle et al. 2005). This approach can also be used in the identification of unidentified bodies or missing persons by comparing or matching their profiles in familial searches. Although, FDP seems to be a promising tool in providing phenotypes of victims and suspects in forensic cases but there are several burdens on it which pull it back, for example, there are many cases where genetic information belongs to more than one person e.g. intimate swab from the rape victim contains genetic material from both the suspect and victim which leads to mixture of their genetic information thus making it a challenging task to predict the phenotype (Machado and Silva 2019). Similarly, there are several other challenges discussed below which in a way or the other hamper the proper utilization of FDP in forensic cases.

27.4.1 Artificially Changed Appearance

EVCs can be changed artificially through many cosmetic procedures or surgery, for example, dyed hair colour, coloured contact lenses, artificial hair extensions, facial plastic surgery, hair transplantation etc. (Walsh 2005). The perpetrators may fake their appearances to avoid any tracing by police investigation which may occur during or after the criminal act. So, the criminals are required to register their feigned appearance in police documentations which is quite difficult (Hernando et al. 2018). In addition to this, passports, driving licences have portrait images along with a record of other physical characteristics of an individual e.g. eye colour, height record which can further assist in the identification of a suspect along with DNA predicted appearance from evidence (Vidaki et al. 2013).

Keeping in mind these complexities, it is necessary to implement the phenotyping techniques with a fair deal of planning to avoid FDP-guided investigations via feigned appearances (Zaorska et al. 2019). However, the criminals often fail to even practice the simplest act of wearing gloves which makes the method of human identification from fingerprints effective and valid even after more than 100 years of its utility (Jing et al. 2019). Even though altered appearance theoretically seems to be a burden on FDP (Forensic DNA Phenotyping) but it is unlikely to challenge FDP directed identification.

27.4.2 Ethical Issues

There are certain ethical issues attached with the use of FDP raised on the basis of privacy associated with EVCs (Sero et al. 2019). Practically, the right-not-to-know does not apply when appearance traits are concerned because they are already well known to the people who have ever seen the person and all the identity cards including driver-licences, passports, voter cards have portrait photographs (Kastelic et al. 2013). Therefore, EVCs are not eligible to be considered as personal data.

On the other hand, the right-not-to-know can be applied on genetic ancestry testing and disease-linked appearance. Utilization of the disease information of an individual for forensic purpose is not ethically suitable to be used. Such reasoning is not universally accepted as the forensic DNA legislation of US state of Texas has permitted the use of information based on genetic diseases for FDP (Kayser 2015). Problems may arise for EVCs where the same gene is associated with the normal as well as diseased character. For example, SNPs related to the cleft-lips disease is also associated with the normal facial width so in such cases it becomes extremely difficult to ignore such an informative SNP (Kayser and de Knijff 2011). In such case, extensive future research is required for the identification of genetic factors responsible for the appearance of cleft thus allowing the differentiation between diseased and normal forms of facial width thus isolating the utilization of disease linked information (Pospiech et al. 2018).

27.4.3 Legal Burdens

In addition to the scientific research based burdens, there are various legal issues associated with practical FDP. Basically, the country specific legal legislations and implications on the regulation of the utilization of human DNA for DNA fingerprinting or DNA profiling do not cover FDP (Walsh et al. 2011). Although, some countries like Netherlands have modified their forensic DNA regulations to permit the application of DNA based appearance prediction in forensic cases. Some countries, such as UK, do not even have particular legislations for the utilization of FDP rather allow its utilization for forensic identification (Mengel-From et al. 2009).

27.4.4 Limited Funding and Scientific Knowledge

Apart from all the other factors imposing challenge to FDP, the significant burden is that of the dearth of immense knowledge on the genetic basis of human appearance (Aulchenko et al. 2009). Along with the restricted scientific knowledge of concerned field, insufficient research funding by government agencies is quite a major factor in not letting FDP flourish as an established technique in forensic identification (Weinberg et al. 2008). Furthermore, the literature present on the topic is quite limited with not a very large number of publications. Provided the novelty and significance of the newly built phenomenon, large number of research studies are required (Slabbert and Heathfield 2018).

Another serious burden for FDP is the unavailability of multiplex genotyping technologies for incorporating in-parallel analysis of a large number of SNPs appropriate for low quality and quantity DNA available on the crime scenes (Salvoro et al. 2019). Continuous efforts are being made to overcome this limitation by implementing massive parallel sequencing technologies viz. next generation sequencing (NGS) or second generation sequencing (SGS) technologies which have the ability to analyse quite a number of SNPs required for a detailed appearance prediction (Bradbury et al. 2019).

27.5 Discussion and Conclusions

FDP is considered to be more dependable and vindicating than racial profiling. Although, as a fact, genetic testing brings along a lot of uncertainties and does not allow the formation of outright conclusions about the appearance, behaviour or predispositions of an individual, but it is incalculably superior to conventional investigative tools such as eyewitness testimony (Scudder et al. 2018).

Although, DNA phenotyping has assisted forensic science in solving various mysteries but critics still remain doubtful of its ability to solve criminal cases. Creation of an image from phenotyping costs a lot of money and additionally, the produced physical traits are merely predictions that surely do not guarantee the

appearance of an individual. There are so many defining characteristics of an individual such as dyed hair, tattoos, facial hair, weight, which cannot be predicted by the DNA of an individual (Kukla-Bartoszek et al. 2018). Phenotyping basically predicts a snapshot image of a person embellished with certain physical characteristics possessing a little margin for error. DNA phenotyping technology is still in infancy and has a long way to go before it is allowed in the courts. There is a need to improve this technology to potentially increase the accuracy rate in the future. The supporters of DNA phenotyping technology stand strong in its favour suggesting the indispensable role this technology can play in bringing justice to victims and families in many unsolved criminal cases (Feng et al. 2018). Apart from solving the forensic cases, it can prove to be a life-saviour by making way into customized medicine and more.

The anticipated capabilities of FDP technologies are quite impactful and have a great potential to yield additional useful information to assist forensic investigations. However, its utility in real case work has not been assessed by many scholarly publications and additionally, their worth for uptake by criminal justice system still remains uncertain as reproducibility of the outcomes and genotype to phenotype predictability is subject to further refinement (Buchanan et al. 2018).

FDP is capable of identifying a source's gender with an accuracy rate of 100% and other EVCs (Externally Visible Characteristics) like iris color, adult height and hair color with an accuracy nearly 70% (Valenzuela et al. 2010). In spite of offering fairly impressive accuracy, there is a reticence among crime agencies in employing FDP as a technique to solve the cases. This diffidence is associated with numerous dilemmas such as, FDP racial profiling, its validity in the criminal court and violation of an individual's privacy but in spite of these substantive dilemmas, continuous efforts are being made to develop FDP with suitable protocols and regulations (Wray et al. 2013). However, there are arguments which rightly state that FDP i.e. a "biological witness" runs in parallel lines with the method of human eye witness (Brewer and Wells 2011). It provides the same kind of information as human eye witness i.e. the outlook of a suspect. So, if there exists no particular guideline to permit the utilization of human eye witness, then what is the need to make special laws for FDP which has even lesser chances of getting manipulated in comparison? DNA phenotyping can prove to be beneficial for the traditional DNA profiling i.e. DNA fingerprinting as it helps to narrow down the pool of possible suspects on the basis of a person's ancestry and appearance (Daniel et al. 2015). After the identification of the suspect by FDP, DNA fingerprinting can be brought into play for proving a successful match. Thus, the traditional and novel technologies can be combined to solve the cases successfully (Caliebe et al. 2018).

In the year 2018, 15 gene loci responsible for human facial features have been discovered by a research group led by Peter Claes et al., out of which seven identified genes are associated to the nose which is a great news for forensic scientists as nose is the most difficult feature to be mapped (Claes et al. 2018). There are a number of highly potential EVCs for which the information of underlying genes is not found out yet. These EVCs can prove to be a boon for FDP purposes due to high

heritability so there is a great need to do fundamental research to identify maximum genetic information capable of identifying trait variance (Sochtig et al. 2015).

There is an urgent need to support the FDP based research by government as well as private research funding agencies along with the development of various research projects by forensic institutions in collaboration with the law enforcement and police agencies (Lima et al. 2015). Extensive genetic studies are required to identify most, if not all, candidate markers which are capable of determining human appearance variations with a very little trait variance. Extensive genome wide studies are required for the mapping of EVC-underlying genes permitting the production of functional and robust markers for EVC prediction studies (Zbiec-Piekarska et al. 2015). Large genome studies are to be undertaken to map the important EVC underlying genes thus making a way for the production of candidate markers for EVC prognosis studies. Furthermore, standardized EVC phenotypes and genotypes can be put together to develop a worldwide database which will have the capability to deliver more accurate DNA-based prediction estimates (Pospiech et al. 2016). Furthermore, this information can be used to predict a concluding set of informative DNA markers for a given EVC which can be utilized by various commercial companies for the development of practical FDP tools (Toom et al. 2016).

Legal burdens on FDP are another source of hampering the growth and utility of this useful technique in forensic cases (Wolinsky 2015). There are many countries which have already permitted the use of biological witness in solving forensic cases. There exist countries which do not welcome the technique with open heart. There is an urgent need for those nations to accept this technique by making specific legislations and allowing its use in the court of law and investigation scenarios (Chaitanya et al. 2017).

In spite of all the controversies and challenges all around, the science of DNA phenotyping is only going to get better and become the future of forensic science. It is the high time to change the conservative attitude of waiting till fundamental sciences discover something for forensic science (Ackerman et al. 2016). The forensic community needs to believe in the practical benefits of FDP utilization. It would be quite effective to invest in areas such as FDP which bear higher potentials to ameliorate forensic analyses in practical scenarios and replacing conventional DNA profiling methods. The forensic researchers as well as investigators need to have substantial faith on the ever growing technology of FDP which can be utilized to assist in creating a more fair and optimistic system.

Acknowledgements The authors are thankful to DST, Govt. of India, for providing financial support to IFSC and Centre for SCTE through the PURSE grant.

References

Ackerman JP, Bartos DC, Kapplinger JD, Tester DJ, Delisle BP, Ackerman MJ (2016) The promise and peril of precision medicine: phenotyping still matters most. Mayo Clin Proc 91:1606. https://doi.org/10.1016/j.mayocp.2016.08.008

Aulchenko YS, Struchalin MV, Belonogova NM, Axenovich TI, Weedon MN, Hofman A, Uitterlinden AG, Kayser M, Oostra BA, van Duijn CM, Janssens ACJW, Borodin PM (2009) Predicting human height by Victorian and genomic methods. Eur J Hum Genet 17:1070. https://doi.org/10.1038/ejhg.2009.5

Bradbury C, Kottgen A, Staubach F (2019) Off-target phenotypes in forensic DNA phenotyping and biogeographic ancestry inference: a resource. Forensic Sci Int Genet 38:93–104. https://doi.org/10.1016/j.fsigen.2018.10.010

Branicki W, Brudnik U, Kupiec T, Wolanska-Nowak P, Wojas-Pelc A (2007) Determination of phenotype associated SNPs in the MC1R gene. J Forensic Sci 52:349–354. https://doi.org/10.1111/j.1556-4029.2006.00361.x

Brewer N, Wells GL (2011) Eyewitness identification. Curr Dir Psychol Sci 20:24–27. https://doi.org/10.1177/0963721410389169

Buchanan N, Staubach F, Wienroth M, Pfaffelhuber P, Surdu M, Lipphardt A, Kottgen A, Syndercombe-Court D, Lipphardt V (2018) Forensic DNA phenotyping legislation cannot be based on “Ideal FDP”—a response to Caliebe, Krawczak and Kayser (2017). Forensic Sci Int Genet 34:e13–e14. https://doi.org/10.1016/j.fsigen.2018.01.009

Budowle B, Bieber FR, Eisenberg AJ (2005) Forensic aspects of mass disasters: strategic considerations for DNA-based human identification. Leg Med (Tokyo) 7:230–243. https://doi.org/10.1016/j.legalmed.2005.01.001

Caliebe A, Krawczak M, Kayser M (2018) Predictive values in Forensic DNA Phenotyping are not necessarily prevalence-dependent. Forensic Sci Int Genet 33:e7–e8. https://doi.org/10.1016/j.fsigen.2017.11.006

Chaitanya L, Pajnic IZ, Walsh S, Balazic J, Zupanc T, Kayser M (2017) Bringing colour back after 70 years: predicting eye and hair colour from skeletal remains of World War II victims using the HIrisPlex system. Forensic Sci Int Genet 26:48–57. https://doi.org/10.1016/j.fsigen.2016.10.004

Chaitanya L, Breslin K, Zuniga S, Wirken L, Pospiech E, Kukla-Bartoszek M, Sijen T, de Knijff P, Liu F, Branicki W, Kayser M, Walsh S (2018) The HIrisPlex-S system for eye, hair and skin colour prediction from DNA: introduction and forensic developmental validation. Forensic Sci Int Genet 35:123–135. https://doi.org/10.1016/j.fsigen.2018.04.004

Claes P, Roosenboom J, White JD, Swigut T, Sero D, Li J, Lee MK, Zaidi A, Mattern BC, Liebowitz C, Pearson L, González T, Leslie EJ, Carlson JC, Orlova E, Suetens P, Vandermeulen D, Feingold E, Marazita ML, Shaffer JR, Wysocka J, Shriver MD, Weinberg SM (2018) Genome-wide mapping of global-to-local genetic effects on human facial shape. Nat Genet 50:414–423. https://doi.org/10.1038/s41588-018-0057-4

Daniel R, Santos C, Phillips C, Fondevila M, van Oorschot RAH, Carracedo A, Lareu MV, McNevin D (2015) A SNaPshot of next generation sequencing for forensic SNP analysis. Forensic Sci Int Genet 14:50–60. https://doi.org/10.1016/j.fsigen.2014.08.013

Feng L, Peng F, Li S, Jiang L, Sun H, Ji A, Zeng C, Li C, Liu F (2018) Systematic feature selection improves accuracy of methylation-based forensic age estimation in Han Chinese males. Forensic Sci Int Genet 35:38–45. https://doi.org/10.1016/j.fsigen.2018.03.009

Hernando B, Ibanez MV, Deserio-Cuesta JA, Soria-Navarro R, Vilar-Sastre I, Martinez-Cadenas C (2018) Genetic determinants of freckle occurrence in the Spanish population: towards ephelides prediction from human DNA samples. Forensic Sci Int Genet 33:38–47. https://doi.org/10.1016/j.fsigen.2017.11.013

Hill WG (2010) Understanding and using quantitative genetic variation. Philos Trans R Soc Lond Ser B Biol Sci 365:73–85. https://doi.org/10.1098/rstb.2009.0203

Hussing C, Huber C, Bytyci R, Mogensen HS, Morling N, Borsting C (2018) Sequencing of 231 forensic genetic markers using the MiSeq FGx forensic genomics system—an evaluation of the assay and software. Forensic Sci Res 3:111–123. https://doi.org/10.1080/20961790.2018.1446672

Jing X, Sun Y, Zhao W, Gao X, Ma M, Liu F, Li C (2019) Predicting adult height from DNA variants in a European-Asian admixed population. Int J Legal Med. https://doi.org/10.1007/s00414-019-02039-8

Kastelic V, Pospiech E, Draus-Barini J, Branicki W, Drobnic K (2013) Prediction of eye color in the Slovenian population using the IrisPlex SNPs. Croat Med J 54:381–386. https://doi.org/10.3325/cmj.2013.54.381

Kayser M (2015) Forensic DNA phenotyping: predicting human appearance from crime scene material for investigative purposes. Forensic Sci Int Genet 18:33–48. https://doi.org/10.1016/j.fsigen.2015.02.003

Kayser M, de Knijff P (2011) Improving human forensics through advances in genetics, genomics and molecular biology. Nat Rev Genet 12:179. https://doi.org/10.1038/nrg2952

Kowalczyk M, Zawadzka E, Szewczuk D, Gryzinska M, Jakubczak A (2018) Molecular markers used in forensic genetics. Med Sci Law 58:201–209. https://doi.org/10.1177/0025802418803852

Kukla-Bartoszek M, Pospiech E, Spolnicka M, Karlowska-Pik J, Strapagiel D, Zadzinska E, Rosset I, Sobalska-Kwapis M, Slomka M, Walsh S, Kayser M, Sitek A, Branicki W (2018) Investigating the impact of age-depended hair colour darkening during childhood on DNA-based hair colour prediction with the HIrisPlex system. Forensic Sci Int Genet 36:26–33. https://doi.org/10.1016/j.fsigen.2018.06.007

Lima FA, Goncalves FT, Fridman C (2015) SLC24A5 and ASIP as phenotypic predictors in Brazilian population for forensic purposes. Leg Med (Tokyo) 17:261–266. https://doi.org/10.1016/j.legalmed.2015.03.001

Liu F, Zhong K, Jing X, Uitterlinden AG, Hendriks AEJ, Drop SLS, Kayser M (2019) Update on the predictability of tall stature from DNA markers in Europeans. Forensic Sci Int Genet 42:8–13. https://doi.org/10.1016/j.fsigen.2019.05.006

Machado H, Silva S (2019) What influences public views on forensic DNA testing in the criminal field? A scoping review of quantitative evidence. Hum Genomics 13:23. https://doi.org/10.1186/s40246-019-0207-5

MacLean CE, Lamparello A (2014) Forensic DNA phenotyping in criminal investigations and criminal courts: assessing and mitigating the dilemmas inherent in the science. Recent Adv DNA Gene Seq 8:104–112

Mengel-From J, Wong TH, Morling N, Rees JL, Jackson IJ (2009) Genetic determinants of hair and eye colours in the Scottish and Danish populations. BMC Genet 10:88. https://doi.org/10.1186/1471-2156-10-88

Morjan CL, Rieseberg LH (2004) How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. Mol Ecol 13:1341–1356. https://doi.org/10.1111/j.1365-294X.2004.02164.x

Norton HL, Kittles RA, Parra E, Mckeigue P, Mao X, Cheng K, Canfield VA, Bradley DG, Mcevoy B, Shriver MD (2005) Genetic evidence for the convergent evolution of light skin in Europeans and East Asians, vol 24, p 710. https://doi.org/10.1093/molbev/msl203

Parson W (2018) Age estimation with DNA: from forensic DNA fingerprinting to forensic (epi)-genomics: a mini-review. Gerontology 64:326–332. https://doi.org/10.1159/000486239

Phillips C, Devesse L, Ballard D, van Weert L, de la Puente M, Melis S, Alvarez Iglesias V, Freire-Aradas A, Oldroyd N, Holt C, Syndercombe Court D, Carracedo A, Lareu MV (2018) Global patterns of STR sequence variation: sequencing the CEPH human genome diversity panel for 58 forensic STRs using the Illumina ForenSeq DNA Signature Prep Kit. Electrophoresis 39:2708–2724. https://doi.org/10.1002/elps.201800117

Pospiech E, Karlowska-Pik J, Ziemkiewicz B, Kukla M, Skowron M, Wojas-Pelc A, Branicki W (2016) Further evidence for population specific differences in the effect of DNA markers and gender on eye colour prediction in forensics. Int J Legal Med 130:923–934. https://doi.org/10.1007/s00414-016-1388-2

Pospiech E, Chen Y, Kukla-Bartoszek M, Breslin K, Aliferi A, Andersen JD, Ballard D, Chaitanya L, Freire-Aradas A, van der Gaag KJ, Giron-Santamaria L, Gross TE, Gysi M,

Huber G, Mosquera-Miguel A, Muralidharan C, Skowron M, Carracedo A, Haas C, Morling N, Parson W, Phillips C, Schneider PM, Sijen T, Syndercombe-Court D, Vennemann M, Wu S, Xu S, Jin L, Wang S, Zhu G, Martin NG, Medland SE, Branicki W, Walsh S, Liu F, Kayser M (2018) Towards broadening Forensic DNA Phenotyping beyond pigmentation: improving the prediction of head hair shape from DNA. Forensic Sci Int Genet 37:241–251. https://doi.org/10.1016/j.fsigen.2018.08.017

Ruff CB (1994) Morphological adaptation to climate in modern and fossil hominids. Am J Phys Anthropol 37:65–107. https://doi.org/10.1002/ajpa.1330370605

Salvoro C, Faccinetto C, Zucchelli L, Porto M, Marino A, Occhi G, de Los Campos G, Vazza G (2019) Performance of four models for eye color prediction in an Italian population sample. Forensic Sci Int Genet 40:192–200. https://doi.org/10.1016/j.fsigen.2019.03.008

Scudder N, McNevin D, Kelty SF, Walsh SJ, Robertson J (2018) Massively parallel sequencing and the emergence of forensic genomics: defining the policy and legal issues for law enforcement. Sci Justice 58:153–158. https://doi.org/10.1016/j.scijus.2017.10.001

Seo HJ, Cho S, Lee JH, Lyoo SH, Kim M-Y, Lee SD (2017) Forensic DNA phenotyping: a review in Korean perspective. Korean J Leg Med 41:23–31. http://synapse.koreamed.org/DOIx.php?id=10.7580%2Fkjlm.2017.41.2.23

Sero D, Zaidi A, Li J, White JD, Zarzar TBG, Marazita ML, Weinberg SM, Suetens P, Vandermeulen D, Wagner JK, Shriver MD, Claes P (2019) Facial recognition from DNA using face-to-DNA classifiers. Nat Commun 10:2557. https://doi.org/10.1038/s41467-019-10617-y

Slabbert N, Heathfield LJ (2018) Ethical, legal and social implications of forensic molecular phenotyping in South Africa. Dev World Bioeth 18:171–181. https://doi.org/10.1111/dewb.12194

Sochtig J, Phillips C, Maronas O, Gomez-Tato A, Cruz R, Alvarez-Dios J, de Cal M-AC, Ruiz Y, Reich K, Fondevila M, Carracedo A, Lareu MV (2015) Exploration of SNP variants affecting hair colour prediction in Europeans. Int J Legal Med 129:963–975. https://doi.org/10.1007/s00414-015-1226-y

Sturm RA, Box NF, Ramsay M (1998) Human pigmentation genetics: the difference is only skin deep. Bioessays 20:712–721. https://doi.org/10.1002/(SICI)1521-1878(199809)20:9<712::AID-BIES4>3.0.CO;2-I

Toom V, Wienroth M, M'charek A, Prainsack B, Williams R, Duster T, Heinemann T, Kruse C, Machado H, Murphy E (2016) Approaching ethical, legal and social issues of emerging forensic DNA phenotyping (FDP) technologies comprehensively: reply to "Forensic DNA phenotyping: predicting human appearance from crime scene material for investigative purposes" by Manfred Kayser. Forensic Sci Int Genet 22:e1–e4. https://doi.org/10.1016/j.fsigen.2016.01.010

Valenzuela RK, Henderson MS, Walsh MH, Garrison NA, Kelch JT, Cohen-Barak O, Erickson DT, John Meaney F, Bruce Walsh J, Cheng KC, Ito S, Wakamatsu K, Frudakis T, Thomas M, Brilliant MH (2010) Predicting phenotype from genotype: normal pigmentation. J Forensic Sci 55:315–322. https://doi.org/10.1111/j.1556-4029.2009.01317.x

Vidaki A, Daniel B, Court DS (2013) Forensic DNA methylation profiling--potential opportunities and challenges. Forensic Sci Int Genet 7:499–507. https://doi.org/10.1016/j.fsigen.2013.05.004

Walsh SJ (2004) Recent advances in forensic genetics. Expert Rev Mol Diagn 4:31–40. https://doi.org/10.1586/14737159.4.1.31

Walsh SJ (2005) Legal perceptions of forensic DNA profiling: Part I. A review of the legal literature. Forensic Sci Int 155:51–60. https://doi.org/10.1016/j.forsciint.2004.11.001

Walsh S, Liu F, Ballantyne KN, van Oven M, Lao O, Kayser M (2011) IrisPlex: a sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. Forensic Sci Int Genet 5:170–180. https://doi.org/10.1016/j.fsigen.2010.02.004

Weinberg SM, Neiswanger K, Richtsmeier JT, Maher BS, Mooney MP, Siegel MI, Marazita ML (2008) Three-dimensional morphometric analysis of craniofacial shape in the unaffected relatives of individuals with nonsyndromic orofacial clefts: a possible marker for genetic susceptibility. Am J Med Genet Part A 146A:409–420. https://doi.org/10.1002/ajmg.a.32177

White D, Rabago-Smith M (2011) Genotype-phenotype associations and human eye color. J Hum Genet 56:5–7. https://doi.org/10.1038/jhg.2010.126

Williams R, Wienroth M (2017) Social and ethical aspects of forensic genetics: a critical review. Forensic Sci Rev 29:145–169

Wolinsky H (2015) CSI on steroids: DNA-based phenotyping is helping police derive visual information from crime scene samples to aid in the hunt for suspects. EMBO Rep 16:782–786. https://doi.org/10.15252/embr.201540714

Wray NR, Yang J, Hayes BJ, Price AL, Goddard ME, Visscher PM (2013) Pitfalls of predicting complex traits from SNPs. Nat Rev Genet 14:507. https://doi.org/10.1038/nrg3457

Zaorska K, Zawierucha P, Nowicki M (2019) Prediction of skin color, tanning and freckling from DNA in Polish population: linear regression, random forest and neural network approaches. Hum Genet 138:635–647. https://doi.org/10.1007/s00439-019-02012-w

Zbiec-Piekarska R, Spolnicka M, Kupiec T, Parys-Proszek A, Makowska Z, Paleczka A, Kucharczyk K, Ploski R, Branicki W (2015) Development of a forensically useful age prediction method based on DNA methylation analysis. Forensic Sci Int Genet 17:173–179. https://doi.org/10.1016/j.fsigen.2015.05.001

Rapid DNA Typing

28

Pankaj Shrivastava, Aditi Mishra, Anand Kumar,
Sumit K. Chaudhary, Sonia Kakkar, and R. K. Kumawat

Abstract

After more than three-decade of the inception of forensic DNA technology, it is accepted today as the most reliable and most believed technology in forensics and criminal investigation. Modern DNA technology is entirely different from the original 35-year-old technology, which is the result of several decadal developments and advancements. Many incremental improvements and continuous progress have made technology fast, robust, and reliable. For handling the increasing load on forensic DNA testing, more rapid techniques have come up in the last decade. The forensic DNA technology of today is discussed with the available faster direct systems globally. The chapter describes the "*Swab in-Profile Out*" options available with present-day forensic DNA typing. Though not successful for all variety of forensic samples, the technology is still getting appreciated and is finding its way in forensic laboratories after initial validation.

P. Shrivastava (✉)
DNA Fingerprinting Unit, State Forensic Science Laboratory, Department of Home (Police), Government of Madhya Pradesh, Sagar, Madhya Pradesh, India

A. Mishra
DNA Fingerprinting Unit, State Forensic Science Laboratory, Department of Home (Police), Government of Madhya Pradesh, Sagar, Madhya Pradesh, India

Raksha Shakti University, Gandhinagar, Gujarat, India

A. Kumar · S. Kakkar
Department of Forensic Medicine, Post Graduate Institute of Medical Education and Research (PGIMER), Chandigarh, India

S. K. Chaudhary
Raksha Shakti University, Gandhinagar, Gujarat, India

R. K. Kumawat
DNA Division, State Forensic Science Laboratory, Jaipur, Rajasthan, India

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_28

Keywords

Rapid DNA · Automation · Capillary electrophoresis

DNA technology has proven to be extremely useful in solving a wide range of criminal and civil cases including disputed paternity, maternity, missing person etc. (Gold 2013). The forensic DNA technology which was invented by Alec Jeffreys is no more being utilized in routine forensic DNA analysis (Roewer 2013). On the contrary the capillary electrophoresis, Short Tandem Repeat (STR) and Polymerase Chain Reaction (PCR) based present DNA technique is Gold Standard (Gold 2013). This Gold standard technology is being utilized for the laboratory examination involves different steps including DNA extraction, quantification, making appropriate dilutions as per the requirement of multiplex kit being used followed by PCR amplification, genotyping and data interpretation. The present day capillary electrophoresis based, multistep forensic DNA technology is described in Fig. 28.1.

Present day forensic DNA typing methods requires multiple expensive instruments along with highly skilled manpower for execution of different steps and analysis. Depending upon the type and nature of sample the forensic DNA analysis may take time from few hours to a whole day or few days. Though, automation in the process of forensic DNA typing has made the process simpler. Now a day's automation is available for DNA extraction. All the three leading brands Thermo Fisher Scientific, Promega and Qiagen are already marketing their forensic validated automated DNA extraction systems. Besides this many liquid handing systems are available now for making the process simpler and automated. With the increasing demand of forensic DNA testing most of the forensic DNA laboratories are overburdened and facing huge backlogs. Also there have been constant pressure from law-enforcement side to provide faster and time bound

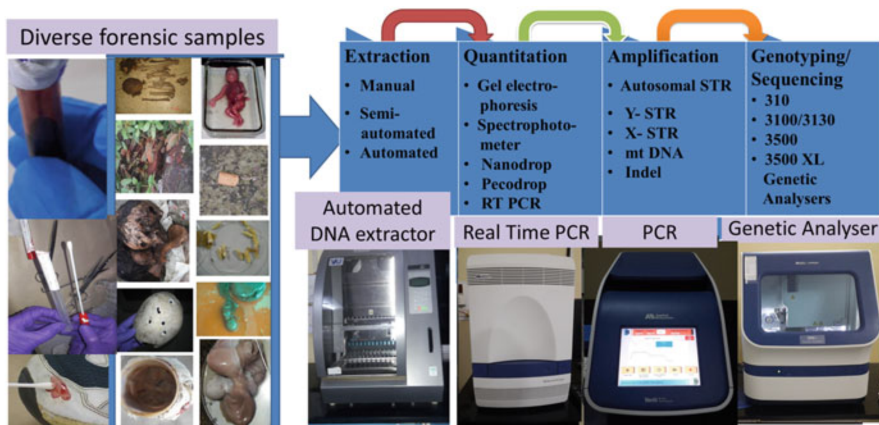


Fig. 28.1 Steps in present day forensic DNA technology

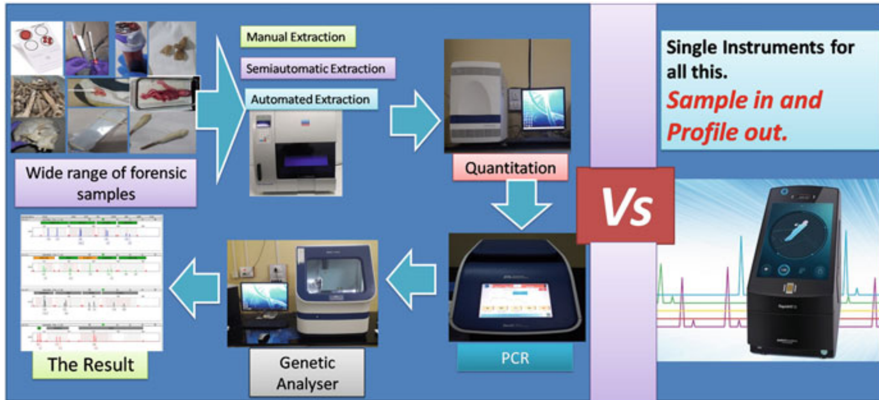


Fig. 28.2 Current practice in Forensic DNA analysis vs Rapid DNA technology

reports from the target samples (Lee 2017). In order to meet the expectations of law and enforcement, fully automated DNA technology is being demanded by many forensic laboratories. This technology is the need of the hour when there is a huge backlog in forensic laboratories across the world. This demand resulted in to the development of instruments which can do all the analysis and provide result in minimum time. The commercialization of an automated system that brings the extraction, amplification, separation and analysis steps together in one platform has been unrealized until recently. Instruments and kits that support the automated DNA process are now available to the few forensic laboratories.

Rapid DNA technology is a fully automated (hands free) process of generating STR profile from buccal swab reference samples. The "swab in—profile out" method (portable) is integrated with the automated DNA extraction, STR amplification, Amplicon separation, detection and allele calling without human involvement which minimizes the analysis time and also cross-contamination of samples (Fig. 28.2) (Gold 2013).

In December 2010, Rapid DNA Program Office was set up by the Federal Bureau of Investigation (FBI) to adopt the rapid DNA technology by law enforcement and other government agencies (https://www.fbi.gov/). Further, many amendments have been considered to the DNA Identification Act of 1994 to utilize the rapid DNA technology in forensic fields (Lee 2017). The Rapid DNA Act of 2017 was signed on August 18, 2017, which amends the DNA Identification Act of 1994. This law gives permission to use fully automated rapid DNA technology or instruments to the various laboratories (Gold 2013). The development of rapid DNA as an advanced technology has many advantages such as:

- It takes less time for analysis
- It is less expensive
- It has minimum risk of sample being contaminated

- It is easy to use (portables) in the field
- It can reduce the case backlogs
- Compact and transportable
- Simple and reliable
- Advanced micro fluids technology

The advanced rapid DNA technology works with computer based software and comes with the pre-filled cartridges. These pre-filled cartridges contain all the chemical reagents required for the further processing such as DNA extraction, amplification, separation and genotyping. These cartridges also contain sample input slots in which forensic samples are loaded. After the sample loading into the instrument, analysis proceeds automatically. The analysis operation is self-calibrating which allows semi-skilled person to handle this equipment with an hour of training. Besides this, the machine allows identifying the criminal (s) within 75–90 min and this helps in exonerate the suspect(s) in very limited time (Roewer 2013).

The developments that helped to advance rapid DNA technology to the prevailing situation are various; but incorporation of advanced micro fluidic principle into rapid DNA system makes it more attractive. The micro-fluidic technology uses fewer amounts of chemicals to minimize the cost, instrument size, reaction time and also size of the cartridges containing chemical. The complete DNA processing steps like extraction, amplification (also commonly called as PCR; Polymerase Chain Reaction), separation (also known as capillary electrophoresis), and interpretation can be achieved with one micro-fluidic chip (Gold 2013). The rapid DNA systems comprise three integral parts:

- Swab
- A chip
- The instrument

The swab is needed for the input of desired quantity of forensic samples to the system for analyses. The chip bears all the materials, regents and waste containments required for the process in order to generate STR profile i.e. reagents for DNA extraction, PCR to capillary electrophoresis. There is no direct contact between the instrument and the sample swab. The sample injection is by pneumatic pressure to avoid any chance for contamination. The steps in processing of samples for generation of DNA profile are shown in Fig. 28.3.

The STR profiles generated from the rapid DNA are compatible to the new CODIS loci, Expanded European Standard Set, National DNA profile databank (United Kingdom), National Criminal Investigation DNA database (Australia), National DNA Data Bank (Canada), National DNA Database (China), Data-Analyse-Datei (Germany), and National DNA profile databank (New Zealand) (Carney 2019).

Rapid DNA analysis can be useful in identification of criminals where they can only be kept for not more than 2 h as it provides faster results. In future, the use or

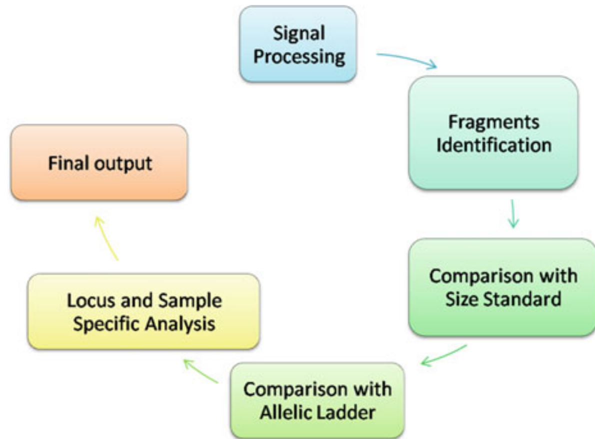


Fig. 28.3 Working of Rapid DNA instruments

availability of such technology can be expanded for acquittal and conviction of suspects in the cases when time is critical. It will allow rapid and more efficient investigations. Rapid DNA technology has shown its efficiency from the forensic laboratories to law enforcement agencies to airport and immigration centers. This Rapid DNA machine is also known as DNA "magic box" by law-enforcement authorities. For the cases of paternity one need not to wait for a day even. Despite of being a costly technology, due to ease of processing and faster results the technology is attracting forensic DNA agencies globally.

28.1 Limitations of the Rapid DNA Technology

The innovation of Rapid DNA technology has been extremely beneficial to the law enforcement agencies and forensic laboratories. As discussed above, it has shown many benefits to the numerous cases. Above all the benefits of the rapid DNA technology, it has some limitations too such as:

- It supports very few sample types. All variety of forensic samples especially tough samples like bone and teeth cannot be processed using this technology.
- It requires pre-processing for many forensic samples.
- It does not give 100% yield rate on all the tested samples.
- In comparison to traditional methods, it is still not affordable for many investigating agencies.
- It still requires proper standardization and validation for the wide variety of forensic samples.
- In many forensic cases the samples are degraded either due to their recovery from adverse environments or due to their late processing for one or the other reason.

Table 28.1 Various details regarding Rapid DNA Analysis instrument

Manufacturer details	Instrument	Chemistry	Time required
Network Biosystems	DNAscan/ ANDE4C DNAscan6C	PowerPlex 16 FlexPlex 27	84 min (Turingan et al. 2016)
IntengenX	RapidHIT 200	GlobalFiler Express	90 min (Thong et al. 2015)
IntengenX	RapidHIT ID	AmpFℓSTR®NGMSElect™ Express	90 min (Salceda 2017)
NEC	Portable DNA Analyzer	–	25 min
LGC	Para DNA® Instruments	HyBeacons™ Biochemistry	75 min (Blackman et al. 2015)

With the advancement in this technology various instruments have been designed by different manufacturer (Table 28.1).

28.2 RapidHIT and RapidHIT™ 200

A previous generation of Rapid DNA technology was developed to address this issue and has enabled integrated sample-in to CODIS-ready profile-out workflow. The RapidHIT® 200 and RapidHIT® ID Systems were developed by California based company known as IntegenX. This is a straight out of box solution which enables DNA profiling to be carried out quickly and easily by normal scientific staff. Both the systems are validated by National DNA Index System (NDIS) and utilize their approved software and chemistry (Roewer 2013).

The first organization to received rapid DNA technology (RapidHIT™ 200 by IntengenX) system was the US Police department of Palm Bay, Florida. These are portable systems which enable law enforcement agencies to develop DNA profile from the reference buccal swab samples (Date-Chong 2016). These systems do not require any certified forensic scientist to work with them. The instrument required minimal training to be used effectively. TheRapidHIT™ 200 instrument is an automated and integrated system which can generate complete STR based DNA profiles within 2 h i.e. from extraction to interpretation. The RapidHIT™ 200 System is plug and play technology that requires no calibration, testing or commissioning. Its ultimate flexibility and the use of global standard STR chemistry means it can be deployed and network across wide geographical area anywhere in the world. The RapidHIT™ 200 System processed the sample and then network to the cloud for cross-referencing to the FBI CODIS's database. This instrument is validated for NDIS-approved GlobalFiler® Express chemistry that includes 21 autosomal STR loci and three sex determining markers (Date-Chong 2016). The RapidHIT™ 200 can accommodate one to eight samples in one run. The cartridges can run five buccal samples, one positive, one negative and an allelic ladder. This instrument has

two sample cartridges for analysis and both of them can be run subsequently. After the cartridges are loaded into instrument and the run has been initiated, the RapidHIT™ 200 uses the DNA IQ™ System from the Promega Corporation to isolate DNA from the buccal swabs. The RapidHIT™ 200 instrument does not require quantification of the samples as the DNA IQ™ method restrict the amount of extracted DNA sent to the PCR chamber (4–20 ng) of input DNA template (Shackleton 2019).

28.3 RapidHIT® ID

The RapidHIT® ID is next generation instrument which processes presumed single source samples to generate CODIS-compatible STR based DNA profiles in less than 90 min. The instrument operates with less-expensive single-use cartridges for each sample tested. The sample cartridges accommodate two types of multiplex kits: Amp*Fl*STR® NG and GlobalFiler® Express. Allelic ladders specific to the kit-dependent cartridges are loaded in the instrument (Buscaino 2018). The cartridges chambers and valves are loaded with the bulk reagents and presumed DNA samples for further processing to generate the DNAprofiles. The RapidHIT system transfers these data to the main computer which is linked with RapidLINK software (IntegenX). The STR data shall be analysed using RapidLINK software and manually also if needed. RapidLINK software also manages reagent supplies and operator access across multiple-geographical network of DNA devices (Fig. 28.4).

The RapidHIT system is slightly smaller than NetBioDNAscan/ANDE system. The system can run up to eight samples at a time. Control cartridges are also loaded along with allelic ladder. RapidHIT® ID comes along with the camera and finger print reader to authenticate the access and for audit tracking (Hennessy et al. 2014). In order to process the buccal swabs, RapidHITACE sample cartridge (IntegenX) is

Fig. 28.4 RapidHIT® IDSystem

used by RapidHIT® ID. However, a second sample cartridge has recently designed by IntegenX for the RapidHIT® ID system to process the extracted DNA samples encountered at crime scene (Buscaino 2018). The complete DNA processing steps like extraction, amplification (also commonly called as PCR; Polymerase Chain Reaction), separation (also known as capillary electrophoresis), and interpretation can be done with RapidHIT® ID system.

28.4 DNAscan/ANDE™ Rapid DNA Analysis System

The DNAscan/ANDE™ Rapid DNA Analysis™ based an integrated Accelerated Nuclear DNAEquipment (ANDE) was developed by Massachusetts based Network BioSystems (NetBio's). It is a compact fully automated system that analyses DNA in less than 85 min. The system is designed for use in police booking stations and forensic laboratories for the analysis of reference samples only. This system works with the combination of three components: ANDE swab; a Chip (for buccal swab processing) and ANDE instrument. It is used with two accessories DNA scan single-use biochip cassettes and DNA scan collection kit with net bio collection swabs (Moreno 2017). The DNAscan/ANDE™ Rapid DNA Analysis™ can run five reference samples simultaneously. The fully integrated DNA biochip cassettes contain all of the necessary components for DNA analysis. All Components are stable at room temperature for 6 months so there is no need for refrigeration or freezer storage. The necessary reagents like DNA extraction, buffers, FlexPlexSTR amplification reagents and polymer shall be loaded pre-filled on the chip and amend for the micro fluidic processing methods and directed by pneumatic pressure to ensure the consistency and accurate results (Moreno 2017). The instrument is not altogether in direct contact with sample or the reagent. The specialized RFID (radio frequency identification) DNA swabs lock into the biochip set cassettes to eliminate the potential of sample mix-up and to preserve chain of custody. The locking RFID system minimizes the potential for sample switch. The bio chips have multiple areas in which the steps of DNA analysis take place. Five samples chambers allow for five samples to be processed at once. The sample chambers are part of the smart cartridge which contains all components for DNA purification. The integrated bio-chips contain all of the reagents to perform PCR and the PCR chambers where the target DNA fragments are amplified. The gel cartridges contain the separation medium and buffer. The instruments interfaces with the biochips set. The analysis and the interpretation of the data are performed by integrated expert system (Software version 1.0.5 and DNAscan/ANDETM Data management system 1.2.0). This system is designed to collect waste material of analysis which can be stored at room temperature (Moreno 2017).The DNAscan collection kit contains all of the components needed to collect the buccal samples. The kit includes barcodes, a mailing envelope, instruction for use, a pair of gloves, NetBio Biochip set Swab and EasiCollect™ device. The analysis and DNAscan/ANDE Rapid DNA identification system is previously compatible with the 4C chemistry based PowerPlex16 system (Promega Corporation, CA). However, upgraded version of this instrument

was specially designed for the detection of six fluorescent dyes (6C) and validated for FlexPlex27 Chemistry. FlexPlex 27 STR Multi-Plexkit has 23 autosomal loci, 3 Y STR loci and Amelogenin loci (Grover 2017). This advanced version of the instrument is referred to as ANDE6C. The additional loci in the FlexPlex27 Multiplexed STR kit having expended set of the CODIS core 20 loci increase the power of discrimination to upgrade the efficacy of kinship analysis and identification matches. Output files shall be encrypted and can be decrypted and exported by using a multi-function user privilege software package known as FAIRS™ (ANDE Corporation).

28.5 Portable DNA Analyzer

Portable DNA analyzer is a fully automated portable instrument manufactured by NEC Corporation. It carries the entire DNA extraction process within 25 min through "single chip technology". The instrument is equipped with the analysis chip and a reagent package (Lyses solution, Wash reagent and Elution reagent) with small quantities of liquid through low cost technologies. The combination of lab-on-chip and reagent package empowers it to be effortlessly worked without pipettes, which are required for existing investigation strategies. NEC's micro-precise "single-chip technology" contains a small plastic chip with 5 mm wells that work as test tubes and fluid-transferring channels that work as pipettes The single layer of resin are coated with layers of silicon film. Portable DNA analyzer is mostly used by Asian countries. The chip is disposable which intrinsically avoids DNA mistyping. The conventional DNA methods carried out all the process i.e. extraction, amplification and genotyping separately however, this analyzer performs each process independently (https://cl.nec.com/es_CL/es/pdf/catalogue.pdf). The smaller sizes also make the connections of each process easier and faster. Hence, the repetitive process of heating and cooling is significantly reduced. This is a world's first portable analyzer from NEC which will be beneficial to speed-up the criminal cases. It is easy to operate and does not need any professional expertise.

28.6 Para DNA® Instruments

Recently, new method of DNA extraction methods has been manufactured by LGC known as Para DNA® instruments to analyze the samples within 75 min. The ParaDNA® instruments can be deployed in a lab or at a crime scene by forensic experts or properly trained non-expert person (e.g. crime scene investigators). The ParaDNA® is quick and robust instrument which generates reliable profile information with high success rate. ParaDNA® instruments also specify whether samples collected at the crime scene contain human DNA and which of them can generate reliable STR profiles for investigations. It consists of two sections; DNA amplification and fluorescent detection software and a consumable kit (sample collector and reaction plate containing required biochemistry for amplification). The ParaDNA®

methods used HyBeacons™ fluorescent probe technology to detect allele's length. This instrument is compatible with five STR loci and amelogenin loci. The ParaDNA® instruments follow Scientific Working Group on DNA Analysis Methods (SWGDAM) guidelines (Blackman et al. 2015). The instrument was tested with touch samples and saliva sample in which saliva has 100% yield rate while touch samples showed 69% yield rate. These data displays that ParaDNA® instruments can be successfully employed in the forensic field as rapid DNA detection system. The addition of ParaDNA® tool shall reduce case backlogs, permits in prioritizing of samples and as well as reducing time and reagents used for analysis. Further testing of a larger sample size is underway as well as additional sample types such as dried blood.

References

Blackman S, Dawnay N, Ball G, Stafford-Allen B, Tribble N, Rendell P, Wells S (2015) Developmental validation of the ParaDNA® Intelligence System—a novel approach to DNA profiling. Forensic Sci Int Genet 17:137–148

Buscaino J (2018) Evaluation of a Rapid DNA process with the RapidHIT® ID system using a specialized cartridge for extracted and quantified human DNA. Forensic Sci Int Genet 34:116–127

Carney C (2019) Developemental validation of the ANDETM Rapid DNA system with FlexPlexTM Assay for arrestee and reference buccal swab processing and database searching. Forensic Sci Int Genet 40:120–130

Date-Chong M (2016) Evaluation of RapidHIT™ 200 and RapidHITGlobelFiler® express kit for fully automated STR genotyping. Forensic Sci Int Genet 23:1–8. https://doi.org/10.1016/j.fsigen.2016.03.001

Gold S (2013) Rapid DNA: a game changer in the law enforcement identification stakes. Biometric Technol Today 2013(1):7–10

Grover R (2017) Flexplex27-highly multiplexed rapid DNA identification for law enforcement, kinship and military applications. Int J Legal Med 131:1489–1501

Hennessy LK et al (2014) Developmental validation of the GlobalFiler® express kit, a 24-marker STR assay, on the RapidHIT® System. Forensic Sci Int Genet 13:247–258

Lee J (2017) The emergence of rapid DNA technology. International Biometrics + Identity Association (IBIA), Washington, DC, pp 1–8

Moreno LI (2017) Internal validation of the DNAscan/ANDE™ Rapid DNAAnalysis™ platform and its associated PowerPlex® 16 high content DNA biochip cassette for use as an expert system with reference buccal swabs. Forensic Sci Int Genet 29:100–108

Roewer L (2013) DNA fingerprinting in forensics: past, present, future. Investig Genet 4(1):22

Salceda S (2017) Validation of a Rapid DNA process with the Rapid HIT® ID system using GlobalFiler® express chemistry, a platform optimized for decentralized testing environments. Forensic Sci Int Genet 28:21–24

Shackleton D (2019) Development of enhanced sensitivity protocol on the RapidHIT™ 200 with a view to processing case work material. Sci Justice 59:411–417

Thong Z et al (2015) Evaluation of the RapidHITTM 200 System: a comparative study of its performance with Maxwell® DNA IQTM/Identifiler® Plus/ABI 3500xL workflow. Forensic Sci Int Genet 19:22–27

Turingan RS et al (2016) Rapid DNA analysis for automates processing and interpretation of low DNA content samples. Investig Genet 7:2

Part VI

Guideline, Quality Control, Legal Aspects and DNA Database in Forensic DNA Typing

Guidelines for Collection and Preservation of Samples for Forensic DNA Testing

29

Pankaj Shrivastava, Shivani Dixit, and Hirak Ranjan Dash

Abstract

Forensic DNA typing is the gold standard in a criminal investigation that has overpowered all other forensic techniques. The materials which are examined in forensic DNA are of biological origin, which is amenable to degradation. Hence, proper preservation and transportation are to be ensured. With the increasing reliance on the DNA technique, it is required that the exhibits being sent for DNA testings should be properly preserved and sent as per the proper guidelines. Also, the chain of custody from collection to transport of material to the DNA laboratory should be ensured. The laboratories practicing forensic DNA have their guidelines that need to be followed for sending the sample to any particular laboratory. But for the benefit of readers, some general guidelines are presented in this chapter.

Keywords

Forensic DNA · Guidelines · Collection · Preservation

Forensic DNA examination has now almost changed the global scenario of Investigation not only for establishing conviction but also to avoid wrongful conviction of innocents (Johnson and Williams 2004). Courts are now waiting for the DNA reports in the cases. On the basis of our experience as forensic DNA experts since last 15 years we have seen dramatic change towards the importance of DNA reports in

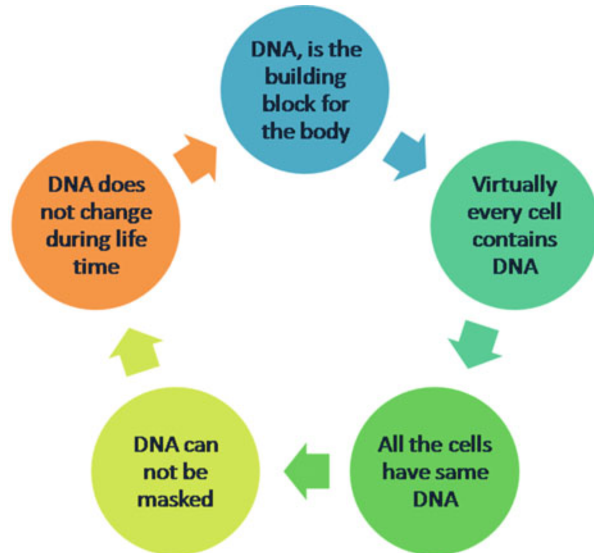
P. Shrivastava (✉) · S. Dixit
DNA Fingerprinting Unit, State Forensic Science Laboratory, Department of Home (Police), Government of Madhya Pradesh, Sagar, Madhya Pradesh, India

H. R. Dash
DNA Fingerprinting Unit, Integrated High-Tech Laboratory, Regional Forensic Science Laboratory, Bhopal, Madhya Pradesh, India

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_29

Fig. 29.1 The reasons for making DNA technology, the method of choice for investigation



criminal cases specially. The peculiar features of DNA have made them the hero in the process of criminal investigation (Fig. 29.1).

Another very important point is the fact that no two individuals have the same DNA profile (except for the identical twins). Due to this fact, any evidence collected from the crime scene can be linked with the suspect or it can be used to eliminate a suspect from any doubt. Similarly in cases of sexual assault the source of body fluid transferred from suspect to victim's body or vice versa is ascertained by using DNA typing (Collins 1998). In case of paternity establishment or identification of mutilated remains, the paternity of concerned is ascertained. The beauty of this technique is the fact that it works with a wide variety of exhibits (Fig. 29.2).

The DNA profile gathered from the evidence sample or from the sample which was lifted and preserved from scene of crime is compared with the reference sample. Also in United States and few other countries having database, this DNA profile can be used to generate hit in to the database like FBI's Combined DNA Index System (CODIS) (Panneerchelvam and Norazmi 2003). For making the effective use of the technology the evidence require proper attention so as to reach the laboratory in a proper way. All the laboratories working on forensic DNA typing work have issued guidelines for proper sampling, preservation and transportation of biological exhibits. Few noticeable guidelines can be accessed from the web-links given in Table 29.1.

The samples in consideration for the DNA analysis can be classified broadly in two categories, i.e. (1) *Reference samples*; (2) *Forensic samples.*

Fig. 29.2 Wide varieties of exhibits received in the laboratory for forensic DNA examination

Table 29.1 Guidelines issued by some reputed laboratories for collection and preservation of DNA evidences

Organization	Link
Federal Bureau of Investigation (FBI)	Handbook of Forensic Services—FBI.gov www.fbi.gov/file-repository/handbook-of-forensic-services-pdf
National Institute of Standards and Technology (NIST)	https://www.nist.gov/topics/forensic-science/evidence-management/biological-evidence-guidance
Central Forensic Science Laboratory (CFSL), Chandigarh, India	http://dfs.nic.in/pdfs/Pamphlet%20.pdf
Centre for DNA Fingerprinting and Diagnostics (CDFD), Hyderabad, India	http://www.cdfd.org.in/images/dna/newinfo28308.pdf
Nepal Police	https://www.nepalpolice.gov.np/images/documents/updates/Dna_Guidelines.pdf
Produced 2014 by the Forensic Working Group (FWG), part of the Partnership for Action against Wildlife Crime (PAW)	https://www.tracenetwork.org/wp-content/uploads/2012/08/Wildlife_DNA_Sampling_Guide_web.pdf

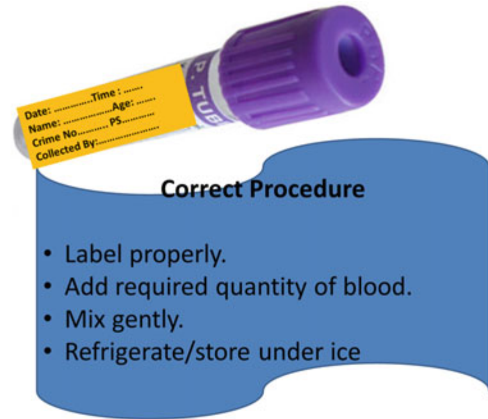


Fig. 29.3 The correct procedure of collection of blood sample in EDTA vial

29.1 Reference Samples

In most of the cases the blood is considered to be the ideal referral sample which is collected from victim as well as from the suspect. Liquid blood is drawn by the medical practitioner after pre-informed written consent as per the DNA examination guidelines (Weedn et al. 1998). In the consent form genetic abnormality (if any) and information about any specific medical condition like blood/organ transplant must be stated and sent to the laboratory along with the sample. 2–5 ml[1] liquid blood samples (standard as per many standard protocols) is collected in pre-sterile EDTA vial. Ethylenediaminetetraacetic acid (EDTA) is a polyprotic acid that chelate Calcium and other ions, thus irreversibly preventing coagulation of blood in the tube (Banfi et al. 2007). 0.5 ml blood should be collected in case of infant/child (in case of paternity establishment). After collection and keeping the blood in EDTA vial, the second important point is to prove the identity of vial. The entire information essential for identification of subject and case viz. name, father's name, case number, name and signature of collector with date should be there in the form of a label as illustrated in Fig. 29.3.

The ideal way is to cover the information by a transparent tape before sealing and preserving the sample for sending to the laboratory. As liquid blood is a biological sample and to avoid any further degradation it is advisable to refrigerate (not frozen) and the sample is transported to the laboratory under ice. Devoid of which, the quality of DNA is decreased due to lysis of WBC (Huang et al. 2017). A suggested method for transport of the referral blood samples is shown in Fig. 29.4. Blood sample can also be collected from dead bodies till 24 h of the last heartbeat by puncturing heart or from the major blood vessel. Postmortem blood should also be

[1]Though with the increased sensitivity of kits and available automations in DNA extraction the requirement of blood samples is less than 100 μl for extracting amplifiable quantity of DNA.

Fig. 29.4 After marking the tube and covering the marking with the transparent tape, the process of keeping sample for transportation

Fig. 29.5 Correct use of a FTA card

collected in EDTA vial. Some of the DNA laboratories also accept the blood samples in pre-sterile blood gauge which authors does not recommend as:

- *It is not expected that all the samples in the laboratory can never be processed immediately*
- *Wet swabs facilitate degradation of sample*
- *Drying of swab is not possible in controlled sterile atmosphere every time*

Alternatively blood sample can also be sent on FTA card (Initially launched by Whatman but now being marketed by Qiagen). For this <125 μl blood sample is transferred on the provided place on the card (Fig. 29.5) (Ahmed et al. 2011). Over-crowding the sample should be provided and the card should be kept for 3–5 min for drying. Each card should be packed separately and the sample on the FTA card can

be transported at room temperature. Blood on FTA card is also required to be sent along with the pre-informed written consent.

In many instances, reference sample is collected in the form of buccal swab. The advantage of using buccal swab is that, it is a non-invasive technique involving less painful and traumatic procedure than venipuncture (Bever and DeGuglielmo 1994). Besides, it is highly useful in newborns, children and adults when blood sample cannot be drawn due to medical or religious grounds. Additionally, buccal swab samples are highly resistant to degradation and the risk of transporting breakable vials containing liquid blood samples is minimized.

29.2 Forensic Samples

29.2.1 Liquid Blood from Scene of Crime

If liquid blood found on any surface then it is recommended to soak up blood in sterile cotton swab/cloth or gauge. Allow it to air dry before packaging in paper envelop, information about exhibit and all details regarding the case should be mentioned on envelope. The written information should be covered with a transparent tape. Liquid blood from crime scene can also be collected either through sterilized syringe or by pipette using pre-sterile disposable tip. After this the liquid can be dispensed in EDTA vial. Blood from multiple spots should be collected and packaged separately to avoid the inter-sample (cross) contamination. In case of clotted blood, it can be lifted with pre-sterile individually packed spatula and can be transported in to a tube. This can also be collected by using sterile cotton swab/cloth or gauge. In collecting sample from a surface, always wear a glove and minimize the touch of surface from where the material is lifted (Kleypas and Gossman 2019).

29.2.2 Wet Blood Stains

If wet blood stains found on any clothing, weapon, small objects or any movable objects which can be moved easily from the crime scene. Then it is recommended to air dry before packaging and then it can be transported to the laboratory. To avoid cross contamination attempt should be made to avoid the overlap of one spot to the other. This can simply be done by separating different stains by keeping paper on different air dried stains. After air drying the item, it should be packaged in paper bag, marked, labeled sealed and then should be transported to the laboratory. Multiple items should be packaged separately to avoid cross contamination (Magalhães et al. 2015).

Wet blood stains from large or immovable objects like wood, surface, wallpaper should be collected by soak up blood through sterilized cotton swab. Swab should be air dried before packaging. Swabs from multiple stained areas should be collected separately to avoid contamination. If wet blood stains are present on large objects

which are difficult to move then it's recommended to take swab or else the object can be cut leaving at least 6 in. space from the periphery of the stain. The stain should be covered with a paper; the paper should be fixed with the use of adhesive tape to avoid the sample loss. The stained area should be allowed to air dry before covering it with paper and cutting it before packaging.

29.2.3 Dried Blood Stains

If dried blood stains found on clothing, weapon, small objects or any movable objects then it is recommended to preserve whole item directly. The collected item should be marked and labeled with the requisite information. Dried blood stains from large or immovable objects can be scraped on a clean paper. If object cannot be scraped then it is recommended to take the swab of the suspected area. Allow to air dry swab before packaging. If the stain is on large object the suspected area can be cut and sent for examination. Blood spatters can be collected through tape lifting from possible areas. Blood samples can also be found in many other possible areas like snow, water etc. if blood stains found on snow then a bigger scoop of snow should be taken and transferred in pre-sterile container. Container should be freeze after taking the sample. If blood stains found in swimming pool, pond or anywhere in still water then collect sample by using sterilized syringe. Transfer sample in a container and then it is recommended to freeze.

29.2.4 Semen, Seminal Stains and Samples from Victims of Sexual Assault

29.2.4.1 Liquid Semen

Many times liquid semen founds on objects and different surfaces at crime scene which will be collected by using sterile cotton swabs. Swabs should be air dried before packaging. Swabs from multiple surfaces or objects should be packaged separately. Source of Victim's body can be swabed for presence of semen by using pre-sterile cotton swab and then can be placed in single use pre sterile vial. Allow to Air dry swab before packaging. Keep swab refrigerated. Swabs collected from multiple body parts should be packaged separately (Magalhães et al. 2015). If seminal stains found on clothing, pillows, bed sheet or any movable object then collect items directly. If any item has wet stains then item should be air dried thoroughly before packaging. Package multiple items separately in paper bags.

If seminal stains found on large objects which cannot be remove from crime scene like carpet then cut stained area and put in a clean paper. Sample must be air dry if item has wet stains. If seminal stains are present on surface or any object which cannot cut then scrape stains on a clean paper otherwise collect stains by using moistened cotton swabs. Every item should be marked properly, labeled then sealed. Depending upon history of case the requisite samples from victim like oral swab, vaginal swab, anal swab, cheeks swab, breast swab should be preserved and sent for DNA testing after drying and individual packing. Commercially available

specialized swabs can also be used for this purpose. Victims clothing can also be preserved and sent for the examination as per requirement in the case. It should be insisted to get medical examination of victim done at earliest to avoid the possible loss of evidence due to change atmospheric pH, enzyme activity, washing and natural cleaning due to urination in the victim's body.

29.2.5 Tissues, Body Parts or Organs

It is difficult to send these samples in dried state, hence if these samples are found on scene of crime then should be picked by sterile forceps or spatula and preserved separately in pre-sterile container in normal saline. To prevent microbial decay it is also suggested to use antibiotics (0.1 mg/ml gentamicin solution). If the sample is being preserved from a postmortem body with varied state of decomposition, it is advisable to collect deep muscle tissues (preferable from more than one organ) by the same procedure. Depending upon the state of degradation of body thigh muscles, heart and brain have been the samples of preference for forensic DNA examination purpose. If it is not possible to send these samples for forensic DNA examinations then its recommended to keep samples under refrigerated condition. If uterus and fetus are to be preserved for establishment of paternity then is highly recommended to preserve both separately to avoid any chance of contamination during possible decay. Tissue or organ samples are never preserved in formal in as it causes DNA damage as well as its known PCR inhibiting capability (Zimmermann et al. 2008).

29.2.6 Bone, Skull, Teeth, or Skeletal Remains

These evidences are already preserved due to their calcified structure (Westen et al. 2008). These samples do not need any preservative and hence can be transported in paper packing at room temperature. Proper information about source and place from where it is collected including the date should be mentioned on the label. For matching of DNA profile of these items, referral samples are also required to be sent as per the guidelines to establish relationship.

29.2.7 Hair

Hairs found in the feast of deceased, on the scene of crime, attached with the clothes of deceased can provide a good link of suspect with the crime. With the advancement in the technology DNA profile can be generated even with a single rooted hair. These samples also do not need any preservative and hence can be transported in paper packing at room temperature. Proper information about source and place from where it is collected including the date should be mentioned on the label. Cut hairs are not normally suitable for DNA examination as they lack nuclear DNA however mtDNA can be extracted from these hairs (Lebedeva et al. 2000) (Table 29.2).

Table 29.2 Quick reference for collection, preservation and transportation of biological samples for forensic DNA examination

Sample preservation guideline at a glance		
Type of sample	How to preserve	How to send/ transport
Blood sample	In EDTA vial	Under ice within 48 h of collection
Blood sample	Whatman filter paper no. 3/FTA	Room temperature
Buccal swabs	Pre-sterile individually packed. Wrapped in a dry paper	Room temperature
Teeth	Wrapped in a dry paper or cloth Preferably molars (without any clinical filling)	Room temperature
Hair with root	Wrapped in a dry paper	Room temperature
Nails	Wrapped in a dry paper	Room temperature
Bones (intact)	Preferably femur wrapped in a dry paper or cloth In case where the bone is in pieces then the largest piece should be sent, however, yield of good quality DNA from the pieces of bone is uncertain owing to the microbial contamination *Completely burnt bones and ash are not useful for DNA analysis*	Room temperature
Vaginal/anal swabs	Pre-sterile individually packed Or dry cotton swab should be placed in a sterile dry glass vial/bottle	Room temperature
Vaginal/anal smear slide	Should be air dried and sent unstained. Swab used to make slide should also be sent in a separate container	Room temperature
Breast swab/finger swab or any other swab	Should be air dried and sent unstained	Room temperature
Penile swab	Use pre-sterile individually packed or Dry cotton swab should be air dried and sent	Room temperature
Condoms dry	In normal paper packing	Room temperature
Condoms wet	Should be tied and sent as soon as possible	Under ice within 48 h of collection
Muscle tissues (about 100 g)/fetus	Should be sent in a clean glass bottle/plastic container with 0.9% DNS (*normal saline sold in medical shops*) on ice or in a crystal salt (sodium chloride) as a preservative Tissue/foetus samples should *never be preserved in formalin*	Should be transported under ice
Bloodstains	In normal paper packing	Room temperature
Weapons	In normal paper packing	Room temperature

(continued)

Table 29.2 (continued)

Sample preservation guideline at a glance		
Type of sample	How to preserve	How to send/ transport
Bidi/cigarette stubs	In normal paper packing	Room temperature
Clothing	In normal paper packing	Room temperature

Dont's

- Do not use direct sunlight, hot air blower, or heater to dry.
- Do not mix stains from different locations.
- Do not swab with bare hands.
- Never preserve any sample in formalin.

Important to Know with Present Day Technology

- Completely charred bones and/or ash are not suitable for DNA examination.
- Cut hairs are normally not suitable for DNA examination.
- Contact of soil with any material may be problematic during forensic DNA analysis.

References

Ahmed HA, MacLeod ET, Hide G, Welburn SC, Picozzi K (2011) The best practice for preparation of samples from FTA®cards for diagnosis of blood borne infections using African trypanosomes as a model system. Parasit Vectors 4:68

Banfi G, Salvagno GL, Lippi G (2007) The role of ethylenediaminetetraacetic acid (EDTA) as in vitro anticoagulant for diagnostic purposes. Clin Chem Lab Med 45:565–576

Bever RA, DeGuglielmo MA (1994) Development and validation of buccal swab collection method for DNA testing for paternity testing. In: Bär W, Fiori A, Rossi U (eds) Advances in forensic haemogenetics, Advances in forensic haemogenetics, vol 5. Springer, Berlin, Heidelberg

Collins KA (1998) The laboratory's role in detecting sexual assault. Lab Med 29:381–385

Huang LH, Lin PH, Tsai KW, Wang LJ, Huang YH, Kuo HC, Li SC (2017) The effects of storage temperature and duration of blood samples on DNA and RNA qualities. PLoS One 12:e0184692

Johnson P, Williams R (2004) Post-conviction DNA testing: the UK's first 'exoneration' case? Sci Justice 44:77–82

Kleypas DA, Gossman WG (2019) Evidence collection. In: StatPearls [Internet]. StatPearls Publishing, Treasure Island, FL. Available from: https://www.ncbi.nlm.nih.gov/books/NBK441852/

Lebedeva IA, Kulikov EE, Ivanova NV, Poltaraus AB (2000) Mitochondrial DNA sequencing of human hair shafts stored for long time. Sud Med Ekspert 43:9–14

Magalhães T, Dinis-Oliveira RJ, Silva B, Corte-Real F, Vieira DN (2015) Biological evidence management for DNA analysis in cases of sexual assault. Sci World J 2015:1. https://doi.org/10.1155/2015/365674

Panneerchelvam S, Norazmi MN (2003) Forensic DNA profiling and database. Malays J Med Sci 10:20–26

Weedn VW, Rogers GS, Henry BE (1998) DNA testing in the forensic laboratory. Lab Med 29:484–489

Westen AA, Gerretsen RRR, Maat GJR (2008) Femur, rib, and tooth sample collection for DNA analysis in disaster victim identification (DVI). Forensic Sci Med Pathol 4:15–21

Zimmermann J, Hajibabaei M, Blackburn DC, Hanken J, Cantin E, Posfai J, Evans TC Jr (2008) DNA damage in preserved specimens and tissue samples: a molecular assessment. Front Zool 5:18

Quality Control in Forensic DNA Typing

30

Ashish Badiye, Neeti Kapoor, Prachi Kathane, and Ritesh K. Shukla

Abstract

Forensic DNA analysis may be defined as the process of identification and individualisation of biological evidence for legal proceedings using DNA technology. It is used in both criminal and in civil cases. Due to dissimilarity in every crime scene and unpredictability of DNA samples collected from the crime scene, the analysis in Forensic Science Laboratories is a tough job for the DNA analyst/ expert. This chapter briefly discusses the various aspects of 'quality control' in DNA forensics. The quality control (QC) in DNA testing is not limited to the quality of the testing laboratory but has to be taken into consideration during every step of the investigation. The chain of custody should be maintained throughout the process. The path of the DNA evidence from the crime scene to the courtroom is quite lengthy and intricate. Different QCs along the whole process shall assist in providing justice by eliminating the chances of errors and thus increasing the admissibility in the court of law.

Keywords

Quality control · Chain of custody · Accreditation · Proficiency testing

A. Badiye (✉) · N. Kapoor
Department of Forensic Science, Government Institute of Forensic Science, Nagpur, Maharashtra, India

P. Kathane
School of Forensic Science & Risk Management, Rashtriya Raksha University, Gandhinagar, Gujarat, India

R. K. Shukla
Biological and Life Sciences, School of Arts and Sciences, Ahmedabad University, Ahmedabad, Gujarat, India

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_30

30.1 Introduction

In criminal and civil trials, DNA is one of the most potential and prime evidence that plays a vital role in the framing of judgement. Personal identification, determination of maternity and paternity, and knowing the involvement of personnel, among others are possible with the invention DNA typing technique. Compared to various others, the DNA technique in some ways is more authentic. It is also considered as 'new gold standard' of scientific evidence and at times even circumvent other established evidence such as fingerprints. During the middle 1980s, the DNA technique was introduced by Sir Alec Jeffreys and his team of the University of Leicester, UK (Lynch 2003; Ansell 2013).

Forensic DNA analysis may be defined as the process of identification and individualisation of biological evidence for legal proceedings using DNA technology. It is used in both criminal and civil cases. Criminal cases mostly involve aggravated assault, murder, rape, arson, grievous hurt, robbery, etc. In contrast, civil cases may include paternity disputes, proof of kinship, custodial disputes, etc. which may involve DNA typing. It helps establish a link between the suspect, victim and the crime scene (National Research Council (US) Committee 1992). In every case, evidence should be handled on a priority basis. It is imperative for every individual coming in contact with the exhibits to maintain the integrity and authenticity of the evidence. Maintaining the high and uniform standards of quality is imperious.

Due to dissimilarity in every crime scene and unpredictability of the type of DNA samples collected from the crime scene, the analysis in Forensic Science Laboratories is a tough job for the DNA analyst/expert. The quality control (QC) and quality assurance (QA) processes play a significant role by providing consistency in results, upgradation of the procedure and tools used for DNA analysis, enhancement of expert skills and maintenance of instruments employed for DNA examination. The need and broad scope of quality control in DNA forensics is illustrated in Fig. 30.1.

'*Quality control*' is the term associated with maintaining the quality of every small entity that proves its adherence to the said set of programs. The ISO 9000 defines quality control as 'a part of quality management focused on fulfilling quality requirements (requirements related to quality)' (ASQ n.d.). So, it applies to every section of testing where quality is the prime motto. Whereas, '*quality assurance*' is referred to as 'the systematic and premeditated approach to build confidence among the service and product, which is required for giving quality result'.

In DNA testing laboratories, quality control is critical because of the vulnerability to produce false (positive/negative) results. Forensic DNA analysis begins with the collection of biological evidences and continues until the report is finally admitted in the court proceedings and has withstood the scrutiny of the defence and the court. When considering the legal purview, the evidence is never treated directly or sent for DNA testing. It has to go through the pipeline before reaching the DNA testing laboratory. In many cases, it first passes through preliminary examination and then after several procedures, at last, is sent for DNA testing. Hence, the quality control in

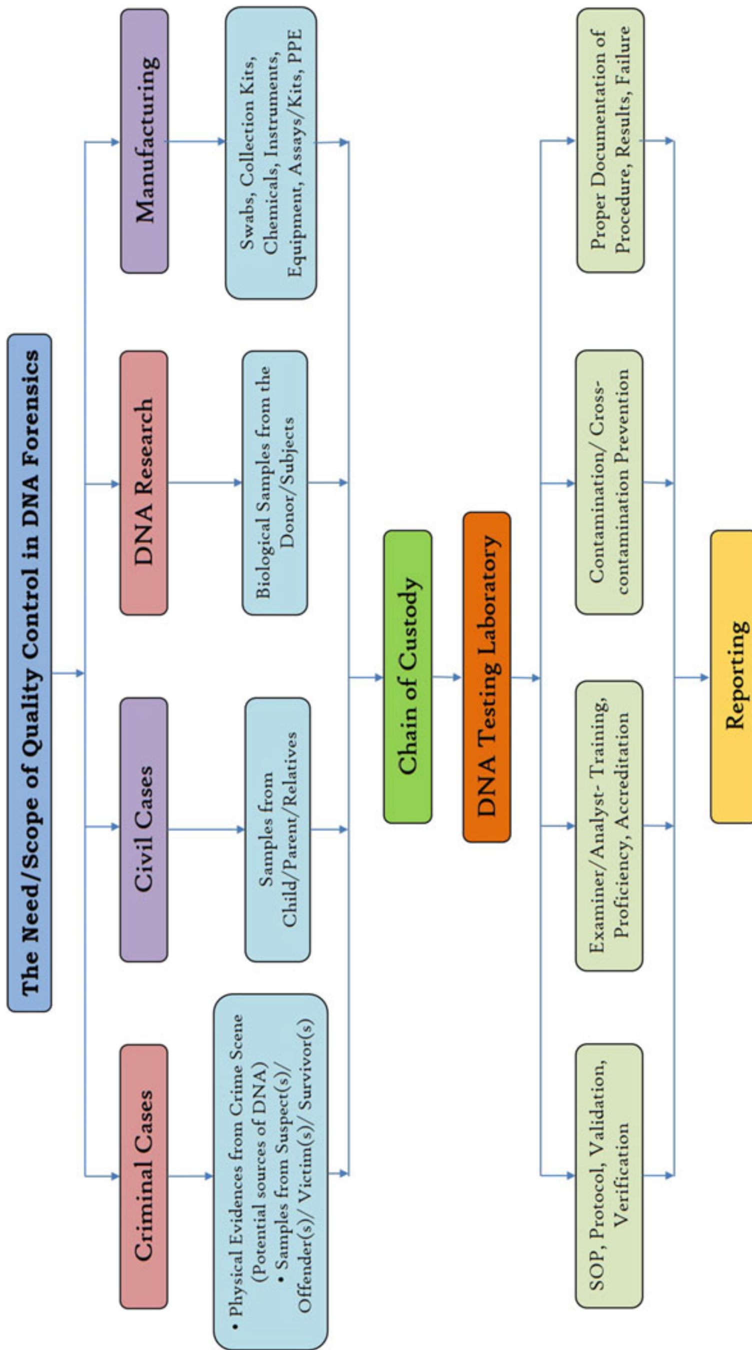


Fig. 30.1 The need and scope of 'quality control' in DNA forensics

DNA testing is not limited to the quality of the testing laboratory but has to be taken into consideration during every step of the investigation. Lacking quality control can lead to the conviction of an innocent and exoneration of the guilty (Balažic and Zupanič 1999). The defence counsel raises not only the question on the DNA typing technique but also the procedure and methods used for DNA collection, preservation, documentation, chain of custody, reporting and even the proficiency of the analyst/expert (to name a few).

The '*quality control*' (QC) is the check on routine operational technique followed by an expert for conducting the DNA analysis. While 'quality assurance' (QA) is referred to as the systematic and premeditated approach to build confidence in the testing to provide a consistent quality result. Each DNA testing laboratory has a set of pretested protocols that needs time to time progression/updating and regular validation of methods and instruments for reliable results. For consistency and transparencies in result, every DNA lab should undergo a regular audit, proficiency test, expert and routine training programs for DNA experts, and along with the accreditation of DNA laboratories (Butler 2005). 'Validation' is described as the procedure followed by the personnel in a laboratory is capable of producing a vigorous, trustworthy and reproducible results (Butler 2005).

The '*proficiency test*' is an assessment of a laboratory's performance in conducting DNA analysis procedures. These tests may be conducted periodically, or on a semi-annual basis, for each DNA examiner in the testing laboratory. The 'ISO/IEC 17043:2010 Conformity Assessment—General Requirements for Proficiency Testing' standard is used by internationally recognised accreditation bodies to accredit proficiency test providers. This standard also provides the purpose of proficiency testing. As per the guidelines of DNA Advisory Board Standard 13.1, this proficiency test can be performed by two ways, i.e., internal proficiency test and external proficiency test. In internal proficiency test, the DNA analyst has to undergo a performance test conducted by some other DNA expert within the lab/organisation/company. In external proficiency test, the proficiency of the analyst examined by some other experts belonging to a different organisation. The most effective way of external proficiency test is when it is performed without the knowledge of DNA analyst being examined. This type of proficiency test is termed as a 'blind external proficiency test'. According to the guidelines, every examiner/analyst handling DNA examination should undergo a proficiency test within 180 days of service. The German DNA profiling group (GEDNAP) has built up a productive and successful blind proficiency-testing program (DNA Box 16.2) (Butler 2005). Further, the analyst engaged in DNA analysis must endure a proficiency test based on new Laboratory Validation Error 393 in typing. This Laboratory Validation Error 393 in typing was proposed by DNA Advisory Board Standards, which was then recognised as the National Standard in the United States (Butler 2005). In India, NABL provides the guidelines for the 'Proficiency Testing Providers' in accordance with ISO/IEC standards.

The '*laboratory audit*' is one of the crucial aspects. It is a systematic examination/audit that may be performed by the laboratory management (internally) or may be conducted by calling experts from another independent organisation/laboratory. It is

a substantial and obligatory requirement for the DNA laboratory to keep the audit record and the course of action undertaken to resolve the points/problems raised by the audit committee.

The '*laboratory accreditation*' is the third-party verification related to the predetermined protocols, instrument performance, personal proficiency and regular audits of a DNA laboratory. Accreditation of a laboratory is the resulting outcome from the successful inspection by a recognised body of a country. It is essential to maintain good lab practices. The 'accreditation process' generally involves several steps such as a 'laboratory self-evaluation, filing the application and supporting documents to initiate the accreditation process, on-site inspection by a team of trained auditors, an inspection report, and an annual accreditation review report'. The inspection evaluates the facilities and equipment, the training of the technical staff, the written operating and technical procedures and the casework reports and supporting documentation of the applicant laboratory (Butler 2005).

30.2 Quality Control at the Crime Scene

Quality control of the samples starts from the crime scene itself. The samples for DNA testing may be collected from crime scenes, victims, suspects and hospitals. When collecting the samples, one should follow the proper method, as making mistakes at this point will cause the failure of the entire procedure.

DNA evidence is collected as blood, hair, bone, teeth, seminal stains, other body fluids, etc. [7]. The type of material used for collecting this evidence also plays a role in quality control. The material used for the collection process should be completely free from any foreign or extraneous substance. The material should be completely sterile, and its quality should be checked before using it to avoid contamination of the sample. This can be achieved by manufacturing the material, especially for collecting DNA evidence under proper guidance. The cotton swab, filter paper, FTA paper modified cellulose, foam, nylon, polyester and rayon-tipped swabs used for collecting the evidence should be kept separately and labelled thoroughly (Eugene 2011). Every sample should be handled carefully to avoid contamination of original evidence with biological and non-biological samples. Contamination of evidence with non-biological origin sample may cause failure in results, whereas contamination with biological origin sample will change the result completely. It will result in a mixed genotype which will further cause problems in the interpretation of the result. This can be avoided only by handling the evidence with precautions, wearing gloves and avoiding mixing of the evidence by keeping them separate from the time of collection itself (Balažic and Zupanič 1999). Care should be taken to avoid contamination as well as cross-contamination.

The investigating officer while collecting evidence from the crime scene should follow some practices like, when the wet sample is collected from the crime scene, it should be air-dried at room temperature only, and the dried sample should be placed in the sterile containers. The wet sample must not be 'blown-upon' in an attempt to hasten the drying process, as it might result in the contamination of the swab by the

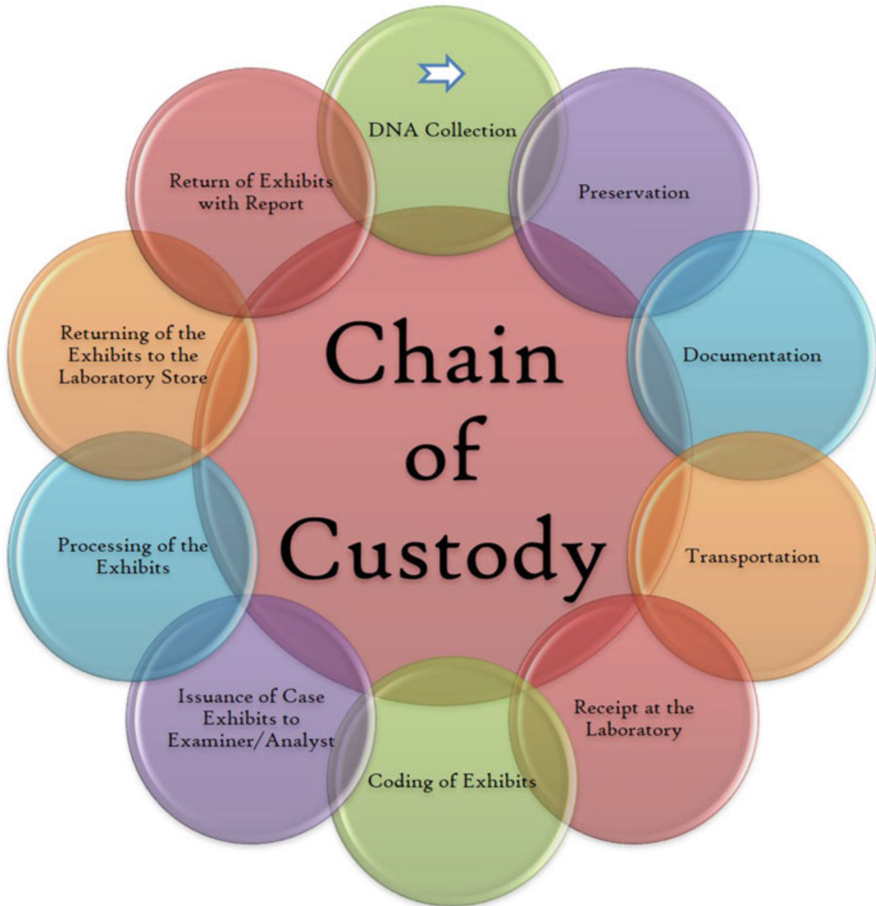


Fig. 30.2 'Chain of custody' of DNA exhibits

saliva coming out from the mouth. The control sample, collected from the victims or suspects, should be stored in the tubes containing anticoagulant EDTA (ethylenediaminetetraacetic acid), which inhibits the sample degradation (Li 2015). The syringe used for the collection of samples should be sterile and new. Every possible effort should be made to avoid contamination of the sample during its collection, preservation and handling until it reaches the lab. A chain of custody of evidence log should be maintained. It is necessary that the chain of custody document has the record of every transfer made from one person to another person, from the moment the evidence is collected. This helps in establishing beyond doubt that nobody else could have accessed the evidence without authorisation. It also serves as a log of possession or chain of evidence (Badiye et al. 2020) (Fig. 30.2).

30.3 Quality Control in the Laboratory

Every DNA testing laboratory should follow Quality Assurance Standards for Forensic DNA Testing Laboratories. The laboratories that perform DNA testing should always work according to the established standards while considering all the ways to maintain the quality of the results produced.

30.3.1 Personnel

The laboratory must have qualified DNA analysts who have experience in examination and testimony. It should have personnel with technical skills and other managerial staff with all required qualifications to complete the assigned work. At the time of hiring laboratory personnel, they must pass the competency test (to confirm his/her knowledge about the field). They must then be trained irrespective of their experience(s), with the updated testing procedures of the laboratory. Based on the expertise and the accuracy required, individuals may be assigned separate tasks.

30.3.2 Laboratory Manuals

The laboratory must have the manuals or SOP (standard operating procedures booklets) soliciting all the analytical procedures performed during testing for maintaining the consistency of testing procedures, and the result produced through it. The laboratory should always follow well established, tested and validated methodologies for DNA testing. The manual should be regularly maintained and updated from time to time, considering the accurate testing and innovations in the field. Along with manual methods, automated or robotic methods should also be validated. Care should always be taken to evaluate the updated procedure for maintaining consistency in the result. Every single modification or update to the software will require a routine check to maintain the integrity of results. Validation of the procedures adopted should be performed by the laboratory (Butler 2001). There should be safety manuals for the laboratory personnel. When dealing with 'unique' samples, the method may be modified as per the existing literature and/or experience of the senior analyst. However, the process should be documented and must be with appropriate justification.

30.3.3 Requirements

Fully equipped lab is the prime requirement for each testing laboratory, including reagents, apparatus, equipment, personal safety wears, etc. The laboratory should always use calibrated equipment. All equipment should be sterilised after every use. The use of reagents should always be according to the written protocol given in the laboratory manual/SOPs. The reagents that are prepared in the laboratory should also

be labelled, including the composition, date of preparation, expiration, the identity of individual preparing reagents, and other relevant details. Depending upon the class of reagent, they should be kept in favourable/suitable storage conditions (dark bottle, away from sunlight, etc.). Logs of the reagents and chemicals used by the laboratory must be maintained as a routine practice. No one should be allowed to enter the laboratory without prior permission and appropriate authorisation. All the entries in and out of the laboratory should be documented and controlled.

30.3.4 Evidence Maintenance

Each laboratory should have a secure and governed area for evidence storage and testing materials (FBI n.d.). The samples, when received at the laboratory, should be appropriately marked using the common standards. The laboratory should also maintain the chain of custody of every piece of evidence, so that when the report is presented in the court, it should stand admissible. The evidence at the laboratory should be accessed by the dedicated personnel only.

Every time only a single sample should be analysed to avoid sample switching. It should be a common practice to try and avoid analysing the reference and suspected sample parallelly.

The result obtained should be uniform throughout the testing procedure. Regular inspection and professional assessments are also the critical factors in maintaining the testing laboratories. Professional assessments will involve internal and external testing. Internal testing will comprise of testing within the laboratory, i.e., the same sample will be tested by multiple personnel to check the consistency of the results. Whereas in external testing, the samples will be tested by authorised external bodies to check the consistency and/or variability of the results (Balažic and Zupanič 1999).

The laboratory should always be cleaned before and after the testing procedures. Care should be taken for not exposing the testing area to direct environment. Primary as well as secondary transfers must be prevented. The appropriate environmental and health safety guidelines and norms must be followed while disposing of the chemicals and waste.

30.3.5 Accreditation

Accreditation boards and agencies function to provide accreditation to the laboratories by assessing their technical competency. The laboratory accreditation services to testing and calibration laboratories are provided following ISO/IEC 17025: 2005 and ISO/IEC 17025: 2017 or other established standards that describes the general and/or specific requirements for the competence of testing and calibration laboratories. The guidelines, standards, and parameters followed vary from country to country. The results of accredited laboratories have the upper hand over their non-accredited counterparts. Regular auditing of the laboratories is the essential

Table 30.1 List of organisations providing accreditation to the forensic DNA testing laboratories

Sr. No.	Name of the body	Abbreviation	Link	Country
1	ANSI National Accreditation Board	ANAB	https://anab.ansi.org/	United States of America
2	American Society of Crime Laboratory Directors Laboratory Accreditation Board	ASCLDLAB	https://www.ascld.org/	United States of America
3	National Accreditation of Testing Authority	NATA	https://www.nata.com.au/	Australia
4	National Accreditation Board for Testing and Calibration Laboratories	NABL	https://nabl-india.org	India
5	The American Association for Laboratory Accreditation	A2LA	https://www.a2la.org/	United States of America
6	United Kingdom Accreditation Service	UKAS	https://www.ukas.com/	United Kingdom
7	New York State Department of Health	NYSDOH	https://www.health.ny.gov/	United States of America
8	The South African National Accreditation System	SANAS	https://www.sanas.co.za/	South Africa
9	Polish Centre for Accreditation	PCA	https://www.pca.gov.pl/	Poland
10	National Accreditation Authority Hungary	NAH	https://www.nah.gov.hu/	Hungary
11	Czech Accreditation Institute	CAI	https://www.cai.cz/	Czech Republic
12	Standards Council of Canada	SCC	https://www.scc.ca/	Canada
13	Finland Accreditation Service	FINAS	https://www.finas.fi/	Finland
14	Deutsche Akkreditierungsstelle	DAkkS	https://www.dakks.de/	Germany
15	The Spanish Association for Standardization and Certification	AENOR	https://www.en.aenor.com/	Spain
16	Swedish Board for Accreditation and Conformity Assessment	SWEDAC	https://www.swedac.se/	Sweden

(continued)

Table 30.1 (continued)

Sr. No.	Name of the body	Abbreviation	Link	Country
17	Swiss Accreditation Service	SAS	https://www.sas.admin.ch/	Switzerland
18	Japan National Laboratory Accreditation	JNLA	https://www.nite.go.jp/	Japan
19	China National Accreditation Service for Conformity Assessment	CNAS	https://www.cnas.org.cn/	China
20	Korea Laboratory Accreditation Scheme	KOLAS		South Korea
21	The Association of Analytical Centers 'Analytics'	AAC 'Analytics'	http://aac-analitica.ru/	Russia

procedure of quality control. Table 30.1 shows the list of organisations providing accreditation to the Forensic DNA Testing Laboratories.

30.3.6 Proficiency Testing

This should be carried out regularly by participating in various proficiency testing schemes. It indicates the ability of the laboratory to produce excellent and valid results independently (Juniper 1999). Some accreditation body requires proficiency testing as their prerequisite. It can be performed in two ways by open testing and blind testing. The open testing states the purpose of testing, whereas blind testing will be done without prior information of sample analyses. Also, this can be done by sending the samples for testing to all participants and then comparing the results produced by them. This is also known as external quality assessment. These schemes focus on increasing the quality control standards of the laboratory (Dequeker et al. 2001). Table 30.2 shows the list of global institutions/organisations conducting proficiency tests for the DNA examiners.

30.3.7 Reporting of Non-conformance and Near-Failures

According to ISO/IEC 17025, the reporting of non-conformance and near-failures plays a vital role in the quality control of the laboratory. The non-conformance can be an instantaneous indication of errors or failure of the test or procedure, or it is possible that it may occur after a while.

Reporting them helps in solving the issues at hand. This can be looked after by identifying the error-causing factor and can be corrected. The failures should also be maintained, which may be the result of human error or technical errors. The laboratory should maintain the Elimination Databases (EDB),which help in reporting

Table 30.2 List of global institutions/organisations conducting proficiency tests for the DNA examiners

Sr. No.	Proficiency testing agency	Website	Proficiency test scheme	Country
1	Forensic Foundations	https://www.forensicfoundations.com.au/proficiency-testing/2020-proficiency-tests/	Forensic Biological Examination and DNA-1. 2020-1 Forensic Biological Examination and DNA-22020-10 Forensic Biological Examination, BPA and DNA. 2020-7	Australia
2	Bode Technology	www.bodetechnology.com	IQAS Proficiency Test	United States of America
3	Forensic Assurance	https://forensicassurance.com/	Forensic Biology	United States of America
4	Collaborative Testing Services	http://www.ctsforensics.com/	CTS Forensics Program: DNA Interpretation CTS Forensics Program: DNA Parentage CTS Forensics Program: DNA Specific Series CTS Forensics Program—DNA Database Saliva CTS Forensics Program: Biology	United States of America
5	InstitutoNacional de ToxicologíayCienciasForenses (INTCF)	https://www.mjusticia.gob.es/	Analysis of DNA polymorphisms in bloodstains and other biological samples. Basic level analysis of DNA polymorphisms in bloodstains and other biological samples. Advanced level	Spain
6	RCPAQAP	https://rcpaqap.com.au/	Forensic diagnostic	Australia
7	Reference Institute for Bioanalytics	https://www.rfb.bio/	DNA isolation [DNA isolation]	Germany
8	Dirección General de ServiciosPericialesFiscalía General del Estado de Chiapas	https://www.chiapas.gob.mx/	–	Mexico
9	Contra Costa County Office of the Sheriff Forensic Services—Division	http://cocosheriff.org/	–	United States of America

(continued)

Table 30.2 (continued)

Sr. No.	Proficiency testing agency	Website	Proficiency test scheme	Country
10	Federal Bureau of Investigation	https://www.fbi.gov/	–	United States of America
11	The Scientific Working Group on DNA Analysis Methods (SWGDAM)	https://www.swgdam.org/	–	United States of America
12	Serological Research Institute	http://www.serological.com/	–	United States of America
13	Central Police Forensic Laboratory	http://clkp.policja.pl/	–	Poland
14	The GEDNAP (German DNA profiling group) proficiency testing system	https://www.gednap.org/	–	Germany
15	Hungarian Institute for Forensic Sciences	http://www.nszkk.gov.hu/	–	Hungary
16	Forenzni DNA Servis	http://www.dna.com.cz/	–	Czech
17	Central Forensic Science Laboratory, Hyderabad	http://cfslhyd.gov.in/	–	India
18	ESR	https://www.esr.cri.nz/	–	New Zealand
19	Bureau Veritas	https://www.bvlabs.com/	–	Canada
20	Forensic Proficiency Testing Service	https://www.forensic-proficiency.co.uk/	–	United Kingdom
21	Easy DNA	https://www.easydna.co.uk/	–	United Kingdom
22	University of Granada—GITAD	www.ugr.es	DNA polymorphisms in bloodstains and other biological samples (autosomal, X- and Y-chromosome, mtDNA); statistical calculations; mixed stains interpretation	Spain

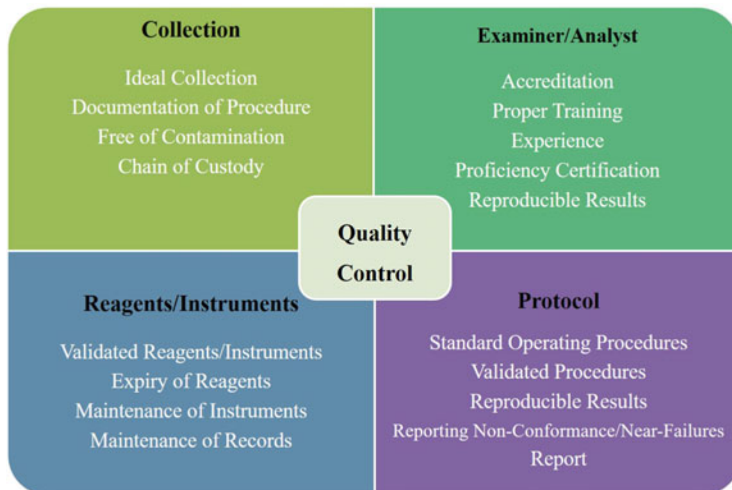


Fig. 30.3 Quality control highlights

eliminations as actual evidence. It mainly contains the DNA profile, which has shown failure and not reported by the laboratory. If the laboratory fails to maintain the EDB, the wrong profile can be used in the legal proceedings and will adversely affect the case in hand. These should be evaluated yearly to compare the growth of the laboratory (Ansell 2013). Figure 30.3 depicts the important points that must be remembered during the quality control cycle.

30.4 Quality Control in Reporting

The reports generated by the testing laboratories are admissible in the court unless the questions are raised on its authenticity. The laboratory should maintain everything in writing to support the conclusions drawn through it. The data should be maintained in hardcopy as well as in softcopy. It should be maintained in a defined format so that everyone could get it when going through it. It should contain all the details of the case and the process. The report will contain information including case identifier/serial number, detailed description of the evidence analysed, description of the procedures and techniques performed, result or conclusion, interpretation of the statements made, date issued, disposition of the evidence and lastly the signature of the expert/authorised person who performed the primary operations in the case. The laboratory should follow the rules for maintaining the privacy of the evidence and information related to it. The final report should be verified by more than one person before dispatch. The reports should never be biased. The laboratory should always hand over the reports to the concerned authorised person. There should not be any leakage of the information from the laboratory. The chain of custody should be maintained throughout the procedure (FBI n.d.).

30.5 Conclusion

In forensic DNA typing setups, '*Quality control*' is one of the critical components in ensuring high quality and standards. The path of the evidence from the crime scene to the courtroom is quite lengthy and intricate. The characteristic unpredictability of every crime and the biological evidence thus recovered and analysed adds to the complexity. Different QCs along the whole process shall assist in providing justice by eliminating the chances of errors and thus being admissible in the court of law. Maintenance of the chain of custody as well as ensuring quality control measures at each and every step during the whole investigation shall be given utmost priority and importance.

Supplementary Materials

See Tables 30.3 and 30.4.

Table 30.3 Compilation of some of the validation studies conducted using commercial STR kits, in-house assays, instrumentation, and software in human identity testing applications

S. No.	Kit, Assay, or Instrument	Reference
1.	19-locus Y-STR system	Daniels et al. (2004)
2.	21-SNP multiplex	Dixon et al. (2005)
3.	ABI 310	Lazaruk et al. (1998), Isenberg et al. (1998), Moretti et al. (2001a, b)
4.	ABI 3100	Sgueglia et al. (2003), Koumi et al. (2004)
5.	ABI 3700	Gill et al. (2001), Koumi et al. (2004)
6.	ABI 377	Frazier et al. (1996), Fregeau et al. (1999)
7.	Amelogenin	LaFountain et al. (1998)
8.	AmpFlSTR Blue	Wallin et al. (1998)
9.	AmpFlSTR Green I	Holt et al. (2002)
10.	Biomek 2000 with DNA IQ	Greenspoon et al. (2004a)
11.	BodeQuant	Fox et al. (2003)
12.	COfiler	LaFountain et al. (2001), Tomsey et al. (2001), Moretti et al. (2001a, b), Holt et al. (2002), Buse et al. (2003), Wallin et al. (2002)
13.	CompareCalls software	Ryan et al. (2004)
14.	CTT	Budowle et al. (1997)
15.	D12S391	Junge and Madea (1999)
16.	D3S1358, D8S1179, D18S51	Potter (2003)

(continued)

Table 30.3 (continued)

S. No.	Kit, Assay, or Instrument	Reference
17.	DNA quant (AluQuant)	Mandrekar et al. (2001)
18.	DNA quant (RT-PCR CSF)	Richard et al. (2003)
19.	DNA quant (RT-PCR *Alu*)	Nicklas and Buel (2003)
20.	genRES MPX-2	Junge et al. (2003)
21.	Identifiler	Collins et al. (2004)
22.	MegaBACE	Koumi et al. (2004)
23.	mtDNAminisequencing	Morley et al. (1999)
24.	mtDNA sequencing	Wilson et al. (1995), Holland and Parsons (1999)
25.	PowerPlex 1.1	Micka et al. (1999), Tomsey et al. (2001), Greenspoon et al. (2000)
26.	PowerPlex 1.1 + D16 primer	Nelson et al. (2002)
27.	PowerPlex 16	Krenke et al. (2002), Tomsey et al. (2001)
28.	PowerPlex 16 BIO	Greenspoon et al. (2004b)
29.	PowerPlex 2.1	Tomsey et al. (2001), Levedakou et al. (2002)
30.	PowerPlex Y	Krenke et al. (2005)
31.	Profiler	Holt et al. (2002)
32.	Profiler Plus	Frank et al. (2001), LaFountain et al. (2001), Tomsey et al. (2001), Holt et al. (2002), Fregeau et al. (2003), Buse et al. (2003), Wallin et al. (2002), Pawlowski and Maciejewska (2000), Moretti et al. (2001a, b)
33	Profiler Plus *ID*	Leibelt et al. (2003)
34.	Quantifiler	Applied Biosystems (2003)
35.	Reduced volume PCR for Profiler Plus STR kit	Gaines et al. (2002), Fregeau et al. (2003)
36.	SEfiler	Coticone et al. (2004)
37.	SGM	Sparkes et al. (1996a, b), Kimpton et al. (1996)
38.	SGM Plus	Cotton et al. (2000)
39.	STR sets	Crouse and Schumm (1995), Micka et al. (1996)
40.	TH01	Van Oorschot et al. (1996), Wiegand et al. (1993)
41.	TH01, VWA, F13A1, FES	Lygo et al. (1994), Clayton et al. (1995), Andersen et al. (1996)
42.	TH01, VWA, F13A1, FES, LPL	Pestoni et al. (1995)
43.	TrueAllele software	Kadash et al. (2004)
44.	Y-PLEX 12	Shewale et al. (2004)
45.	Y-PLEX 5	Sinha et al. (2003b)
46.	Y-PLEX 6	Sinha et al. (2003a)
47.	Y-STR 10plex	Johnson et al. (2003)
48.	Y-STR 4plex	Prinz et al. (2001)

Available from: https://strbase.nist.gov/

Table 30.4 Summary of articles published on rapid DNA instruments since 2016

S. No.	Instrument	STR Primer Set	Reference
1.	ANDE 6C	FlexPlex (6-dye, 27plex STR assay)	Carney et al. (2019)
2.	DNAscan/ ANDE 4C	PowerPlex 16	Turingan et al. (2016), Della Manna et al. (2016), Moreno et al. (2017)
3.	RapidHIT 200	GlobalFiler Express	Date-Chong et al. (2016)
4.	RapidHIT 200	NGM SElect Express	Boiso et al. (2017a, b), Shackleton et al. (2019a, b)
5.	RapidHIT ID	GlobalFiler Express	Wiley et al. (2017), Salceda et al. (2017), Amick and Swiger (2019)
6.	RapidHIT ID	GlobalFiler Express and NGM SElect Express	Buscaino et al. (2018)

Adapted from: https://doi.org/10.1016/j.fsisyn.2019.12.002

References

Amick GD, Swiger RR (2019) Internal validation of RapidHIT® ID ACE sample cartridge and assessment of the EXT sample cartridge. J Forensic Sci 64(3):857–868

Andersen JF, Greenhalgh MJ, Butler HR, Kilpatrick SR, Piercy RC, Way KA, Myhill HS, Wright JC, Hallett R, Parkin BH (1996) Further validation of a multiplex STR system for use in routine forensic identity testing. Forensic Sci Int 78:47–64

Ansell R (2013) Internal quality control in forensic DNA analysis. Accred Qual Assur 18:279–289. https://doi.org/10.1007/s00769-013-0968-9

Applied Biosystems (2003) Quantifiler human DNA quantification kit user's manual. Applied Biosystems, Foster City, CA

ASQ (n.d.). https://asq.org/quality-resources/iso-9000

Badiye A, Kapoor N, Menezes RG (2020) Chain of custody (chain of evidence). In: StatPearls. StatPearls Publishing LLC, Treasure Island, FL. http://www.ncbi.nlm.nih.gov/pubmed/31869141

Balažic J, Zupanič I (1999) Quality control and quality assurance in DNA laboratories: legal, civil and ethical aspects. Forensic Sci Int 103:S1–S5. https://doi.org/10.1016/S0379-0738(99)00064-X

Boiso S, Dalin E, Seidlitz H, Trygg E, Hedman J, Ansell R (2017a) Experiences from operating the RapidHIT® System and identified issues processing crime scene samples. February 2017 report from the Swedish National Forensic Centre. Available from: https://nfc.polisen.se/siteassets/dokument/informationsmaterial/rapporter/nfc-rapport-2017-02_experiences-fromoperating-the-rapidhit-system.pdf

Boiso S, Dalin E, Seidlitz H, Sidstedt M, Trygg E, Hedman J, Ansell R (2017b) RapidHIT for the purpose of stain analyses e an interrupted implementation. Forensic Sci Int Genet Suppl Ser 6: e589–e590

Budowle B, Moretti TR, Keys KM, Koons BW, Smerick JB (1997) Validation studies of the CTT STR multiplex system. J Forensic Sci 42(4):701–707

Buscaino J, Barican A, Farrales L, Goldman B, Klevenberg J, Kuhn M et al (2018) Evaluation of a rapid DNA process with the RapidHIT® ID system using a specialized cartridge for extracted and quantified human DNA. Forensic Sci Int Genet 34:116–127

Buse EL, Putinier JC, Hong MM, Yap AE, Hartmann JM (2003) Performance evaluation of two multiplexes used in fluorescent short tandem repeat DNA analysis. J Forensic Sci 48 (2):348–357

Butler JM (2005) Forensic DNA typing: biology, technology, and genetics of STR markers, 2nd edn. Academic Press, London. https://www.elsevier.com/books/forensic-dna-typing/butler/978-0-08-047061-0

Butler JM (2001) Forensic DNA Typing: Biology and Technology behind STR markers. Elsevier Academic Press, San Diego

Carney C, Whitney S, Vaidyanathan J, Persick R, Noel F, Vallone PM et al (2019) Developmental validation of the ANDE rapid DNA system with FlexPlex assay for arrestee and reference buccal swab processing and database searching. Forensic Sci Int Genet 40:120–130

Clayton TM, Whitaker JP, Fisher DL, Lee DA, Holland MM, Weedn VW, Maguire CN, DiZinno JA, Kimpton CP, Gill P (1995) Further validation of a quadruplex STR DNA typing system: a collaborative effort to identify victims of a mass disaster. Forensic Sci Int 76:17–25

Collins PJ, Hennessy LK, Leibelt CS, Roby RK, Reeder DJ, Foxall PA (2004) Developmental validation of a single-tube amplification of the 13 CODIS STR loci, D2S1338, D19S433, and amelogenin: the AmpFlSTRIdentifiler PCR amplification kit. J Forensic Sci 49(6):1265–1277

Coticone SR, Oldroyd N, Philips H, Foxall P (2004) Development of the AmpFISTRSEfiler PCR amplification kit: a new multiplex containing the highly discriminating ACTBP2 (SE33) locus. Int J Legal Med 118(4):224–234

Cotton EA, Allsop RF, Guest JL, Frazier RR, Koumi P, Callow IP, Seager A, Sparkes RL (2000) Validation of the AMPFlSTR((R)) SGM plus system for use in forensic casework. Forensic Sci Int 112(2-3):151–161

Crouse C, Schumm JW (1995) Investigation of species specificity using nine PCR-based human STR systems. J Forensic Sci 40(6):952–956

Daniels DL, Hall AM, Ballantyne J (2004) SWGDAM developmental validation of a 19-locus Y-STR system for forensic casework. J Forensic Sci 49(4):668–683

Date-Chong M, Hudlow WR, Buoncristiani MR (2016) Evaluation of the RapidHIT 200 and RapidHITGlobalFiler® express kit for fully automated STR genotyping. Forensic Sci Int Genet 23:1–8

Della Manna A, Nye JV, Carney C, Hammons JS, Mann M, Al Shamali F et al (2016) Developmental validation of the DNAscan Rapid DNA Analysis instrument and expert system for reference sample processing. Forensic Sci Int Genet 25:145–156

Dequeker E, Ramsden S, Grody W et al (2001) Quality control in molecular genetic testing. Nat Rev Genet 2:717–723. https://doi.org/10.1038/35088588

Dixon LA, Murray CM, Archer EJ, Dobbins AE, Koumi P, Gill P (2005) Validation of a 21-locus autosomal SNP multiplex for forensic identification purposes. Forensic Sci Int 154:62–77

Eugene T (2011) Sample collection system for DNA analysis of forensic evidence: towards practical, fully-integrated str analysis. https://www.ncjrs.gov/pdffiles1/nij/grants/236826.pdf

FBI (n.d.). https://www.fbi.gov/file-repository/quality-assurance-standards-for-forensic-dna-testing-laboratories.pdf/view. Accessed 30 Mar 2020

Fox JC, Cave CA, Schumm JW (2003) Development, characterization, and validation of a sensitive primate-specific quantification assay for forensic analysis. Biotechniques 34(2):314–318. 320, 322

Frank WE, Llewellyn BE, Fish PA, Riech AK, Marcacci TL, Gandor DW, Parker D, Carter RR, Thibault SM (2001) Validation of the AmpFlSTR Profiler Plus PCR amplification kit for use in forensic casework. J Forensic Sci 46(3):642–646

Frazier RR, Millican ES, Watson SK, Oldroyd NJ, Sparkes RL, Taylor KM, Panchal S, Bark L, Kimpton CP, Gill PD (1996) Validation of the Applied BiosystemsPrism™ 377 automated sequencer for forensic short tandem repeat analysis. Electrophoresis 17(10):1550–1552

Fregeau CJ, Bowen KL, Fourney RM (1999) Validation of highly polymorphic fluorescent multiplex short tandem repeat systems using two generations of DNA sequencers. J Forensic Sci 44(1):133–166

Fregeau CJ, Bowen KL, Leclair B, Trudel I, Bishop L, Fourney RM (2003) AmpFlSTR Profiler Plus short tandem repeat DNA analysis of casework samples, mixture samples, and nonhuman DNA samples amplified under reduced PCR volume conditions (25 microL). J Forensic Sci 48 (5):1014–1034

Gaines ML, Wojtkiewicz PW, Valentine JA, Brown CL (2002) Reduced volume PCR amplification reactions using the AmpFlSTR Profiler Plus kit. J Forensic Sci 47(6):1224–1237

Gill P, Koumi P, Allen H (2001) Sizing short tandem repeat alleles in capillary array gel electrophoresis instruments. Electrophoresis 22:2670–2678

Greenspoon SA, Lytle PJ, Turek SA, Rolands JM, Scarpetta MA, Carr CD (2000) Validation of the PowerPlex 1.1 loci for use in human identification. J Forensic Sci 45(3):677–683

Greenspoon SA, Ban JD, Sykes K, Ballard EJ, Edler SS, Baisden M, Covington BL (2004a) Application of the BioMek 2000 Laboratory Automation Workstation and the DNA IQ System to the extraction of forensic casework samples. J Forensic Sci 49(1):29–39

Greenspoon SA, Ban JD, Pablo L, Crouse CA, Kist FG, Tomsey CS, Glessner AL, Mihalacki LR, Long TM, Heidebrecht BJ, Braunstein CA, Freeman DA, Soberalski C, Nathan B, Amin AS, Douglas EK, Schumm JW (2004b) Validation and implementation of the PowerPlex 16 BIO System STR multiplex for forensic casework. J Forensic Sci 49(1):71–80

Holland MM, Parsons TJ (1999) Mitochondrial DNA sequence analysis - validation and use for forensic casework. Forensic Sci Rev 11(1):22–50

Holt CL, Buoncristiani M, Wallin JM, Nguyen T, Lazaruk KD, Walsh PS (2002) TWGDAM validation of AmpFlSTR PCR amplification kits for forensic DNA casework. J Forensic Sci 47 (1):66–96

Isenberg AR, Allen RO, Keys KM, Smerick JB, Budowle B, McCord BR (1998) Analysis of two multiplexed short tandem repeat systems using capillary electrophoresis with multiwavelength florescence detection. Electrophoresis 19:94–100

Johnson CL, Warren JH, Giles RC, Staub RW (2003) Validation and uses of a Y-chromosome STR 10-plex for forensic and paternity laboratories. J Forensic Sci 48(6):1260–1268

Junge A, Madea B (1999) Validation studies and characterization of variant alleles at the short tandem repeat locus D12S391. Int J Legal Med 112(1):67–69

Junge A, Lederer T, Braunschweiger G, Madea B (2003) Validation of the multiplex kit genRESMPX-2 for forensic casework analysis. Int J Legal Med 117(6):317–325

Juniper I (1999) Quality issues in proficiency testing. Accred Qual Assur 4:336–341. https://doi.org/10.1007/s007690050377

Kadash K, Kozlowski BE, Biega LA, Duceman BW (2004) Validation study of the TrueAllele automated data review system. J Forensic Sci 49(4):660–667

Kimpton CP, Oldroyd NJ, Watson SK, Frazier RR, Johnson PE, Millican ES, Urquhart A, Sparkes BL, Gill P (1996) Validation of highly discriminating multiplex short tandem repeat amplification systems for individual identification. Electrophoresis 17(8):1283–1293

Koumi P, Green HE, Hartley S, Jordan D, Lahec S, Livett RJ, Tsang KW, Ward DM (2004) Evaluation and validation of the ABI 3700, ABI 3100, and the MegaBACE 1000 capillary array electrophoresis instruments for use with short tandem repeat microsatellite typing in a forensic environment. Electrophoresis 25(14):2227–2241

Krenke BE, Tereba A, Anderson SJ, Buel E, Culhane S, Finis CJ, Tomsey CS, Zachetti JM, Masibay A, Rabbach DR, Amiott EA, Sprecher CJ (2002) Validation of a 16-locus fluorescent multiplex system. J Forensic Sci 47(4):773–785

Krenke BE, Viculis L, Richard ML, Prinz M, Milne SC, Ladd C, Gross AM, Gornall T, Frappier JRH, Eisenberg AJ, Barna C, Aranda XG, Adamowicz BB (2005) Validation of a male-specific, 12-locus fluorescent short tandem repeat (STR) multiplex. Forensic Sci Int 148:1–14

LaFountain M, Schwartz M, Cormier J, Buel E (1998) Validation of capillary electrophoresis for analysis of the X-Y homologous amelogenin gene. J Forensic Sci 43(6):1188–1194

LaFountain MJ, Schwartz MB, Svete PA, Walkinshaw MA, Buel E (2001) TWGDAM validation of the AmpFlSTR Profiler Plus and AmpFlSTRCOfiler STR multiplex systems using capillary electrophoresis. J Forensic Sci 46(5):1191–1198

Lazaruk K, Walsh PS, Oaks F, Gilbert D, Rosenblum BB, Menchen S, Scheibler D, Wenz HM, Holt C, Wallin J (1998) Genotyping of forensic short tandem repeat (STR) systems based on sizing precision in a capillary electrophoresis instrument. Electrophoresis 19(1):86–93

Leibelt C, Budowle B, Collins P, Daoudi Y, Moretti T, Nunn G, Reeder D, Roby R (2003) Identification of a D8S1179 primer binding site mutation and the validation of a primer designed to recover null alleles. Forensic Sci Int 133(3):220–227

Levedakou EN, Freeman DA, Budzynski MJ, Early BE, Damaso RC, Pollard AM, Townley AJ, Gombos JL, Lewis JL, Kist FG, Hockensmith ME, Terwilliger ML, Amiott E, McElfresh KC, Schumm JW, Ulery SR, Konotop F, Sessa TL, Sailus JS, Crouse CA, Tomsey CS, Ban JD, Nelson MS (2002) Characterization and validation studies of powerPlex 2.1, a nine-locus short tandem repeat (STR) multiplex system and penta D monoplex. J Forensic Sci 47(4):757–772

Li R (2015) Forensic biology, 2nd edn. CRC Press, Taylor & Francis Group, Boca Raton

Lygo JE, Johnson PE, Holdaway DJ, Woodroffe S, Whitaker JP, Clayton TM, Kimpton CP, Gill P (1994) The validation of short tandem repeat (STR) loci for use in forensic casework. Int J Legal Med 107(2):77–89

Lynch M (2003) God's signature: DNA profiling, the new gold standard in forensic science. Endeavour 27:93–97. https://doi.org/10.1016/S0160-9327(03)00068-1

Mandrekar MN, Erickson AM, Kopp K, Krenke BE, Mandrekar PV, Nelson R, Peterson K, Shultz J, Tereba A, Westphal N (2001) Development of a human DNA quantitation system. Croat Med J 42(3):336–339

Micka KA, Sprecher CJ, Lins AM, Comey CT, Koons BW, Crouse C, Endean D, Pirelli K, Lee SB, Duda N, Ma M, Schumm JW (1996) Validation of multiplex polymorphic STR amplification sets developed for personal identification applications. J Forensic Sci 41(4):582–590

Micka KA, Amiott EA, Hockenberry TL, Sprecher CJ, Lins AM, Rabbach DR, Taylor JA, Bacher JW, Glidewell DE, Gibson SD, Crouse CA, Schumm JW (1999) TWGDAM validation of a nine-locus and a four-locus fluorescent STR multiplex system. J Forensic Sci 44(6):1243–1257

Moreno LI, Brown AL, Callaghan TF (2017) Internal validation of the DNAscan/ANDE Rapid DNA Analysis platform and its associated PowerPlex® 16 high content DNA biochip cassette for use as an expert system with reference buccal swabs. Forensic Sci Int Genet 29:100–108

Moretti TR, Baumstark AL, Defenbaugh DA, Keys KM, Brown AL, Budowle B (2001a) Validation of STR typing by capillary electrophoresis. J Forensic Sci 46(3):661–676

Moretti TR, Baumstark AL, Defenbaugh DA, Keys KM, Smerick JB, Budowle B (2001b) Validation of short tandem repeats (STRs) for forensic usage: performance testing of fluorescent multiplex STR systems and analysis of authentic and simulated forensic samples. J Forensic Sci 46(3):647–660

Morley JM, Bark JE, Evans CE, Perry JG, Hewitt CA, Tully G (1999) Validation of mitochondrial DNA minisequencing for forensic casework. Int J Legal Med 112(4):241–248

National Research Council (US) Committee on DNA Technology in Forensic Science (1992) DNA technology in forensic science. National Academies Press (US), Washington, DC. 6. Use of DNA information in the legal system. Available from: https://www.ncbi.nlm.nih.gov/books/NBK234535/

Nelson MS, Levedakou EN, Matthews JR, Early BE, Freeman DA, Kuhn CA, Sprecher CJ, Amin AS, McElfresh KC, Schumm JW (2002) Detection of a primer-binding site polymorphism for the STR locus D16S539 using the Powerplex 1.1 system and validation of a degenerate primer to correct for the polymorphism. J Forensic Sci 47(2):345–349

Nicklas JA, Buel E (2003) Development of an Alu-based, real-time PCR method for quantitation of human DNA in forensic samples. J Forensic Sci 48(5):936–944

van Oorschot RA, Gutowski SJ, Robinson SL, Hedley JA, Andrew IR (1996) HUMTH01 validation studies: effect of substrate, environment, and mixtures. J Forensic Sci 41(1):142–145

Pawlowski R, Maciejewska A (2000) Forensic validation of a multiplex containing nine STRs--population genetics in northern Poland. Int J Legal Med 114(1–2):45–49

Pestoni C, Lareu MV, Rodriguez MS, Munoz I, Barros F, Carracedo A (1995) The use of the STRs HUMTH01, HUMVWA31/A, HUMF13A1, HUMFES/FPS, HUMLPL in forensic application: validation studies and population data for Galicia (NW Spain). Int J Legal Med 107(6):283–290

Potter T (2003) Co-amplification of ENFSI-loci D3S1358, D8S1179 and D18S51: validation of new primer sequences and allelic distribution among 2874 individuals. Forensic Sci Int 138 (1–3):104–110

Prinz M, Ishii A, Coleman A, Baum HJ, Shaler RC (2001) Validation and casework application of a Y chromosome specific STR multiplex. Forensic Sci Int 120(3):177–188

Richard ML, Frappier RH, Newman JC (2003) Developmental validation of a real-time quantitative PCR assay for automated quantification of human DNA. J Forensic Sci 48(5):1041–1046

Ryan JH, Barrus JK, Budowle B, Shannon CM, Thompson VW, Ward BE (2004) The application of an automated allele concordance analysis system (CompareCalls) to ensure the accuracy of single-source STR DNA profiles. J Forensic Sci 49(3):492–499

Salceda S, Barican A, Buscaino J, Goldman B, Klevenberg J, Kuhn M et al (2017) Validation of a rapid DNA process with the RapidHIT® ID system using GlobalFiler® Express chemistry, a platform optimized for decentralized testing environments. Forensic Sci Int Genet 28:21–34

Sgueglia JB, Geiger S, Davis J (2003) Precision studies using the ABI prism 3100 genetic analyzer for forensic DNA analysis. Anal Bioanal Chem 376(8):1247–1254

Shackleton D, Gray N, Ives L, Malsom S, Vanhinsbergh D (2019a) Development of RapidHIT® ID using NGM SElect™ Express chemistry for the processing of reference samples within the UK criminal justice system. Forensic Sci Int 295:179–188

Shackleton D, Pagram J, Andrews N, Malsom S, Ives L, Vanhinsbergh D (2019b) Development of enhanced sensitivity protocols on the RapidHIT™ 200 with a view to processing casework material. Sci Justice 59:411–417

Shewale JG, Nasir H, Schneida E, Gross AM, Budowle B, Sinha SK (2004) Y-chromosome STR system, Y-PLEX 12, for forensic casework: development and validation. J Forensic Sci 49 (6):1278–1290

Sinha SK, Budowle B, Arcot SS, Richey SL, Chakrabor R, Jones MD, Wojtkiewicz PW, Schoenbauer DA, Gross AM, Sinha SK, Shewale JG (2003a) Development and validation of a multiplexed Y-chromosome STR genotyping system, Y-PLEX 6, for forensic casework. J Forensic Sci 48(1):93–103

Sinha SK, Nasir H, Gross AM, Budowle B, Shewale JG (2003b) Development and validation of the Y-PLEX 5, a Y-chromosome STR genotyping system, for forensic casework. J Forensic Sci 48 (5):985–1000

Sparkes R, Kimpton C, Gilbard S, Carne P, Andersen J, Oldroyd N, Thomas D, Urquhart A, Gill P (1996a) The validation of a 7-locus multiplex STR test for use in forensic casework. (II), Artefacts, casework studies and success rates. Int J Legal Med 109(4):195–204

Sparkes R, Kimpton CP, Watson S, Oldroyd NJ, Clayton TM, Barnett L, Arnold J, Thompson C, Hale R, Chapman J, Urquhart A, Gill P (1996b) The validation of a 7-locus multiplex STR test for use in forensic casework: I. Mixtures, ageing, degradation and species studies. Int J Legal Med 109(4):186–194

Tomsey CS, Kurtz M, Kist F, Hockensmith M, Call P (2001) Comparison of PowerPlex 16, PowerPlex1.1/2.1, and ABI AmpFISTR Profiler Plus/COfiler for forensic use. Croat Med J 42(3):239–243

Turingan RS, Vasantgadkar S, Palombo L, Hogan C, Jiang H, Tan E, Selden RF (2016) Rapid DNA analysis for automated processing and interpretation of low DNA content samples. Invest Genet 7:2

Wallin JM, Buoncristiani MR, Lazaruk KD, Fildes N, Holt CL, Walsh PS (1998) TWGDAM validation of the AmpFISTR Blue PCR amplification kit for forensic casework analysis. J Forensic Sci 43(4):854–870

Wallin JM, Holt CL, Lazaruk KD, Nguyen TH, Walsh PS (2002) Constructing universal multiplex PCR systems for comparative genotyping. J Forensic Sci 47(1):52–65

Wiegand P, Budowle B, Rand S, Brinkmann B (1993) Forensic validation of the STR systems SE 33 and TC 11. Int J Legal Med 105(6):315–320
Wiley R, Sage K, LaRue B, Budowle B (2017) Internal validation of the RapidHIT® ID system. Forensic Sci Int Genet 31:180–188
Wilson MR, DiZinno JA, Polanskey D, Replogle J, Budowle B (1995) Validation of mitochondrial DNA sequencing for forensic casework analysis. Int J Legal Med 108(2):68–74

Legal Aspects of Forensic DNA Typing 31

Ankit Srivastava, Abhimanyu Harshey, and Pankaj Shrivastava

Abstract

DNA testing is a powerful tool in criminal investigation. The advent of DNA profiling has revolutionized the criminal justice system. Since the discovery of DNA till now, it has provided the cogent evidential base to corroborate guilt or innocence in the court of law. The anticipation of technological advancement in the criminal justice system puts two significant concerns that are to protect the fundamental rights of the citizen, and other is to eliminate the probability of misuse of technique. Many countries have enacted DNA legislation to provide the legal background to DNA testing along with the establishment of DNA databases, while India is in the process of adopting the DNA Technology (Use and Application) bill 2019. Legal perspectives of DNA testing in different nations are discussed in this chapter as to build an understanding of the status of DNA in the court of law, its admissibility, and ethical concerns.

Keywords

Forensics · Criminal justice system · DNA profiling · DNA databank · Legal perspective · The DNA Technology (Use and Application) bill 2019

A. Srivastava (✉) · A. Harshey
Dr. A.P.J. Abdul Kalam Institute of Forensic Science and Criminology, Bundelkhand University, Jhansi, Uttar Pradesh, India

P. Shrivastava
DNA Fingerprinting Unit, State Forensic Science Laboratory, Department of Home (Police), Government of Madhya Pradesh, Sagar, Madhya Pradesh, India

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_31

31.1 Introduction

The beginning of human civilization and its evolution, from the stone-age to the modern era, is one of the significant areas of interest in social sciences. Since humans are the social element, therefore, the civilized society has its value-based rules and regulations to ensure one's activities and to protect the rights. In socio-legal terms, the breach of these rules and regulations is anti-social and is often referred to as a crime. Although in the legal context, the definition of crime also depends upon the law of the land. To ensure justice against crime is essential to maintain law and order. Scientific and technological advancements have greatly influenced the lifestyle of a common man. It is worthy to note that now technological advances are being used as a tool of crime. That is the reason, with progressive changes and advances in science and technology, new trends of crime are also being emerging. The potential of science in the administration of justice has been well recognized. The information extracted from the scientific analysis is of the crucial evidential significance. Contribution of the science in a criminal investigation is referred to as forensic science that is an integral entity of the criminal justice system (CJS) that provides strong corroborative evidence. Forensics not only provides scientific evidence but has also helped to exonerate the innocent and thus strengthens the path of fair justice (Jayewardene 1988; Brown 1998; Sedley 2005).

Adoption of the scientific discoveries and inventions to both life of a common man and CJS put several considerations to the policymakers. Laws and regulations that govern the society ensure some fundamental human rights such as the Right to Privacy and Right to Self-Incrimination. Before the adoption of new scientific advent, it must be ensured that conventional rules, regulations, and human rights are intact and are not getting affected. It is of utmost importance to create balance among the newly developed technique and traditional and constitutional rights, e.g., Right of Privacy and Right of Self-Incrimination (Kirby 2001; Walsh 2005).

Sometimes, it is impossible to optimally utilize the scientific advances without the breach of human rights; therefore, any other legislation or modification to the existing regime is required. At the end of policymakers, it is essential to reduce the possibilities of misuses of technology and new rules and regulation against the public interest.

Research and development echelon of forensic genetics raised several new possibilities with the discovery of the deoxyribonucleic acid (DNA). DNA technology has efficaciously contributed as strong corroborating evidence in the court of law (Kirby 2001; Adhikary 2007; Kumar et al. 2016).

31.2 Development of DNA Technology and Impact Across the Globe

The discovery of DNA, in 1950, tagged DNA as the universal genetic material that inherits from one generation to another. With time, the technique of DNA profiling has been revolutionized with the introduction of several technological

developments. Many developments are going onwards, but it is notable that from restriction fragment length polymorphism (RFLP) to the short tandem repeat (STR) and next-generation sequencing (NGS), fundamental principle, i.e., DNA is a unique genetic code that can demonstrate the genealogy, remains the same (Butler 2015a).

The advent and result oriented increasing use of DNA fingerprinting has broadly influenced the criminal justice system. In 1985, it was an immigration dispute (a civil dispute) in which DNA technology was applied for the first time. However, in 1986, the case of Colin Pitchfork was the first criminal case that was decided on the grounds of DNA evidence. Endeavors of Prof. Alec Jeffrey of Leicester University, UK, resulted in the identification of the culprit (Colin Pitchfork Case, 1986) in England. In this case, DNA also served as an evidential tool corroborating the acquittal of an innocent (Jeffreys et al. 1985; Jobling and Gill 2004).

Development of polymerase chain reaction (PCR) and STR-based DNA analysis significantly contributed to a forensic DNA analysis. Since, many times the biological sample (source of DNA) undergoes several adverse conditions (such as exposure to the high temperature, humidity) followed by degradation as well as the sample may present in minimal quantity. These techniques enabled the generation of DNA profiles from the degraded samples, thus countered the problem. DNA phenotyping is an emerging trend in forensic genetics, attracting experts and scientists across the globe (Edwards et al. 1992; Butler 2015a; Shrivastava et al. 2016).

The ‘establishment of event or linkage to crime’ is referred to as the body of crime for what the term corpus delicti is often used in forensics and legal proceedings. It requires the establishment of personal identity. Several methods such as fingerprints, anthropometry and poroscopy are well-recognized and -accepted practices for personal identification. Identification from skeletal remains, body fluids, and mutilated bodies presents limitations to the techniques described above. DNA typing has the potential to counter these limitations with a high level of accuracy. Enormous aid of DNA to the criminal investigation has designated it as a “new gold standard” in forensic science (Lynch 2003; The Royel Society and The Royel Society of Edinburgh 2017; Wrobel et al. 2018).

DNA analysis has provided scientific evidence beyond reasonable doubt and played a vital role in the administrations of justice in various criminal and civil cases. The DNA report has produced the grounds indicating the innocence of a suspect or a person who has wrongfully convicted. Project Innocence is the demonstration that DNA cogent gateway to the fair justice (Shrivastava et al. 2016; Steinback 2007; Goswami and Goswami 2018).

Human trafficking, child abuse, illegal adoption, etc. are some significant concerns to law enforcement agencies. DNA has found to be a useful tool for identification and rehabilitation. DNA PROKID is a project initiated by the University of Granada Genetic Information Laboratory, Spain, with the support of the Government of Spain, in 2006, with the prime objective of the identification and rehabilitation of the potential victims of human trafficking, i.e., children and women. Questioned and reference databases are being generated with the help of DNA PROKID Kit under this project. DNA PROKID kit offers a simple collection of

samples and nullifies the probabilities of the contamination (https://dna-prokids.org/que-es-dna-pro-kids/). The identification and rehabilitation of a considerable number of children (more than 900) to their families pronounce the potential of DNA to deal with these kinds of issues (Alvarez-Cubero et al. 2018).

31.3 DNA in Courtroom: Influence to the CJS

DNA typing has significantly contributed to the administration of justice in various civil as well as criminal cases. The immigration case in the UK was the first civil dispute solved with DNA evidence. In the famous Colin Pitchfork case, DNA evidence not only identified the offender but also saved an innocent. The potential of DNA evidence and its application in various investigations has established it as "Genetic eyewitness" (Bureau of Justice Statistics 1991; Dodd and Laville 2012).

DNA is a genetic signature of an individual and cannot be the same in two persons (exception of identical twins). This fact is originated from the scientific research and found the grounds of admissibility of the DNA evidence in the court of law. The technological and scientific developments in the field of genetics have influenced society as well as CJS to a great extent (Adhikary 2007; Jeffreys et al. 1985; Verma and Goswami 2014).

Technological developments have enhanced the credibility of the DNA evidence. But here it is notable that DNA evidence is not infallible, and its admissibility and validity have also debated in the court of law (Verma and Goswami 2014; GrubbA 1993; Virkler and Lendev 2009). These debates and trials yield some protocols and standards, e.g., Frye test, Daubert test, prejudicial effect test, and usefulness standard, directing the grounds for validity assessment of the evidence. Information presented by the evidence forms the territories to decide the validity. Undoubtedly, if any evidence is supposed to produce confusion to the jury or any misleading information, it cannot be admitted in the trial (Lynch and McNally 2003). The 'tampering with evidence' is the most common allegation that dominantly affects the validity and the admissibility of the evidence—while in the view of the improper practice of the technique and substandard methodology, the accused may be given the benefit of the doubt. While admitting the scientific evidence, chances of contamination, secondary and tertiary transfer of DNA, false inclusion, and exclusion have utmost significance in legal and experimental contexts and put roles of great magnitude on the reliability of the evidence (Alvarez-Cubero et al. 2018; Murphy 2017).

To anticipate the advances of DNA technology to the CJS in compliance with the law, many nations have enacted DNA Act to provide the statutory basis to the DNA profiling as well to establish the DNA databank. In 1995, England established the world's first databank. Later on, several countries appreciated databank and proceeded in this direction. Establishment of DNA databank as well as the applicability of the DNA evidence also presents some challenges to the judiciary (Adhikary 2007; Goswami and Goswami 2018; Verma and Goswami 2014; Committee on DNA Technology in Forensic Science, USA Board of Biology, Commission on Life

Science, National Research Council of USA 1992; Kunkel 2018). This chapter looks toward the painstaking attempts of the judiciary members and policymakers to keep the balance between the rights and the technology. Indian legislative perspective of the DNA evidence becomes significant as the government of India is in the process of enacting the bill.

31.4 Legal Dimensions of DNA Evidence in the United States of America

The court evaluates the admissibility of the scientific information before including it. In the USA, the assessment of the reliability of the evidence initiated in 1923 when the court, in *United States v Frye*,[1] denied to relay upon the lie detector evidence and stated that the relevant scientific community had not accepted the technology. The court emphasized that the technique, procedure, or scientific principle deducing the expert testimony must be sufficiently established in the particular field. To keep adverse scientific evidence out of the trial, prosecution and defense pleaded in the light of Frye standard. In brief, the Frye standard identifies the particular field into which the discovery or scientific principle may be categorized and acceptance of the principle, invention, or experimental procedures within the concerned scientific community (Committee on DNA Technology in Forensic Science, USA Board of Biology, Commission on Life Science, National Research Council of USA 1992).

It is notable that theoretically, the role of the court is limited to the assessment for the acceptance of scientific technique underlying the evidence in the concerned regime and reliability of scientific technology for the intended purpose. When selecting the expert from forensic genetics, due to its complexities, the role of the court becomes essential and critical. While dealing with the DNA evidence, admissibility is judged on the validity of assumption, i.e., DNA is a unique genetic code (except for the identical twins), the used procedure allows the profiling of the DNA samples, and available databank and statistical methods enable the assessment of the match probability. Acceptance of these assumptions is followed by the evaluation of the procedure performed in the particular case in compliance with the available standards. Frye standard delineates rejection of the DNA evidence extracted from the improper and substandard practice even if DNA profiling is well established and accepted in the scientific domain.

Despite the reliability of DNA typing, its admissibility was successfully challenged in the court of law. In 1976, *People v Kelly*,[2] held in California court, was a landmark decision concluding that correct scientific procedures were exercised must be demonstrated by the proponent of the evidence. *People v Castro*[3] is another flagship case which did not accept the DNA evidence due to the flawed practice of

[1] 293 F 1013 (DC Cir 1923).

[2] 144 Misc2d 955 NVS 22 d 985 (Sup Ct 1989).

[3] 545 N.Y.S.2d 985 (Sup. Ct.1989).

the technique. *People v Castro* also emphasized that the admissibility not only requires the correct procedures to be used but also the assurance that the correct procedure was performed correctly. DNA evidence was rejected by the appellate court for the first time in 1989 when the Supreme Court of Minnesota deduced that compliance of the laboratory with available proper standard and testing data is the foundation on which admissibility of the evidence rests. The quality of the procedures followed hereby decides the probative force of the evidence. California Court of Appeals made a significant remark that admissibility of the evidence, as per Frye standard, requires technique used to be supported by a clear majority of the experts of concerned filed rather than unanimity of views on support (Adhikary 2007).

31.4.1 The Federal Standard and Daubert's Standard

Overtly conservative approach and vulnerability to the manipulation of the Frye standard was criticized for excluding the novel scientific pieces of evidence. In 1975, the Federal Rules of Evidence (FRE) were introduced. Since appellate courts address few concerns related to the facts of the case, FRE governs in the trial court during the initial presentation of the evidence. FRE does not repudiate the Frye rule though it follows a more permissive approach. Efficiency, reliability, relevance, etc. are central themes of the FRE. It has provisions to exclude the evidence producing misleading facts and undue prejudice or confusion to the jury (Adhikary 2007; Committee on DNA Technology in Forensic Science, USA Board of Biology, Commission on Life Science, National Research Council of USA 1992).

In 1993, Supreme Court rejected the criticism of FRE in the hearing of *Daubert v Merrell Dow Pharmaceutical Inc.*[4] The court pronounced that the incompetency of the Frye has superseded by FRE. Daubert's case tested the reliability in terms whether the falsifiability and testability of any evidence have been tested or can be tested; whether the technique has peer-reviewed publication as it increases the chance of flaw detection; maintenance of technique operational standard and assessment of potential error rate and assurance whether the technique has accepted sufficiently in the relevant community. Admissibility and reliability of the opinion evidence are judged by scientific analysis in combination with opinion differences.

While administrating the reliability of evidence, the choice of standard has been subjected to the discussion among the legal fraternity of the USA. It is worthy to note Justice Blackmun's pronouncement over the admissibility of the evidence. According to him, the inclusion of the evidence should be restrictive and cautious as well as there should be a logical rejection of the relevant evidence if evidence may cause confusion, unfair prejudice or mislead to the Court. Therefore, as per rule 403 of FRE, the probative force of expert testimony should be assessed. It is not appropriate to take only the strengths and weaknesses of the scientific evidence into

[4] 113 S Ct 2786 (1993).

account. Though, the comparison of scientific evidence with other evidence should be made to assess the justifiability of differential treatment of scientific evidence (Adhikary 2007).

'Convicted by Juries, Exonerated by Science' project unmarked the error in 28 cases by exonerating the previously convict of rape or murder by means of DNA testing. In most cases, DNA evidence served as evidence discovered after the trial. DNA tests were conducted independently, and in some cases, results from DNA tests were sent to another laboratory for cross-validation. Authorities asked for the preservation of the evidence for further testing since the previous examination destroyed the sample or proof. Furthermore, in some cases, in compliance with the chain of custody was observed to be broken, and thus, the integrity of the investigation was put into the peril. This project also emphasized the establishment of the criteria for the inclusion and reliability of the eyewitness. The literature recommends that the reliability of eyewitness testimony (pros and cons) should be tested with the use of expert testimony (National Department of Justice (NIJ) Research Report, US Department of Justice 1996).

31.4.2 The DNA Identification Act of 1994

Accreditation and quality control are two points that have been asked by the judiciary. The National Research Council found that the adequacy of laboratory procedures and its compliance with standards and competence of experts are always questionable. Consequently, the National Research Council has emphasized the need for external accreditation. In 1988, the Technical Working Group on DNA Analysis Methods (TWGDAM), consisting of forensic DNA experts of government and private sector from the USA and Canada, was established under the sponsorship of FBI laboratory division. TWGDAM issued guidelines from time to time (in 1989, 1991, and 1995) for the quality assurance of DNA analysis. In 1991, joint House and Senate Congressional hearings recommended the federal legislation by recognizing the application of DNA technology for criminal identification and the need for procedural and quality control standards for the DNA test (Adhikary 2007).

Enactment of DNA Identification Act 1994, led the establishment and improvement of the Forensic DNA testing laboratories through the provided grants. In 1995, the DNA Advisory Board (DAB) was established and funded to formulate DNA testing standards under the provisions of this act. This bill also emphasized the national DNA profile indexing of convicted offenders. During the tenure of Nobel Laureate Dr. Joshua Lederberg, in 1998, DAB published Quality Assurance standards for Forensic Testing Laboratories. Under the chairmanship of Dr. Arthur Eisenberg, DAB finalized and produced Quality Assurance Standards for Convicted Offender DNA Databasing Laboratories. The director approved both the measures of the FBI. This act is the legal framework of DNA testing and database in the USA, and these standards ensure high-quality testing (Committee on DNA Technology in Forensic Science, USA Board of Biology, Commission on Life Science, National Research Council of USA 1992).

31.4.3 The NAS Report, 2009

In light of some cases of erroneous forensic examination, to ensure the quality control, the accuracy of the investigation, development, and enrichment of forensics, in 2005, the US Congress authorized the National Academy of Science (NAS) to initiate a study on forensics. In this consequence, after the number of meetings and deliberations, NAS' National Research Council (NRC) published a report titled "Strengthening Forensic Science in the United States: A Path Forward" in February 2009. This framework is often referred to as NAS report, and the need and significance of its recommendations have been felt across the globe (National Academy of Sciences 2009; Butler 2015b).

The NAS report proposed 13 recommendations to counter the existing problems as well as to strengthen the future pathway. NAS report advocated setting up of National Institute of Forensic Science (NIFS) to establish and enforce good laboratory practices, formulation of standards, arrangement of funding for compulsory accreditation and certification to promote the quality and peer-reviewed research, to provide smooth funding thus giving substantial financial support to the Forensics, to enhance forensic research and education system, to identify the existing problems of forensic science and its limitations in the legal arena and anticipate newly discovered scientific technology in the judicial filed to reinforce the crime investigation. Judge Harry Edwards (Co-Chair NAS committee), outlined the importance of the NAS report and challenges to the forensic science in the courtroom while anticipating discoveries and conversed as

> "What our committee found is that, although there are many dedicated and skilled forensic professionals, the quality of practice in the forensic disciplines varies widely, and the conclusions reached by forensic practitioners are not always reliable. From my vantage point, the response to the Report has been very positive, and I have seen a groundswell of support in favor of major reforms to correct the ills of the forensic science community. If courts blindly follow the precedent that rests on unfounded scientific premises, this will lead to unjust results. When scientific methodologies once considered sacrosanct are modified or discredited, the judicial system must accommodate the changed scientific landscape." (Edwards 2010)

Several commissions and committees were founded to strengthen the forensic practice, quality assurance, and the inclusion of new technological advents. In 2009, the National Science and Technology Council Subcommittee on Forensic Science (NSTC SoFS) was established. For the purpose of policymaking and to improve the practice in forensic science, in 2013, the US Department of Justice (DOJ) and National Institute of Standards and Technology (NIST) jointly founded National Commission on Forensic Science (NCFS) as Federal Advisory for the DOJ (Butler 2015b; Charter of the Subcommittee on Forensic Science, Committee on Science National Science and Technology Council 2015; Renewed Charter of the Subcommittee on Forensic Science 2015).

31.4.4 Existing Issues Concerning DNA Databank and DNA Phenotyping

The USA established, in November 1997, 13 STR markers-based Combined DNA Index System (CODIS) followed by the establishment of the National DNA Index System (NDIS) (Butler 2006). The impact and need for DNA databases have been discussed in the literature. It may be a very reliable tool in the investigation of cross-border crimes. The advent of phenotyping in 2008 has created a new perspective in criminal investigation and research and development arena of the Forensic Genetics domain. Several debates related to databank and NGS system are available, demonstrating the response of existing law to these issues, impact of these on the Right to Privacy, scope of further development within the limits of the law, lack of public awareness on the significance of DNA database, possibilities to modify the regulations for optimal anticipations of scientific advancements, and reducing the risk of misuse of technology against the society.

Here it is worthy to note the concept of transitional exchange of DNA database that involves an international DNA database that includes linking of national databases and request-based exchange of data. In 2002, the INTERPOL DNA gateway platform was established as a global international database. INTERPOL DNA database is contributed by 84 countries and stores only DNA profiles of convicts, suspects, and missing persons rather than the personal information of the subjects. Profiles originating from any contributing country are subject to the national laws of that concerned country. As far as the regional international databases are concerned, the Europol Information System (EIS) established in 2005 (repository of DNA profiles from European Union member states) is one of the examples of this. The application of these kinds of the database is limited to criminal investigation only (INTERPOL 2009).

To facilitate the optimal application of the DNA technology with its new advents and DNA databank in the courtroom, a need for public awareness has felt. Studies on views of the public on DNA databank are needed to counter the existing problems. Despite being applied in a few criminal cases, in most jurisdictions, the use of forensic DNA phenotyping and familial searches remain unregulated. Whether the police should be permitted to access the profile from the DNA database is a question that is subjected to discussion for the entire scientific and law enforcement echelon. A need for more harmonized policies to protect public security and individual liberty and rights has been felt (Amankwaa 2019).

31.5 Legal Dimensions of DNA Evidence in the United Kingdom (UK)

DNA profiling was first applied in England in 1986 for the investigation of a criminal case. It not only apprehended the culprit but also saved the innocent. The legal foundation in the UK classifies blood as an intimate sample, and therefore, consent is needed for sample collection (certain conditions are an exception where sample

could be collected against the consent, e.g., testing for diseases such as HIV) and thus protects the privacy. On the other hand, the law designates buccal scrapes and hair roots as a non-intimate sample; therefore, sample could be collected by force if necessary. As per the recommendations of the Royal Commission on Criminal Justice 1993, police should be allowed to collect non-intimate samples from suspects. Forensic Science Service contributed efficaciously to the feasible establishment of the National DNA Database (NDNAD).

In 1994, with enacting the Criminal Justice and Public Order Act, the legal foundation for the National DNA database was established. This act authorized the police to collect the samples from any person charged with any recordable offense and search for matching the profiles that can be performed. This bill enabled the collection of DNA samples of arrested individuals and thus makes the investigation process rapid. It also created a complete record of the active criminal population, thereby eliminating the innocent as a suspect first. Retention or recording of physical samples has been subject of debate. The UK is credited as the first country that established a national DNA database for criminal investigation and detection in 1995. The Criminal Justice and Public Order Act 1994, Criminal Procedure (Scotland) Act 1995, and Police (Amendment) (Northern Ireland) Order 1995 made provision authorizing the retention of the information even if the subject is not guilty in England (in Wales also), Scotland, and Northern Ireland, respectively. Earlier, in 2001, White Paper on CJS suggested that in Scotland, there should be legislative grounds for the retention of voluntarily given fingerprints and DNA samples. As per the Paper, samples, derived from an individual who has progressed through the CJS and found to be not guilty, cannot be retained. This reflects the contrast to the Criminal Justice and Police Act 2000 (UK) that allows the retention of samples regardless of the aforesaid conditions. White Paper recommended that sample collection by non-intrusive means does not require any consent. All of these recommendations were included in the Criminal Justice (Scotland) Bill 2002.

The use of the information derived from the samples or database has been challenged in the court of law. Earlier, the House of Lords Select committee on science and technology suggested the formation of an independent body to supervise the functioning of the database. *R v Willoughby* is one of the landmark cases in which the application of the database was pleaded as a contravention of the Criminal Justice and Police Order Act 1994. As per the provisions of this act, samples should be destroyed after used, and this information cannot be used for the investigation purpose. Attempts to calculate statistical probabilities are out of the investigation sphere (Adhikary 2007; Amankwaa 2019; Christopher and Asplen 2001).

In another case, *R (on the application of S) v Chief Constable of South Yorkshire*, the court of appeal distinguished among the "taking," "retention," and "use" of samples, i.e., fingerprints or DNA. The amendment to the Police and Criminal Evidence Act 1984 as the insertion of section 64(1A) provided the statutory basis for the retention of the samples. It was pleaded that these amendments are not in compliance with the provision of protection of privacy that is provided in articles 8 and 14 of the Human Rights Act. The Court of Appeal upheld the legislation directing the preservation of the samples. *Saunders v United Kingdom* demonstrates

the court's view distinguishing the identification and self-incrimination. While in *S and Marper v. the United Kingdom*, the court found retention of samples to be a substantial threat to privacy (Law Commission of India, Department of legal affairs, Govt. of India 2017).

Enactment of the DNA legislation in compliance with fundamental human rights is a big task for policymakers. In the past, various countries have enacted the DNA specific legislation demonstrating provision and guidelines to collect samples, the statutory basis for the retentions of the sample, and the use of information for the investigation purpose. Table 31.1 presents a glance view of the legal provision of different countries for the sample collection and its retention.

31.6 DNA Evidence in India

31.6.1 Current Perspective and Existing Facilities

In India, DNA technology was first applied in the case of a paternity dispute. In 1989, the Center for Cellular and Molecular Biology (CCMB) was the first institute that conducted the DNA test for forensic purposes. Later, in 1998, a specialized center named as Center for DNA Fingerprinting and Diagnostics (CDFD) was established to conduct DNA testing for legal purposes. All the Central Forensic Science Laboratories (CFSLs) and State Forensic Science Laboratories (SFSLs) are well equipped to perform DNA profiling. Recently, the Government of India has provided funds to strengthen the DNA facilities in the laboratories. Under this step, the laboratories are getting strengthened with the all modern sophisticated and updated instrument. In India, there are two institutions severing DNA testing for the wildlife forensic purpose that are Wildlife Institute of India (WII) and Laboratory for the Conservation of Endangered Species (LaCONES) situated at Dehradun and Hyderabad, respectively. As far as the area of the research is concerned, STR marker-based studies of the diverse Indian population and studies establishing the genealogy are current trends (Verma and Goswami 2014; Shrivastava et al. 2019; Srivastava et al. 2019a, b, 2020; Gupta et al. 2016).

Some existing laws regulate the practice of DNA. The Indian judiciary has appreciated DNA evidence. DNA has revolutionized the criminal justice system and played a crucial role in resolving several criminal as well as civil cases. Since the Indian laws do not anticipate modern technology, contradictions and challenges arose due to the absence of specific legislation that puts forth a need for it.

31.6.2 Right to Privacy and Right Against Self-Incrimination vs. DNA Evidence

The Constitution of India (COI) protects the rights of the citizens. Right to Privacy and Right against Self Incrimination is defined as fundamental rights under Article 21 and Article 20 (3) of COI (Patel et al. 2013). Maintaining a balance between

Table 31.1 Legal provision in different countries: at a glance (Law Commission of India, Department of legal affairs, Govt. of India 2017)

Sr. No.	Country	Subjects of sample collection	Policies over sample retention	Concerning legal foundations
1	Austria	Convicted persons, suspects charged of a serious offences and all crime scene stains	Convicted persons "samples must be destroyed when individual reaches age 80, suspects samples are retained despite suspects" acquittal [a written request for destruction must be submitted]	State Police Law (SPG)
2	Belgium	Persons convicted of a "serious offence" and crime scene stains when ordered by a prosecutor	Convicted persons "samples must be destroyed once DNA profile is created, suspects" profiles must be destroyed once the prosecutor has determined that a suspects" request for independent DNA analysis will not be granted or when the result of such request has been communicated to the suspect	Law of March 22, 1999 Royal Decree of February 4, 2002
3	China (Hong Kong)	Persons convicted of any "serious arrestable offence" and all crime scene stains	Convicted persons' samples must be destroyed as soon as they are practicable from such time as there is no other charge against the person in relation to an offence which renders the retention of the sample necessary and all proceedings (including any appeal) arising out of the conviction have been concluded; suspects' samples must be destroyed as soon as they are practicable 12 months after the sample is taken if they	The Dangerous Drugs, Independent Commission Against Corruption and Police Force Ordinance Independent Commission Against Corruption Ordinance Police Force Ordinance

(continued)

Table 31.1 (continued)

Sr. No.	Country	Subjects of sample collection	Policies over sample retention	Concerning legal foundations
			are not charged with any offense, or if so charged when all charges are withdrawn, the person is discharged by a court before conviction of the offence or all the offences, or they are acquitted of all charges	
4	Cyprus	All convicted persons, suspects, and crime scene stains	All samples follow fate of DNA profile	Police Law
5	Czech Republic	All convicted persons and crime scene stains	All samples follow fate of DNA profile	Criminal Procedure Act Law on Police Binding Instruction No. 88/2002 of the President of the Police
6	Denmark	Convicted persons, suspects charged of an offence that could lead to a prison sentence of 1½ years or more, and all crime scene stains	All samples follow fate of DNA profile	Law Establishing a Central DNA Profile Register
7	Finland	Convicted persons serving a prison sentence of 3 years or more, suspects charged of a crime that could lead to a prison sentence of 6 months or more, and all crime scene stains	Convicted persons' samples must be destroyed 10 years after their death, suspects' samples must be destroyed within 1 year of a prosecutorial determination that there is no evidence of an offence, charges have been dismissed, when their sentence has been nullified, or 10 years after the suspects' death if not removed earlier	Coercive Measures Act Police Act Police Personal Data File Act
8	France	Persons convicted of or charged with a serious offence (list in law) and crime scene	Convicted persons' samples are retained for 40 years after their conviction or until	Code of Criminal Procedure Law No. 98-468 of June 17, 1998; Law

(continued)

Table 31.1 (continued)

Sr. No.	Country	Subjects of sample collection	Policies over sample retention	Concerning legal foundations
		stains when deemed relevant	their eightieth birthday; suspects' samples are kept until conviction or acquittal, i.e., procedurally, DNA samples are treated as regular evidence	No. 2001-1062 of November 15, 2001; Law No. 2003-239 of March 18, 2003; Decree No. 2000-413 of May 18, 2000; Decree No. 2002-697 of April 30, 2002; Decree No. 2004-470 of May 25, 2004; Decree No. 2004-71 of May 25, 2004; Decree No. 2009-785 of June 23, 2009. Deliberation No. 2008-113 of May 14, 2008; Circular of the Ministry of Justice of 27 July 2004
9	Germany	Persons convicted of a serious offence or repeatedly committing the same minor offence, suspects charged of a serious offence, and crime scene stains when related to any recordable offence	Convicted persons" and suspects" samples must be destroyed when they are no longer considered useful for investigatory purposes	Act LXXXV of 1999 on the criminal records and certificates on criminal record
10	Hungary	Persons convicted of one of the crime categories which are listed in law, suspects charged with an offence that could lead to a prison sentence of 5 years or more or that is listed in law, and all crime scene stains	Convicted persons' samples must be destroyed 20 years after their conviction, and suspects' samples must be destroyed upon their acquittal or abandonment of the underlying investigation or proceeding	Law on Police DNA File Police DNA File Regulations
11	Netherlands	Persons convicted or suspected of any recordable offence and all crime scene stains	Convicted persons' samples are retained indefinitely; suspects' samples must be destroyed upon their acquittal	
12	Slovakia	Persons condemned to punishment other than a fine, all	All samples must be destroyed "as soon as	The Act n. 417/ 2002—use of DNA analysis for

(continued)

Table 31.1 (continued)

Sr. No.	Country	Subjects of sample collection	Policies over sample retention	Concerning legal foundations
		suspects, if warranted by possible prison sentence, and all crime scene stains	possible" [GET QT FROM LAW]	identification of persons
13	Sweden	Persons serving a prison sentence of 4 years or more, suspects charged of an offence that could lead to a prison sentence of 4 years or more (approval of prosecutor is required), and all crime scene stains	Have to be destroyed 20 years after their creation for individuals sentenced to no more than 6 years, 30 years for individuals sentenced to more than 6 years, or at most 20 years after the individual's death; suspects' samples must be destroyed upon their acquittal	Code of Judicial Procedure Police Data Act
14	UK	Persons convicted of any recordable offence, arrested for any recordable offense, and all crime scene stains. For Scotland: Persons convicted of any recordable offence, arrested for any recordable offense, and all crime scene stains	All samples are retained indefinitely. For Scotland: Convicted persons' samples are retained indefinitely, but suspects' samples must be destroyed upon their acquittal or when no criminal proceedings are initiated	Police and Criminal Evidence Act (PACE) Criminal Justice and Public Order Act (CJPOA) Criminal Evidence (Amendment) Act Criminal Justice and Police Act (CJPA) Criminal Justice Act (CJA) 2003 Data Protection Act (DPA) Human Rights Act (HRA)
15	USA	Twelve states have laws authorizing arrestee sampling. All 50 states require that convicted sex offenders provide a DNA sample; 46 states require that all convicted felons provide a DNA sample. Eleven states specify certain misdemeanor among those who must provide a sample. There are 28 states	The criteria for retention vary from immediate removal, if a sample is not used, to retention of a sample for at least 35 years, to permanent retention for certain specified offences	The DNA Identification Act 1994 Justice for All Act 2004 Violence Against Women Act 2005

(continued)

Table 31.1 (continued)

Sr. No.	Country	Subjects of sample collection	Policies over sample retention	Concerning legal foundations
		that include DNA from delinquent juveniles in the database		

Also see: Thibedeau AD (2011) National DNA database 2011. National Forensic DNA database council for the responsible genetics. http://www.antoniocasella.eu/dnlaw/DNA-data2011.pdf

technology and constitutional rights is a painstaking task to the judiciary. *Selvi v. the State of Karnataka*[5] is one of the flagship cases in the history of the Indian judiciary in which it was decided that the court does not possess an inherent power to direct the sample collection of an Individual. Further, the court observed that consent becomes essential if the technique in invasive. Court has more right to order a DNA test in the criminal case rather than a civil trial. Furthermore, the court finds that if a person refuses to submit the sample, it cannot be a ground to take an adverse inference. At the same time, section 114 of the Indian Evidence In 1872 (IEA) directs to draw an adverse conclusion if a person denies producing a piece of evidence under his possession (Goswami and Goswami 2018).

The supreme court of India found that, in *State of Bombay v. KathiKaluOghad*,[6] submission of a specimen of finger or foot does, handwriting or signature prints of palms cannot be considered a breach of the Right to Privacy, and therefore, consent is not essential. While in the case of *BhabaniPrasadJena v. Convener, Secretary, Orissa State Commission for Women,*[7] the apex court observed that the court should decide the important need of the DNA test in search of the truth, on its discretion, after balancing the rights. The insertion of sections 53-A and 164-A in the Code of Criminal Procedure (CrPC) by CrPC (Amendment) Act 2005 provisions direct investigator and the registered medical examination practitioner to conduct the medical test of a rape victim as well to collect the sample of the accused or culprit regardless of his consent (Gupta et al. 2016).

In this consonance, few more judgments are there that demonstrate the views of the Indian judiciary on the Right to Privacy and Right against Self-incrimination. Neither CrPC nor IEA has provisions directing the sample collection of a minor or his mother in connection with maternity dispute.[8] As per the Court's observation in *SubayyaGounder v. BhoopalaSubramaniam*,[9] against the will of the subject, the court cannot order to give the blood sample. Collection of blood from the veins cannot be regarded as the breach of Right to Privacy as held by Rajasthan High

[5]*51. 2010 7 SCC 263.*

[6]AIR 2014 SC 932).

[7]*AIR 2010 SC 2851.*

[8]Venkateshwarlu vs. Saubbayya (AIR 1951 MAD 910).

[9](1959 Cr LJ 1087 MAD).

Court.[10] A positive view of the court toward the DNA testing in matrimonial cases may be observed in *Sharda* vs. *Dharmpal*[11] (Adhikary 2007; Goswami and Goswami 2018; Law Commission of India, Department of legal affairs, Govt. of India 2017).

31.6.3 Presumption of Paternity

The introduction of DNA technology in Indian courts was initiated with the paternity dispute. Indian legislation gives a special status to the legitimacy.[12] If the husband fails to prove no access to the wife, the child shall be considered as born from their legal wedlock as per the provisions of Section 112 of IEA. Actually, in paternity disputes, special care and discretion are needed to satisfy the rights of the child as well as the father. DNA reveals the biological truth of the child and may put the right of a child in peril. Therefore, legitimacy has referred to as conclusive presumption (with the exception of a particular condition), and the court observed that although DNA technique is well established and as well as accurate, nonetheless, it is not only enough to escape from Section 112 of IEA.[13] *GoutamKundu v State of Bengal*[14] is a very famous paternity dispute in India. In this case, the apex court formulated a few guidelines to conduct the DNA test in complaints with the legal provisions. In brief, the instructions may be extracted as the DNA test is not a matter of the routine. The husband must prove no access to the wife as a strong prima facie. The court may order DNA on its discretion only after deciding that it does not threaten the rights of children and women. The subject cannot be forced to give a sample, and his refusal to provide a sample is watched over by Article 20(3) and 21 of COI. But it appears that the court itself took the way out of these guidelines in *N.D. Tiwari v RohitShekhar*.[15] These guidelines have supposed to create obstacles in the criminal investigation. The case studies reveal that with time the court has relied on the DNAtest. To solicit the scientific advances on legal grounds, Malimath Committee[16] has suggested some revision to be incorporated in the Section 112 of IEA (Goswami and Goswami 2018; Verma and Goswami 2014).

---

[10] SwatiLodha vs. state of Rajasthan (1959 Cr LJ 1087 MAD).

[11] 2003 4 SCC 493.

[12] Yasu vs. Santh (1975, Ker Lt 533).

[13] Smt. KaomtiDevi v. Poshiram(2001) 5 SCC 311.

[14] 1993 (2) Scale 994. O. 6/95.

[15] FAO(OS) No. 547/2011.

[16] Ministry of Home Affairs, Govt. of India Report, Reforms of Criminal Justice System (2003) vol 1.

31.6.4 Other Issues and Ethical Concerns

Several ethical challenges are associated with the practice of DNA. It is of utmost care to reduce the chances of misuse against the public interest. Specific legislation provides the statutory grounds for the practice of DNA tests. Standards of examination, accreditation of the laboratory, qualification of the expert, etc. are some significant issues that directly influence the reliability of the evidence. With the emerging technological advances, some concepts such as fabrication and theft of DNA have arisen (Verma and Goswami 2014; Harshey and Srivastava 2018). The UK has categorized DNA theft as an offense (Laurie 2003). A DNA profile cannot distinguish between the genuine and falsely planted samples. In the recent past, a study demonstrating in vitro synthesis of specific DNA, particularly with any biological sample, has been reported. A strict guideline for DNA practice is a must to avoid the misuse of it as well to safeguard to the privacy since DNA databank is supposed to be a threat to privacy. Care should be there, as it may lead to the miscarriage of justice (Bolden 2011; Frumkin et al. 2010; Kunkel 2018; Tang et al. 2020). The advent of the assisted reproductive technique (ART) and bone marrow transplantation (BMT) has posed a need to take special care while conducting the DNA test (Batch et al. 1968; Gatti et al. 1968; Pope and Chapman 2006; Motluk 2011; Goswami 2015, 2016; Dauber et al. 2004). The need and benefits of the establishment of a DNA databank along with its ethical concerns have been well discussed. Databank also bears some ethical issues along with it (Amankwaa 2019; Guillén et al. 2000; Yadav 2017; Joy et al. 2018; Wallace et al. 2014; Cordner 2001; García et al. 2017; Machado and Silva 2019; Jakovski et al. 2017).

DNA Profiling Bill 2019 also directs the provisions for the establishment of the databank. Indian Government machinery is exercising to draft a bill for the DNA practice since 2003. For several times, the draft framework was criticized as privacy issues were not properly addressed. In 2019 lower house of Indian Parliament, i.e., Lok Sabha, has approved the DNA Technology (Use and Application) Regulation Bill 2019, and now it is pending in Rajya Sabha that is Upper house. This bill will provide the statutory basis for the identification of the victim, culprit, missing person, and unknown dead bodies utilizing DNA profiling. This bill constitutes ten chapters directing provisions of the DNA Regulatory Board, funding for smooth operation, the establishment of DNA databank, accreditation procedures, and organization of a DNA laboratory. This bill also describes the mechanism to secure the genetic information of the subjects and penalties for the disclosure or misuse of the data. This bill will strengthen the future pathway of DNA profiling in India on legal grounds (The DNA Technology 2019).

31.7 Conclusion

Since the development of DNA technology, it has been appreciated and well believed in the criminal justice system. It reveals that the accuracy of DNA profiling does not mean that it is infallible. Distinguished care is needed while admitting the

DNA as evidence. A number of ethical issues are associated with DNA technology, posing serious challenges. For the optimal anticipation of the DNA in the criminal justice system, a need for public awareness has felt. Balancing the rights with technology is of utmost importance. On reviewing the legal frameworks of different nations, it is evident that all the countries have erected the law intending to deliver justice by guarding the fundamental rights and eliminating the chances of miscarriage of justice. Legal provision has emphasized the standard procedures to ensure the reliability of the DNA evidence. Accreditation and the qualification of the expert are crucial factors while assessing the authenticity of the evidence. In the current scenario, to strengthen the future perspective, a need for publicly available information about the forensic potential of DNA is needed since it will lead to increase public awareness, followed by increasing the impact of DNA. Along with this, there are several burning questions, such as whether police should be granted access or not that are needed to be answered. Legal frameworks anticipating the future perspective of DNA technology and avoiding substantial threats of criminal justice will increase its impact in the CJS, and it is likely to be a significant contribution toward humanity.

References

Adhikary J (2007) DNA technology in administration of justice. LexisNexis Butterworths (a division of Reed Elsevier India Pvt Ltd), New Delhi, India

Alvarez-Cubero MJ, Saiz M, Martinez-Gonzalez LJ, Alvarez JC, Lorente JA (2018) Application of DNA fingerprinting: DNA and human trafficking. In: Dash HR et al (eds) DNA fingerprinting: advancements and future endeavors. Springer, Singapore. https://doi.org/10.1007/978-981-13-1583-1_11

Amankwaa AO (2019) Trends in forensic DNA database: transnational exchange of DNA data. Forensic Sci Res 5:8. https://doi.org/10.1080/20961790.2019.1565651

Batch FH, Albertini RJ, Joo P, Anderson JL, Bortin MM (1968) Bone marrow transplantation in a patient with the Wiskott-Aldrich syndrome. Lancet 2:1364–1366

Bolden K (2011) DNA fabrication, a wake up call: the need to reevaluate the admissibility and reliability of DNA evidence. Georgia State Univ Law Rev 227:1–34

Brown A (1998) DNA as an investigative technique. Sci Justice 38(4):263–265

Bureau of Justice Statistics (1991) Forensic DNA analysis: issues. U.S. Department of Justice, Bureau of Justice Statistics, Washington, DC, p 4. note 8

Butler JM (2006) Genetics and genomics of core short tandem repeat loci used in human identity testing. J Forensic Sci 51:253–265

Butler JM (2015a) The future of forensic DNA analysis. Philos Trans R Soc B 370:20140252. https://doi.org/10.1098/rstb.2014.0252

Butler JM (2015b) U.S. initiatives to strengthen forensic science and international standards in forensic DNA. Forensic Sci Int Genet 18:4–20. https://doi.org/10.1016/j.fsigen.2015.06.008

Charter of the Subcommittee on Forensic Science, Committee on Science National Science and Technology Council (2015). Available from: https://www.whitehouse.gov/sites/default/files/microsites/ostp/forensic-science-subcommittee-charter.pdf

Christopher H, Asplen JD (2001) The application of DNA technology. U.S. Department of Justice. https://www.ncjrs.gov/pdffiles1/nij/grants/203971.pdf

Committee on DNA Technology in Forensic Science, USA Board of Biology, Commission on Life Science, National Research Council of USA (1992) DNA technology in forensic science

Cordner S (2001) Ethics of forensic applications and databanks. In: Encyclopedia of forensic and legal medicine 2. Elsevier Ltd, Amsterdam

Dauber EM, Dorner G, Mitterbauer M, Wenda S, Fae I, Glock B, Mayr WR (2004) Discrepant results of sample taken from different tissues of a single individual. Prog Forensic Genet 10:48–49

Dodd V, Laville S (2012) Stephen Lawrence verdict: Dobson and Norris guilty of racist murder. The Guardian, 03 Jan 2012

Edwards HT (2010) The National Academy of Sciences Report on Forensic Sciences: what it means for the bench and the bar. Presentation at the Superior Court of the District of Columbia conference on the role of the courts in an age of developing science & technology, held in Washington, DC, 6 May 2010. Available from: http://www.law.yale.edu/documents/pdf/alumni_affairs/stith_edwards_nas_report_forensic_science.pdf

Edwards A, Hammond HA, Jin L, Caskey CT, Chakraborty R (1992) Genetic variation at five trimeric and tetrameric tandem repeat locus in four human population groups. Genomics 12 (2):241–253

Frumkin D, Wasserstrom A, Davidson A, Grafit A (2010) Authentication of forensic DNA samples. Forensic Sci Int Genet 4(2):95–103

García O, Crespillob M, Yurrebasoa I (2017) Suspects identification through "familial searching" in DNA databases of criminal interest. Social, ethical and scientific implications. Rev Esp Med Legal 43(1):26–34. https://doi.org/10.1016/j.reml.2016.07.003

Gatti RA, Meuwissen HJ, Allen HD, Hong R, Good RA (1968) Immunological reconstitution of sex linked lymphopenic immunological deficiency. Lancet 2:1366–1369

Goswami GK (2015) The genetic truth of surrogate parentage. Med Leg J 83(4):188–193

Goswami GK (2016) Assisted reproduction and conflicts in rights. Satyam Law International, New Delhi

Goswami GK, Goswami S (2018) Three decades of DNA evidence: judicial perspective and future challenges in India. In: Dash HR et al (eds) DNA fingerprinting: advancements and future endeavors. Springer, Singapore. https://doi.org/10.1007/978-981-13-1583-1_11

GrubbA (1993) Legal aspect of DNA profiling. JFSS 33(4):228–233

Guillén M, Lareu MV, Pestoni C et al (2000) Ethical-legal problems of DNA databases in criminal investigation. J Med Ethics 26:266–271

Gupta R, Gupta S, Gupta M (2016) Journey of DNA evidence in legal arena: An insight on its legal perspective worldwide and highlights on admissibility in India. J Forensic Sci Med 2 (2):102–106. https://doi.org/10.4103/2349-5014.184196

Harshey A, Srivastava A (2018) DNA theft, DNA fabrication: serious challenges in ethical practice of DNA. J Forensic Sci Crim Investig 8(5)

INTERPOL (2009) INTERPOL handbook on DNA data exchange and practice, 2nd edn. INTERPOL, Lyon

Jakovski Z, Ajanovska RJ, Stankov A, Popska V, Bitoljanu N, Belakaposka V (2017) The power of forensic DNA database in solving crime cases. Forensic Sci Int Genet Suppl Ser 6:e275–e276

Jayewardene CHS (1988) Forensic science and the law. Forensic Sci Int 36:41–45

Jeffreys AJ, Wilson V, Thein SL (1985) Hyper variable minisatellite regions in human DNA. Nature 314:67–73

Jobling MA, Gill P (2004) Encoded evidence: DNA in forensic analysis. Nat Rev 5(10):739–751

Joy S, Bastian TS, Selvamani M, Abraham S (2018) Indian-national DNA offender database as tool for criminal surveillance: need for public debate!!! J Med Radiol Pathol Surg 5:11–14

Kirby M (2001) DNA evidence: Proceed with care. Aust J Forensic Sci 33:9–13

Kumar S, Verma AK, Singh P, Singh R (2016) Current scenario of forensic DNA database in or outside India and their relative risk. Egypt J Forensic Sci 6:1–5

Kunkel DR (2018) A moratorium on the presentation of DNA evidence. Int Soc Sci Rev 94(3):2

Laurie G (2003) DNA theft: new crime in the UK. Nat Rev Genet 4(8):584

Law Commission of India, Department of legal affairs, Govt. of India (2017) Human DNA profiling—A draft bill for the use and regulation of DNA-based technology. http://lawcommissionofindia.nic.in/reports/Report271.pdf

Lynch M (2003) Gods signature: DNA profiling, the new gold standard in forensic science. Endeavor 27(2):93–97

Lynch M, McNally R (2003) "Science," "common sense," and DNA evidence: a legal controversy about the public understanding of science. Public Understand Sci 12:83–103

Machado H, Silva S (2019) What influences public views on forensic DNA testing in the criminal field? A scoping review of quantitative evidence. Hum Genomics 13:23. https://doi.org/10.1186/s40246-019-0207-5

Motluk A (2011) Canadian court bans anonymous sperm and egg donation. Nature. https://doi.org/10.1038/news.2011.329

Murphy E (2017) Forensic DNA typing. Annu Rev Criminol 1:13.1–13.19. https://doi.org/10.1146/annurev-criminol-032317-092127

National Academy of Sciences (2009) Strengthening forensic science in the United States: a path forward. National Academies Press, Washington, DC

National Department of Justice (NIJ) Research Report, US Department of Justice (1996) Convicted by juries, exonerated by science: case studies in the DNA evidence series

Patel N, Gautaman VK, Jangir SS (2013) The role of DNA in criminal investigation – admissibility in Indian legal system and future perspectives. Int J Humanit Soc Sci Invent 2(7):15–21

Pope S, Chapman H (2006) The effect of bone marrow transplantation on DNA profiles: a case example. Sci Justice 46(4):231–237

Renewed Charter of the Subcommittee on Forensic Science (2015). Available from: https://www.whitehouse.gov/sites/default/files/sofs_charter_2012_signed.pdf

Sedley S (2005) Science & justice—DNA and the courts. Sci Justice 45(2):59–60

Shrivastava P, Jain T, Trivedi VB (2016) DNA fingerprinting: a substantial and imperative aid to forensic investigation. Eur J Forensic Sci 3(3):1–8. https://doi.org/10.5455/ejfs.204929

Shrivastava P, Kaitholiya K, Kumawat RK, Dixit S, Dash HR, Srivastava A, Baroniya Y, Mishra A, Kushwaha K, Rana M, Dixit A, Sarawgi T, Bhatt A, Sharma H, Sinha M, Chaubey G (2019) Forensic effectiveness and genetic distribution of 23 autosomal STRs included in VerifilerPlusTM multiplex in a population sample from Madhya Pradesh, India. Int J Legal Med 134:1327. https://doi.org/10.1007/s00414-019-02172-4

Srivastava A, Kumawat R, Dixit S, Kaitholia K, Shrivastava D, Yadav VK, Nigam K, Sharma H, Trivedi VB, Chaubey G, Shrivastava P (2019a) Genetic data for Power Plex 21™ autosomal and Power Plex 23 Y-STR™ loci from population of the state of Uttar Pradesh, India. Int J Legal Med 133(5):1381–1383

Srivastava A, Yadav VK, Nigam K, Dixit S, Kumawat RK, Shrivastava D, Shrivastava P (2019b) Genomic portrait of population of Uttar Pradesh, India drawn with autosomal STRs and Y-STRs. Forensic Sci Int Genet Suppl Ser 7(1):800–801. https://doi.org/10.1016/j.fsigss.2019.10.182

Srivastava A, Singh PP, Bandopadhyay A, Singh P, Das D, Tamang R, Chaubey AK, Shrivastava P, Driem GV, Chaubey G (2020) Genetic and linguistic noncorrespondence suggests evidence for collective social climbing in the Kol tribe of South Asia. Sci Rep 10:5593. https://doi.org/10.1038/s41598-020-61941-z

Steinback R (2007) The fight for post-conviction DNA testing is not yet over: an analysis of the eight remaining holdout states and suggestions for strategies to bring vital relief to the wrongfully convicted. J Crim Law Criminol 98(1):8

Tang J, Ostrander J, Wickenheiser R, Hall A (2020) Touch DNA in forensic science: the use of laboratory created eccrine fingerprints to quantify DNA loss. Forensic Sci Int Synergy 2:1–16. https://doi.org/10.1016/j.fsisyn.2019.10.004

The DNA Technology (use and application) regulation bill (2019). Retrieved from: http://164.100.47.4/billstexts/lsbilltexts/PassedLokSabha/142C_%202018_LS_Eng.pdfhttps://pib.gov.in/Pressreleaseshare.aspx?PRID=1559099

The Royel Society & The Royel Society of Edinburgh (2017) Forensic DNA analysis: a primer for the court. ISBN: 978-1-78252-301-7. https://royalsociety.org/about-us/programmes/science-and-law

Verma SK, Goswami GK (2014) DNA evidence: current perspective and future challenges in India. Foensic Sci Int 241:183–189

Virkler K, Lendev IK (2009) Analysis of body fluids for forensic purposes: from laboratory testing to non-destructive rapid confirmatory identification at a crime scene. Forensic Sci Int 188:1–17

Wallace HM, Jackson AR, Gruber J, Thibedeau AD (2014) Forensic DNA databases ethical and legal standards: a global review. Egypt J Forensic Sci 4:57–63

Walsh SJ (2005) Legal perceptions of forensic DNA profiling: Part I. A review of the legal literature. Forensic Sci Int 155:51–60

Wrobel K, Doroz R, Porwik P, Bernas M (2018) Personal identification utilizing lip print furrow base pattern. A new approach. Pattern Recognit 81:585–800. https://doi.org/10.1016/j.patcog.2018.04.030

Yadav PK (2017) Ethical issues across the different field of forensic science. Egypt J Forensic Sci 7 (1):10

32 DNA Databases

Anna Barbaro

Abstract

Forensic DNA analysis of evidence is gold standard in criminal investigation. Since with the advent of forensic DNA technology, the importance of DNA database is well understood for the purpose of statistical interpretation of results and also for making a criminal or offender database. This chapter describes the information on various DNA databases.

Keywords

Database · DNA · YHRD · X-STR · STRbase · EMPOP

32.1 Introduction

The introduction in forensics of DNA typing as an identification tool produced an enormous change in criminal investigations. During years, several DNA databases that include relevant information on commonly used forensic markers, available technologies, and population data have been developed as useful tools for the forensic community. In addition, many countries introduced national databases where DNA profiles of evidences and of people involved in violent crimes are stored, with the aim to link crimes and to favorite the exchange of data between countries.

A. Barbaro (✉)
Department of Forensic Genetics, Studio Indagini Mediche E Forensi (SIMEF), Reggio Calabria, Italy

Worldwide Association of Women Forensic Experts (WAWFE), Reggio Calabria, Italy
e-mail: simef_dna@tiscali.it

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_32

32.2 Online Forensic Databases

Evaluation of autosomal STR profiles and of sexual chromosomes/mitochondrial haplotypes requires an accurate knowledge of several statistical parameters. Numerous papers about STR population studies and probabilistic calculations have been published, and they highlight also the importance of internal and external quality control (QC) during dataset creation, in order to prevent errors and to ensure more reliable allele frequency estimate.

For example, right sample collection procedures are required to exclude close relatives in order to obtain correct and high quality data. In this perspective, during years, different platforms have been created, which are free available online, which offer to the international community useful information about commonly used forensic markers, available technologies, validation studies, and also several calculation tools.

32.2.1 STR Database (STRbase)

Since years, STR analysis is routinely used for human identity testing: new loci and new systems are continuously studied and validated by forensic scientists. In July 1997, the Applied Genetics Group of the National Institute of Standards and Technology created the STRBase (available at www.cstl.nist.gov/biotech/strbase) that became an important resource for the forensic community because it contains updated information about commonly used STRs (Butler and Reeder 1997; Ruitberg et al. 2001; Butler 2008) (Fig. 32.1).

It also includes several sections with information about commercial multiplex, PCR primers sequences, observed alleles or “off-ladder” alleles, new microvariant, and mutation rates. In addition, data about available typing technologies, population studies, quality assurance, and validation procedures are provided. A list of STR

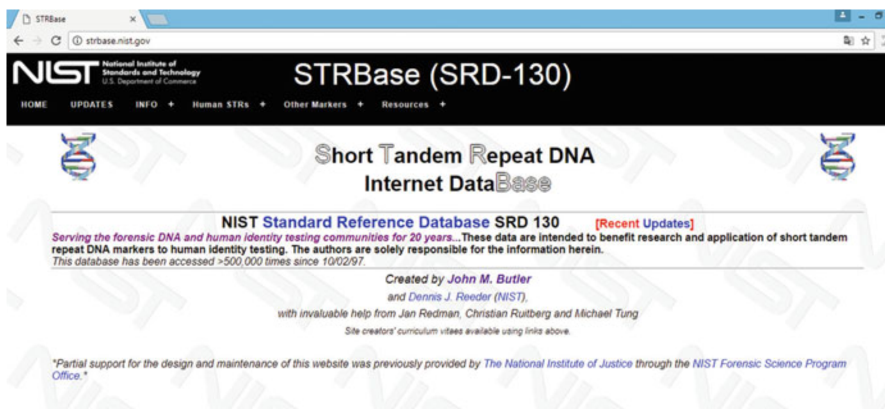


Fig. 32.1 STR database (STRbase) home page

references with relevant web links and addresses of forensic experts is also reported. Additional sections include information about non-traditional DNA markers (SNPs and mt-DNA) and non-human DNA resources.

32.2.2 Y-STR Database (YHRD)

STR markers on chromosome Y (ChrY) are inherited as a block of linked haplotypes; hence, an estimate of a haplotype frequency is based on the observation of how many times a particular haplotype is observed in a given database. Because of this, a Y-STR database consists of haplotype frequencies rather than allele frequencies (Willuweit and Roewer 2007).

Obviously the possibility to obtain a reliable estimate and the confidence in the accuracy of frequencies estimation depend on some database characteristics, such as the ethno-geographic composition of the database, the number of individuals included, and the number of searchable Y-STR loci. In 2000, a *Y-STR haplotype reference database (YHRD)*, free accessible online at www.yhrd.org/index.html, was created by Lutz Roewe and Sascha Willuweit from the Institute of Legal Medicine and Forensic Sciences, Charité Universitätsmedizin Berlin with some objectives Roewer et al. (2001) (Fig. 32.2), such as:

- To collect reliable frequencies for Y-STR haplotypes and Y-SNP haplotypes, which can be used in forensic and kinship analysis
- To study the characteristics of male lineages in order to obtain informations about the evolution of human populations

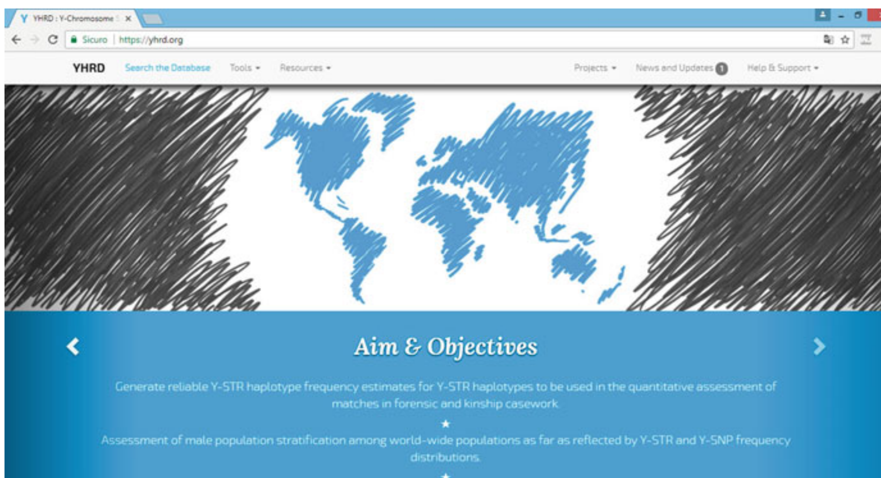


Fig. 32.2 Y-STR database (YHRD) home page

The database represents now the result of international collaborative efforts to collect population data: by October 2017, Y-HRD contains more than 197,100 of 9-loci minimal haplotypes, 39,414 23-locus haplotypes, and 22,832 Y-SNP profiles, from Asia, Europe, Latin America, North America, Africa, and Oceania/Australia. In addition, the database provides advanced tools for the calculation of likelihood ratio (LR) of patrilineal relationships and for Y-STR mixture calculations.

32.2.3 X-STR Database (ChrX-STR.org 2.0)

The analysis of X-chromosome (ChrX) markers may support autosomal and Y-STR typing especially in complex kinship cases. While ChrY markers do not recombine during meiosis, for ChrX markers, the localization is relevant because all loci are placed on the same chromosome within an area of 240 cM. Physical and genetic localization of chromosome X loci are related, but not linearly; this implies that working with X-STRs requires an exact knowledge about their genetic localization because the possibility of a crossing-over is directly related to the genetic distance between markers. Hence, loci are often linked and exhibit linkage disequilibrium. Thus, X-STR markers are clustered into linkage groups, and each set may be considered as a haplotype (Hering et al. 2006; Szibor 2007; Szibor et al. 2003).

Haplotype frequencies are estimated by the number of observations in a population sample: if two or more STRs loci are analyzed, the count of haplotypes may include up to thousand possible haplotypes. A *X-chromosome 2.0 database* (ChrX-STR.org) freely accessible online (http://www.chrx-str.org/) was launched in 2006 by the Forensic ChrX Research Group in order to provide a reference database for ChrX STR haplotypes (Fig. 32.3). It includes published population data from several countries and information about the most commonly used X-STRs multiplexes and commercial software packages (Szibor et al. 2006).

Fig. 32.3 X-STR 2.0 database (ChrX-STR.org) home page

In addition, a free calculation software is accessible for the calculation of some population genetics parameters useful in forensic such as MEC, PIC, HET, PDfemale, and PDmale and for Hardy–Weinberg Equilibrium checking. An ideogram about markers distribution on the chromosome, and a table with information about physical and genetical localization of several markers are also available.

32.2.4 Mitochondrial DNA Database (EMPOP)

Since years, mitochondrial DNA (mt-DNA) is used to study several genetic diseases, to map out ancestry and migration during years, but also to analyze variations between human and other species in order to find any relationships between them. Mt-DNA is also useful, in forensics, for the analysis of scarce or high degraded samples, because it is available in several copies per cell, while nuclear DNA has only two copies per cell. In addition, while nuclear DNA is inherited from father and mother to the offspring, mitochondrial DNA is passed through maternal lineage, so it can be used when a direct comparison with a person is not possible because any other individual from the maternal line may give a reference sample. Obviously, this means that the analysis of mt-DNA is not enough for identification purposes.

A common problem with mt-DNA is the heteroplasmy that consists in the co-existence in an individual of different mitochondrial DNA types, generally differing at only one base (Butler 2009). This must be accurately evaluated during data interpretation in order to avoid wrong exclusions. Many collaborative studies showed that several errors may occur during mt-DNA analysis or data interpretation (i.e., nomenclature inconsistencies, sample mix-up), so the European DNA Profiling Group (EDNAP) on October 16, 2006, launched online a EDNAP forensic mt-DNA population database (EMPOP) with the aim to develop the standards for data generation, analysis, and quality control (Parson and Bandelt 2007; Parson et al. 2014; Parson and Dür 2007).

The database is maintained and updated by the Institute of Legal Medicine of Innsbruck (Austria), and it is freely accessible online at www.empop.org (Parson and Dür 2007; Prieto et al. 2011; Roewer and Parson 2015) (Fig. 32.4). Mitochondrial DNA databases are the basis for frequency estimations of mt-DNA haplotypes, and hence, they are useful for calculating the probability of a match by determining the frequency of a given haplotype in a population. The EMPOP database now comprises 5173 mt-DNA sequences from worldwide populations sent by laboratories participating in some collaborative exercises: among them, 4527 sequences are forensic data (high quality sequences), and 646 sequences are from published papers. The majority of the sequences in the database are from Western Eurasian populations, less data are from East Asian, South East Asian, and Sub-Sahara African populations, but data are continuously increasing. The database includes also a section that provides some software packages for mt-DNA analysis, data interpretation, and quality control.

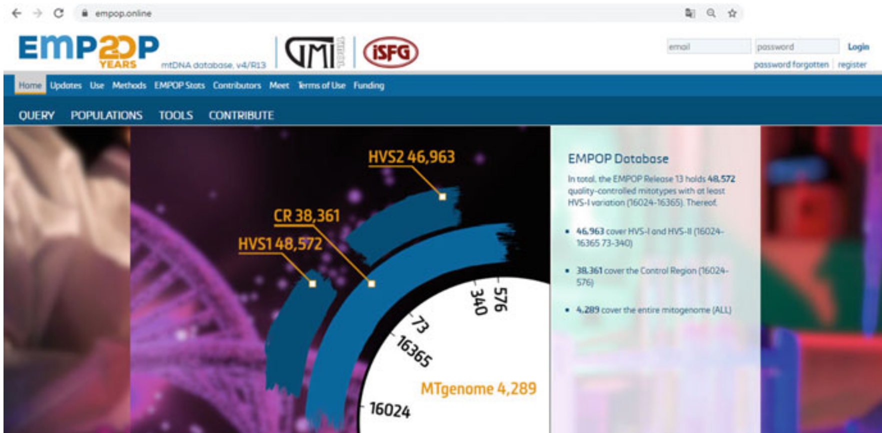


Fig. 32.4 Mitochondrial DNA database (EMPOP) home page

32.2.5 SNPs Database (SNP*for*ID browser)

Single-nucleotide polymorphisms (SNPs) represent approximately 90% of the natural genetic variation in the human genome. An SNP is a single base change in a DNA sequence and occurs when a base (A, T, C, or G) is replaced by any of the other three bases. Even if STRs are considered the ideal markers for personal identification and paternity tests, however SNPs can be also useful due to their abundance in the genome, low mutation rates, reduced amplicon sizes (ability to analyze degraded DNA), high-throughput genotyping, ancestry, and phenotypic trait prediction (Gill et al. 2004).

In fact an advantage of SNPs analysis toward STRs typing is ability to recover high information from degraded DNA samples because with SNPs only a single nucleotide is typed instead of hundreds of nucleotides as with STRs (Sobrino and Carracedo 2005; Butler et al. 2007; Budowle and van Daal 2008). In April 1999, a SNP Consortium (TSC) was established under the leading of Arthur L. Holden as a collaboration of ten large pharmaceutical companies and the U.K. Wellcome Trust philanthropy. The initial purpose was to discover in 2 years 300,000 SNPs and to produce a public widely accepted, high-quality SNPs map resource, but final results largely exceeded the aim and over 1.8 million SNPs were mapped.

A public website (http://snp.cshl.org), maintained at Cold Spring Harbor Laboratory, was set up in order to provide to the scientific community some information about the project itself and also to improve existing data and searching tools (Sherry et al. 2001). The SNP*for*ID Consortium was established in 2003 as collaboration between different institutes of legal medicine with the aim to identify SNPs suitable for individual identification and to develop highly efficient strategy and high-throughput platforms for reliable and accurate SNPs multiplex typing. The group in fact developed in 2005 a multiplex of 52 SNPs useful for forensic analysis (Sanchez et al. 2006).

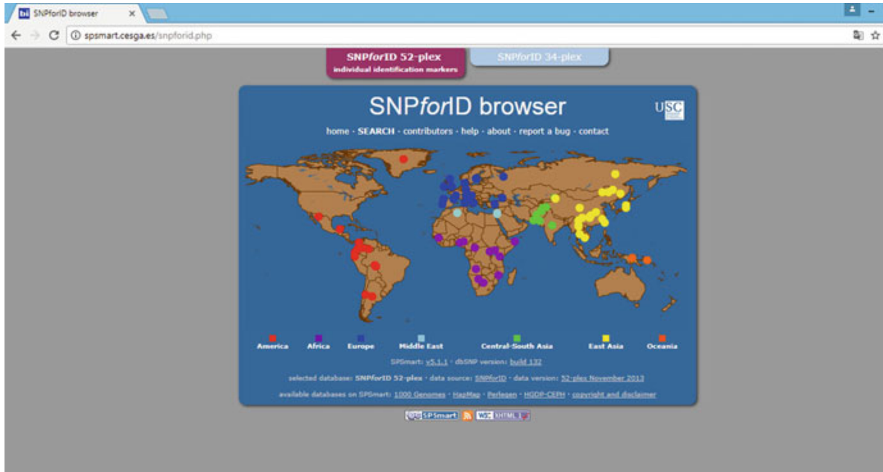


Fig. 32.5 SNPforID browser home page

All data generated by the SNPforID consortium are freely available online on the SNPforID browser (http://spsmart.cesga.es/snpforid.php) that is a public accessible database launched in 2005 and maintained by the Institute of Legal Medicine from the University of Santiago de Compostela (Spain) (Amigo et al. 2008) (Fig. 32.5). The database allows to search and review SNP allele frequencies from several populations, and in addition, it permits the combination of populations in groupings, the comparison between populations individually or among groupings or with equivalent HapMap data.

32.3 Criminal Databases

During years national DNA databases have become very efficient tools during criminal investigations for providing information about unknown individuals with the aim to identify the author of a crime. Originally developed for sexual assault caseworks, DNA databases have been extended to all criminals, because it was observed a majority of crimes are committed by repeat offenders.

The utility of a DNA database consists of the possibility to link crimes or criminals who are not directly related to a specific case and to prevent further crimes committed by the same persons. The first governmental database was established in the United Kingdom in April 1995, the second one in the same year in New Zealand, and the third one in France in 1998 (Schneider and Martin 2001, Linacre 2003). In USA, the DNA Identification Act of 1994 authorized the creation of the USA National DNA database (CODIS) that was launched in 1997 by FBI. CODIS software allows comparisons of DNA profiles at different levels (national, statal and local) in accordance with actual laws.

For the inclusion in CODIS, DNA profiles must be analyzed following the FBI Director's Quality Assurance Standards and the guidelines of the Scientific Working Group on DNA Analysis Methods (SWGDAM). In Europe, the DNA working group of the European Network of Forensic Sciences Institutes (ENFSI) has developed recommendations for DNA databases management, describing all criteria required for including or deleting DNA profiles (ENFSI 2009/2010). Many countries have implemented their DNA analysis programs for database generation, according to guidelines of coordinating bodies or DNA working groups.

Generally, a national database includes a section with DNA profiles of several categories of persons such as convicted persons, suspects, arrestees, and sometimes also volunteers. In another section of the database, DNA profiles of the evidences recovered at the crime scene are stored. Obviously, this may change in different countries according to their law. Database access is obviously exclusive for police purposes, and there are strong penalties for anyone different from the law enforcement, using information or samples for other purposes. To assure respect of privacy and security of the information, the database does not contain any information about samples donor or about the criminal case.

The software is able to perform multiples comparisons so as to determine if an evidence profile matches with any known donor. In case of a match, a new DNA sample from the suspect is required for performing a confirmation test. After a match confirmation, evidence test laboratory is notified and may exchange additional information and case details with the central laboratory. If there is not any match between the evidence profile and the reference profiles, then the evidence profile is compared with DNA profiles of samples coming from other unsolved crimes, to verify if the same person (even if unknown) is the author of different crimes.

In USA, 13 loci (D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, CSF1PO, FGA, TH01, TPOX, and vWA) were initially required for inclusion of a reference profile in CODIS. In 1999, the ENFSI (European Network of Forensic Science Institutes) DNA Working proposed a European Standard Set (ESS) of seven loci (TH01, vWA, FGA, D21S11, D3S1358, D8S1179, and D18S51) as a minimum number of markers required to allow comparison of DNA profiles between different European countries. This proposal was accepted and confirmed by the EU-Council resolution no. 2001/C 187/01.

In 2009, following to the Treaty of Prüm, that enabled all 27 EU countries to exchange DNA profiles, ENFSI recommended the extension of the European Standard Set of Loci (ESS) including five additional loci, with the aim to reduce false matches in case of massive data exchange. This proposal was accepted and confirmed by the resolution of the European Council no. 2009/C 296/01. In order to favorite the amplification of degraded/low DNA samples, it was established to include between the five new core loci, also three new miniSTR loci (D10S1248, D2S441, D22S1045).

In 2012, in USA, FBI also proposed the expansion of the original CODIS core loci from 13 to 20 loci with the aim to increase international compatibility and the discrimination power. In particular, in 2015 it was decided to include the five ESS loci (D1S1656, D2S441, D10S1248, D12S391, and D22S1045) plus D2S1338 and

D19S443. This resolution became effective by January 2017. As of September 2017, CODIS contains over 13,041,408 offender profiles, 2,860,423 arrestee profiles, 804,902 forensic profiles, and produced over 392,684 matches assisting in more than 377,507 criminal investigations (www.fbi.gov). It is now the largest DNA database in the world.

References

Amigo J, Phillips C, Lareu M, Carracedo A (2008) The SNPforID browser: an online tool for query and display of frequency data from the SNPforID project. Int J Legal Med 122(5):435–440

Budowle B, van Daal A (2008) Forensically relevant SNP classes. BioTechniques 44:603–610

Butler JM, Reeder DJ (1997) STRBase: a short tandem repeat DNA database for the DNA typing community. Profiles in DNA, Promega Corporation, Volume 1, No. 2, p. 10

Butler JM (2008) New resources for the forensic genetics community available on the NIST STRBase website. Forensic Science International: Genetics Supplement Series (Progress in Forensic Genetics 12) 1(1):97–99

Butler JM (2009) Fundamentals of Forensic DNA Typing. Academic Press, Elsevier, London

Butler JM, Coble MD, Vallone PM (2007) STRs vs. SNPs: thoughts on the future of forensic DNA testing. Forensic Sci Med Pathol 3:200–205

DNA-Database Management Review and Recommendation ENFSI DNA Working Group—April (2009/2010)

Gill P, Werrett DJ, Budowle B, Guerrieri R (2004) An assessment of whether SNPs will replace STRs in national DNA database: joint considerations of the DNA working group of the European Network of Forensic Science Institutes (ENFSI) and the Scientific Working Group on DNA Analysis Methods (SWGDAM). Sci Justice 44:51–53

Hering S, Augustin C, Edelmann J, Heidel M, Dressler J, Rodig H, Kuhlisch E, Szibor R (2006) DXS10079, DXS10074 and DXS 10075 are STRs located within a 280-kb region of Xq12 and provide stable haplotypes useful for complex kinship cases. Int J Legal Med 120:337–345

Linacre A (2003) The UK national DNA database. Lancet 361:1842

Parson W, Bandelt HJ (2007) Extended guidelines for mtDNA typing of population data in forensic science. Forensic Sci Int Genet 1(1):13–19

Parson W, Gusmao L, Hares DR, Irwin JA, Mayr WR, Morling N, Pokorak E, Prinz M, Salas A, Schneider PM, Parsons TJ (2014) DNA Commission of the International Society for Forensic Genetics: revised and extended guidelines for mitochondrial DNA typing. Forensic Sci Int Genet 1(1):13–19

Parson W, Dür A (2007) EMPOP-a forensic mtDNA database. Forensic Sci Int Genet 1(2):88–92

Prieto L, Zimmermann B, Goios A, Rodriguez-Monge A, Paneto GG, Alves C, Alonso A, Fridman C, Cardoso S, Lima G, Anjos MJ, Whittle MR, Montesino M, Cicarelli RM, Rocha AM, Albarran C, de Pancorbo MM, Pinheiro MF, Carvalho M, Sumita DR, Parson W (2011) The GHEP-EMPOP collaboration on mtDNA population data—a new resource for forensic casework. Forensic Sci Int Genet 5(2):146–151

Roewer L, Parson W (2015) Internet accessible population databases: YHRD and EMPOP. Forensic Biol 311

Roewer L, Krawczak M, Willuweit S, Nagy M, Alves C, Amorim A, Anslinger K, Augustin C, Betz A, Bosch E, Caglia A, Carracedo KM et al (2001) Online reference database of European Y-chromosomal short tandem repeat (STR) haplotypes. Forensic Sci Int 118:106–113

Ruitberg CM, Reeder DJ, Butler JM (2001) STRBase: a short tandem repeat DNA database for the human identity testing community. Nucleic Acids Res 29(1):320–322

Sanchez JJ, Phillips C, Børsting C, Balogh K, Bogus M, Fondevila M, Harrison CD, Musgrave-Brown E, Salas A, Syndercombe-Court D, Schneider PM, Carracedo A, Morling N (2006) A

multiplex assay with 52 single nucleotide polymorphisms for human identification. Electrophoresis 27(9):1713–1724

Schneider PM, Martin PD (2001) Criminal DNA database: the Europe situation. Forensic Sci Int 119(2):232–238

Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K (2001) dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 29:308–331

Sobrino B, Carracedo A (2005) SNP typing in forensic genetics. In: Carracedo A (ed) Forensic DNA typing protocols, Methods in molecular biology, vol 297. Humana Press, Totowa, NJ, pp 1064–3745

Szibor R (2007) X-chromosomal markers: past, present and future. Forensic Sci Int Genet 1 (2):93–99

Szibor R, Krawczak M, Hering S, Edelmann J, Kuhlisch E, Krause D (2003) Use of X-linked markers for forensic purposes. Int J Legal Med 117:67–74

Szibor R, Hering S, Edelmann J (2006) A new web site compiling forensic chromosome X research is now online. Int J Legal Med 120:252

Willuweit S, Roewer L (2007) Y chromosome haplotype reference database (YHRD): update. Forensic Sci Int Genet 1(2):83–87

33 Building of the World's Largest DNA Database: The China Case

Ausma Bernotaite

Abstract

Started only in 2005, China has already entered 68 million profiles into its National DNA Database (NDNAD), according to the data presented by NDNAD governing agency—the Chinese Institute of Forensic Science, Ministry of Public Security, at the Asian Forensic Sciences Network in September 2018, Beijing, China. Following the data presented by the same government agency at the International Society for Forensic Genetics in 2017, Seoul, Korea, the database grew to the said number from a reported 55 million profiles in less than 1 year. This number implies that by the number of profiles entered, the Chinese NDNAD is by far the largest in the world, followed by 17,530,781 profiles in the National DNA Index of the United States of America in 2018 and 6,024,032 profiles in the NDNAD of the United Kingdom in 2017. Additionally, National Missing Children DNA Database was created in 2009 to include genetic data from parents and their children. While large in the mere count of profiles included, it currently covers a relatively low percentage of approximately 4.5% of the Chinese population and has not provided globally comparable match, or hit, rates for criminal case inquiries. Despite its rapid expansion and swift adoption of emerging genetic technologies, little effort has been made to share the developmental details of the database. This chapter serves as a short summary to trace the development timeline, goals, technological applications, main actors, as well as the biggest achievements and future plans of the Chinese NDNAD. It is concluded with the emphasis for the need to ground national databases in strong socio-legal considerations.

A. Bernotaite (✉)
School of Criminology and Criminal Justice, Griffith University, Mount Gravatt, QLD, Australia
e-mail: ausma.b@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_33

Keywords

DNA database · China · National DNA Index · NDNAD

33.1 Introduction

The DNA database has become the most important technological tool in the field of criminal technology in China. (Zhao et al. 2017)

Analyzing DNA evidence with the assistance of DNA databases is rapidly becoming a part of routine criminal police workflow globally. Started only in 2005, Mainland China[1] has already entered 68 million profiles into its National DNA Database (further referred to as the China's NDNAD), according to the data presented by China's NDNAD governing agency—the Chinese Institute of Forensic Science (IFS), Ministry of Public Security, at the Asian Forensic Sciences Network in September 2018, Beijing, the People's Republic of China. This makes the China's NDNAD the largest DNA database globally by the number of profiles entered, with the USA and the UK ranking second and third, respectively. The largest growing portion of the database is the reference profile portion, which constitutes over 95% of all profiles preserved. Apart from the NDNAD, China also has a Y-Chromosome STR Haplotype Reference Database (YHRD), as well as National Missing Children DNA Database (NMCDD) founded in 2009. The IFS responsible for governing the NDNAD, the NMCDD, as well as more around 600 government DNA laboratories of provincial, prefectural, and county levels. The complex forensic database system employs a top to bottom governance mode across a growing network of forensic laboratories.

Despite the rapid growth of the Chinese DNA databases, little is known about the governing principles of their implementation and management internationally. Nationally, there is also little debate in the country due to the authoritarian nature of its governance, where important decisions are made by the Party-state and then trickled down to provincial levels for implementation. Discussing the governance of China's genetic databases provides an interesting case study: the NDNAD is operates in a country without free speech and run by one Party, which is the sole actor taking and implementing decisions for public security, and where public debate is not encouraged. The government's ability to make quick decisions without reaching public consensus has been the most important factor determining the speedy adoption of new technologies within the criminal justice systems.

To shed light on how the intersection of big data, genetics, and criminal justice works in China, this chapter will provide a basic comprehensive overview of the history of DNA databasing internationally and in Mainland China, types of databases created, technologies used, standards followed, their relevant governing

[1]In this chapter, we refer to "China" as the region of Mainland China without the areas of Hong Kong or Macao, as these areas have independent criminal justice governance mechanisms.

agencies, and outline of the main points of contestation. Conference outcomes, articles published by government representatives, and texts reflecting the views of government representatives are used where possible to give an overview of the blueprint that the Chinese government has followed in the adoption of the NDNAD. As China's forensic DNA governance is centralized and mainly hidden from the public eye, texts released by the governing agencies are the most direct reflection of what work has been completed and what developmental plans are remaining. The chapter is concluded with the overall evaluation of the forensic database governance system in China.

33.2 The International Context

Technology and big data are becoming increasingly intertwined with our everyday lives, and this also applies to criminal justice. Digitalized forms of fingerprints, DNA, facial scans, and criminal records—all of these examples of digitalized personal data used for policing—are increasingly becoming the standard for efficient implementation of criminal justice. Following the dramatic drop in cost of the genetic analysis technologies, forensic DNA databases have boomed in popularity. A well-managed DNA fingerprinting database is a powerful and cost-effective tool for untangling crime cases and has been proven to deter crime and reduce crime rates (Doleac 2017). The most recent global report on the use of DNA profiling published by Interpol in 2008 reported that 54 countries have national DNA databases and 120 countries use DNA profiling as evidence in criminal investigations (Interpol Global DNA Profiling Survey 2008). There have been many more databases created, but there has not been a more recent global summary publication since then. A strong effort to collect comprehensive information, not only about the presence and size of DNA databases, but also on the local public discussions around them, has been made by the Forensic Genetics Policy Initiative—a collaboration of three charity organizations, namely GeneWatch UK, Privacy International, and the Council for Responsible Genetics. The organization estimates that there currently are 60 countries with operational and at least 34 with planned DNA databases globally (The Forensic Genetics Policy Initiative 2018). The same source reports that there have been proposals of whole-population database creation. Tracy and Morgan (1999) called this type of DNA database expansion plans "DNA fever" and reasoned that whole-population databases are a financial burden on the government that does not provide statistically valuable outcomes. Instead it creates a Big Brother state, while potentially violating constitutional rights to personal privacy. Following the growth of forensic database adoption around the world, the scope of public and academic discussions on this multifaceted topic has also grown.

Just like any complex method of policing, DNA databases require to find the right balance between ensuring the implementation of justice and individual privacy protection. While the safety and security of the people in the country is the obvious goal of all forensic DNA databases, governing authorities must take into account a variety of intersecting factors of database operation ethics. To explain this in depth,

we can revert to the history of the two other DNA databases that are currently second and third in size after China's NDNAD. As mentioned in the introduction, the USA and the UK have the two largest databases in the Americas and Europe, respectively, not only by their size but also by the percentage of the population covered. According to the Federal Bureau of Investigation website, the USA, the National DNA Index had a total of 17,785,074 profiles in October 2018, which is approximately 5% of the total population (CODIS—NDIS Statistics 2018). The National DNA Database of the United Kingdom held an estimate of 5,258,600 DNA profile records in March 2017, the number accounting for an approximate 8% population coverage; the match rate, or "the chance that a crime scene profile, once loaded onto NDNAD, matched against a subject profile stored on NDNAD" (National DNA 2018), of the database is high at 66% for the reporting year of 2016/2017 (National DNA 2018). The National DNA Database in the UK is also the oldest one, and one that has undergone detailed scrutiny by the public. The UK's database had allowed for overly loose profile inclusion regulations, which are now partially addressed by the state: the Criminal Justice and Police Act 2001 decreased the levels of authority requirements for taking of DNA samples without consent and eliminated the need to remove the DNA data and samples collected from the database after the end of investigation (McCartney 2004). There has been a strong public backlash against the new regulations, which culminated in the case of *S. and Marper v. the UK* in the European Court—the first of its kind. In 2008, the Grand Chamber of the European Court of Human Rights passed a judgment rule that the law in England and Wales breached the European Convention on Human Rights. Following the judgment, the UK removed a total of 7,753,000 DNA profiles (including over 1.7 million DNA profiles taken from innocent people and from children) from its database without an observance of reduced crime detection rates (NDNAD 2013). The match rate of the NDNAD has seen stable increase despite the removal of a large number of profiles (National DNA 2018). Following the UK NDNAD's privacy dispute, "a growing global consensus on the need for legislative provisions for the destruction of biological samples and deletion of innocent people's DNA profiles" was gaining speed (Wallace et al. 2014). A second question brought to the public eye in the recent years is the issue of racialized inequalities of DNA data inclusion and consecutive criminalization disproportionately increasing the policing of black men observed but not resolved (Skinner 2013). Randerson estimated that a staggering number of 77% of black men between the ages of 15 and 34 years had a profile on the DNA database, and up to 37% of the total DNA database constituted profiles of black men as compared to 13% of Asian and 9% of white men (Randerson 2006). Skinner (2013) raised concerns over potential evidence of racism in DNA profile inclusion resulting in ethnic monitoring by the police.

In the USA, DNA evidence became increasingly used in the 1980s, but it was not until 1994 that the Federal Bureau of Investigations formalized the Combined DNA Index System (CODIS) database, which serves as a joint platform for DNA profile sharing between states (Deray 2011). In terms of racial inequalities, the USA database has witnessed criticisms of similar nature as to those of the UK NDNAD: Cole (1999) traced archival knowledge of black policing in the USA and expressed

concern over the racial discrimination in the database reproduced by historical racist biases. A further criticism of the CODIS pointed at the expansive policies of databasing and referred back to the *Marper* case against the United Kingdom and the consecutive ruling of the European Court of Human Rights:

> *As federal and state legislation in the United States become increasingly expansive, legislators and law enforcement need to be reminded that just because they have the technology to do something does not mean they should do it.* (Cole 1999)

Despite the public criticisms, the two databases have proven success in assisting the solving of crimes. In the annual UK NDNAD report (2018), the National DNA Database Strategy Board provided the following statistics: "In 2016/17 the NDNAD provided 31,743 routine matches, including to 514 homicides and 612 rapes, and 493 urgent matches, including to 122 homicides and 141 rapes. This is an increase in the number of matches reported in the previous year, demonstrating the continued effectiveness of the NDNAD. The percentage of crime-scene profiles which matched a subject profile on load to the NDNAD (referred to as the match rate) was 66%" (National DNA 2018). In the USA, the official FBI website provides wider statistics: "As of October 2018, CODIS has produced over 440,346 hits assisting in more than 428,808 investigations" (CODIS—NDIS Statistics 2018). The joint historical and modern experiences of the two largest DNA databases, following the Chinese one, suggest that DNA evidence can serve as an important tool in procedural justice by enabling the police to bring serious offenders to justice faster as well as ability to prove suspects innocent. While acknowledging the benefits of DNA databases, it is also important to recognize the fact that the use of technology does not happen in a socio-political vacuum. In light of multifaceted considerations of DNA database ethics, it is of crucial importance that ongoing and transparent efforts are made to ensure that the genetic databases are managed ethically. To borrow the words from the Forensic Genetics Policy Initiative mission statement: "An appropriate middle ground between the legitimate needs of law enforcement and a respect for individual rights is achievable" (The Forensic Genetics Policy Initiative 2018).

33.3 Breaking Ground for the Chinese DNA Database

The Chinese DNA databases have grown immensely with the advancement and scaling of genetic technologies. In the recent years, we have seen more publications talking about the history of this—now grandiose—public security project. The process of setting up the database has been top to bottom: led by a small group of experts from the Ministry of Public Security and then passed down to provincial public security branches. The initial small DNA database was created to protect children from going missing—still a serious issue in modern-day China. Ge, Liu, and Peng (2017) suggest that there have been four main stages of DNA database development: technological observation, active search of suitable technical DNA database implementation strategies, promoting its creation, and the stage of rapid

development. The authors shared a few landmark events on the road to NDNAD establishment: the idea to create a "DNA Fingerprinting Database" was first introduced in a meeting co-organized by the Chinese Ministry of Public Security and the Ministry of Science and Technology, the second time at "The First Academic Conference on Forensic Material Evidence," and was finally decidedly approved for government use for the Ninth Five-Year Plan, for the years 1996–2000. An IFS publication echoes the dates and adds commentary on what was exactly achieved in each year:

> *With reference to the advanced experience of developed countries, the Physical Evidence Identification Center, the Ministry of Public Security, first proposed the establishment of an experimental DNA database in China in 1996. 'Forensic DNA Quality Control Technology and DNA Database' was a complex project that combined quality control and database technologies. It set the ground for the standardization of forensic DNA testing through systematic and pioneering research. In terms of quality control, the project completed a standard named 'Forensic DNA Laboratory Quality Assurance System' as well as 6 standards for forensic DNA testing and DNA database*. (Hu et al. 2003)

Following the setting of basic standards, various provincial laboratories started initial work with small-scale databases of their own: for example, the North Eastern city of Liaoyang built a small database of 2211 Han Chinese people with 14 preliminary gene targets in 1999 (Wang et al. 2003). The initial goal of the DNA database was to solve child kidnapping cases, and in the year 2000, five government laboratories securely connected their database DNA data of 48,000 profiles enabling the solving of 418 missing children cases. In the year 2003, the number of profiles in the shared databases was 65,012: 11,685 of those were criminal profiles, 3737 profiles retrieved from crime scenes, 37,600 profiles of missing people, and 11,990 profiles of family members of the aforementioned missing persons (Hu et al. 2003). The same source points out that after being set up for over 2 years, the database was helping solve serious crimes, such as murders, attempted murders, and rapes. Up until the year 2005, preparatory work in different cities continued, followed by an ordinance from the Ministry of Public Security for database construction issued in 2005. Thereafter, the work continued nationally: new laboratories were established, database capacity expanded, and work regulations set.

33.4 Types of DNA Databases in China

Over a seemingly short period of time, China has adopted a wide range of social surveillance technology measures. While it is the current owner of the largest DNA database in the world, the personal data collected in China is not limited to that. Currently available databases include: facial recognition, fingerprint, STR (direct and familial searches), and Y-STR profile (direct or mass screening) databases. The databases can then be connected together for various purposes of policing and public

Picture 33.1 A police station at a crossroads at an Eastern city of Ningbo using a facial recognition and traffic control robot connected to the digital database and an LED screen displaying the faces and surnames of people crossing the street on red light

security. For example, a connection of facial recognition connected to personal records database used for catching and shaming jaywalkers is shown in Picture 33.1.

There are three different DNA databases that are being run in China: the National DNA Database (中国法庭科学*DNA*数据库), the National Missing Children DNA Database (NMCDD) (全国失踪儿童*DNA*数据库), and the most recently added Y-STR database(*Y-STR*数据库) adjoined with the NDNAD. By the number of profiles, both the NDNAD and the Y-STR database are the biggest databases of their own kind in the world. All of the databases are run by the Ministry of Public Security of the People's Republic of China.

During the fifth National Congress of Forensic DNA Database in 2017, a presentation on the overview of forensic DNA technologies provided useful insights into the growth of forensic DNA use in the country (Institute of Forensic Science 2017). As of September 2017, there were 598 forensic DNA laboratories with 3344 technical staff and 1050 DNA analysis instruments. Data of crime cases solved with the help of the database for the years 2011–2016 was presented (Institute of Forensic Science 2017), as seen in Fig. 33.1:

The overall percentile numbers of profiles entered in different provinces of China vary greatly and depend on provincial public security policies, as shown in Table 33.1.

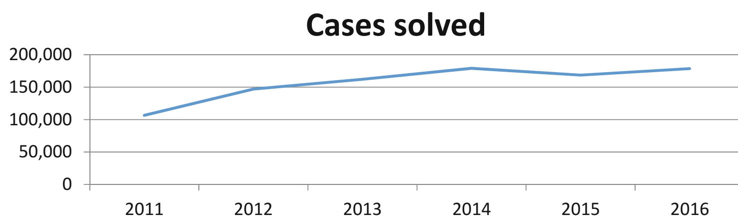


Fig. 33.1 Cases solved with the help of the NDNAD

Table 33.1 Percentile population coverage in different provinces and regions of China

Province/ region	Number of profiles	Province/region coverage percent	Province/ region	Number of profiles	Province/region coverage percent
Beijing	217,390	1.11%	Hubei	1,946,219	3.40%
Tianjin	213,068	1.65%	Hunan	2,181,102	3.32%
Hebei	1,352,379	1.88%	Guangdong	1,707,206	1.64%
Shanxi	919,004	2.57%	Guangxi	1,593,857	3.46%
Inner Mongolia	354,541	1.44%	Hainan	199,495	2.30%
Liaoning	930,050	2.13%	Chongqing	1,504,827	5.22%
Jilin	1,003,642	3.65%	Sichuan	2,746,677	3.42%
Heilongjiang	1,855,512	4.84%	Guizhou	1,553,093	4.47%
Shanghai	58,968	0.26%	Yunnan	879,757	1.91%
Jiangsu	2,997,317	3.81%	Tibet	2625	0.09%
Zhejiang	1,236,539	2.27%	Shaanxi	1,332,932	3.57%
Anhui	2,096,514	3.52%	Gansu	981,509	3.84%
Fujian	1,119,043	3.03%	Qinghai	224,768	3.99%
Jiangxi	1,596,364	3.58%	Ningxia	129,897	2.06%
Shandong	2,392,837	2.50%	Xinjiang	2,851,289	13.07%
Henan	4,232,469	4.50%			

The total number of profiles in the database comes to 3.1% of population coverage; however, it can be noted from the table that different provinces have varying coverage rates as each province is managed separately. It is worth noting once more that while increasing the percentile population coverage by the DNA database is prioritized, a very low number of those people included in the Chinese NDNAD are actual offenders (Ge et al. 2017). According to the proceedings of the fifth National Congress of Forensic DNA Database (Institute of Forensic Science 2017), an overwhelming majority of 95% of the profiles in the DNA database are reference samples collected from target populations. The ways in which target populations are set varies from province to province. The provinces of Xinjiang, Shandong, Henan, Jiangsu, Zhejiang, Guangdong, and the autonomous region of Beijing had over three million profiles inserted into the database. A particularly strong percentile coverage rate of Xinjiang is by far the most notable. Xinjiang is the only case where a whole-population database was ordered by the authorities to be

Picture 33.2 Polymerase chain reaction machines processing thousands of reference samples at the Center for Applied Forensic Sciences and Technologies in Beijing, China

built (Human Rights Watch 2018). The province of Xinjiang, or the Xinjiang Uyghur Autonomous Region, is one of the most politically sensitive regions in China. The region has a large population of Uyghur Muslims, and tensions between Han and Uyghur ethnicities in the region are common; reported surveillance and control, peaking in the recent Chinese government recognition of "re-education camp" establishment for "thought reform" of hundreds of thousands Uyghurs, are justified by terrorism threats and the need to ensure the safety of the people (Human Rights Watch 2018). This can be related back to the large percentile coverage of the Xinjiang overall percentile population coverage (Picture 33.2).

According to the September 2017 presentation of IFS at the International Society for Forensic Genetics, 23.6% of crime scene samples match a person or another crime sample when entered into the database for the first time. For the China's NDNAD, a clear goal outlined by the government on multiple occasions is to increase the database match rate between crime case profiles (match rate of evidence-individual) and profiles in the database by continuously increasing population coverage (Bing 2015). Liu Yao, who is a representative of Institute of Forensic Science, has noted that China is lagging behind the percentile match rates of the UK and including more DNA profiles into the database would be a strong priority at IFS (Bing 2015). The author does not explain whether or not these profiles would be taken from offenders. While talking about the larger structures of running the DNA

database, "the assessment of the state of public order has relied on two indicators: the number of files places on police files (li'an) and the police clear-up rate (po'anlu)" (Sapio et al. 2017). This partially explains the reasons of assigning high priority to increasing NDNAD match rate. Besides ethics and financial viability, the ability to quickly match a crime scene DNA profile to a known person in DNA database would dramatically cut down on the work of the police, thus raising the statistical value of police work quality.

Y-STR database was officially announced in September 2012, as a response to the need for further DNA evidence to solve crime and the statistical conclusion that most violent crime is being committed by men. Ge et al. (2014) argue that since it is a well-established fact that men commit the overwhelming majority of violent crimes, male-specific Y-STR data can be used to rapidly and efficiently identify the perpetrators of crimes. Y-STR database in China has over seven million entries. Four main uses of the Y-STR database have been outlined: determining connections between surnames and Y-STR heritage in regions with little migration, familial searching using autosomal and Y-STRs to determine the identity of criminals, determining the inference of source population, analyzing of mixed samples (especially rape samples with small trace of any present male DNA), and the searching of missing persons or the relatives of missing persons (Ge et al. 2013). Determining statistically significant connections between Y-STR data and surnames, as well as being able to determine the inference of source population still need a lot published research from different areas in China. Currently, Y-STR database is identified as particularly important because genetic familial linkage information can assist the police in tracing close male relatives of a criminal in case an instant match is not available upon crime case profile search in the database. High discriminatory power and the advantage of being able to trace paternal lineage mean that while sometimes a database search might not provide a match, the related results would suggest a male family member of a crime perpetrator. In areas of China that have not seen much migration, genetic heritage is strongly connected and familial linkages can be quickly uncovered through data derived from Y-STRs, while maintaining sufficient discrimination power for direct comparisons of single-source samples as well. IFS confirms that, although increasing the number of crime scene and reference sample profiles in the database remains a continuous goal, familial searching, indeed, is being considered as a potential means of maximizing the efficiency of the database. In 2012, the work of building the Y-STR database was started in nine laboratories in six provinces, with the goal of further expanding and uniting this database. This type of database is not common in other places of the world, where most of the databases focus on non-gendered autosomal STR loci. China has taken a bold step to build a large national Y-STR database, starting with provincial crime laboratories. Since the Chinese judicial system is focused on being able to "break the case," match rates are strongly emphasized as prime indicators of database success. They note that the gender distribution of crime statistics should be an important driving factor in Y-STR database creation (Ge et al. 2017).

The third type of the DNA database created in China is the National Missing Children DNA Database. Issues of missing children were the reason why the work of

genetics was started in China is the late 1990s—NMCDD was the starting point of all of the databases. The issue of missing children has started in the 1970s, and around 200,000 missing children are reported each year, only 0.1% of them are found (Ge et al. 2013). If thinking chronologically, the database for missing children was the first one to be established; however, it took a long time for the project to be approved as a stand-alone database. In 2009, the Ministry of Public Security gave the go-ahead for the project.

> *(...) the Fifth Bureau of the Ministry of Public Security will re-develop and construct the "National Public Security Organization DNA Database Information System for Search of Trafficked/Missing Children" and the "National Public Security Organization DNA Database Information System for Search of Trafficked/Missing Children", which will be handed over to the Ministry of Public Security for evidence management and management. (...) Up to now, the library has accumulatively recorded 557,000 DNA data of abducted/missing children and their relatives. There are 4420 trafficked/missing children that have been found, and the longest time that a child has been lost is 36 years.* (Ge et al. 2017: 4–5)

In their doctoral thesis Chen Zeyu (2013) notes that in the years of database running, it has faced many issues ranging from ethical issues of DNA sample collection and management to the technical questions around the running of the database and suggests that two main advancements—resolving ethical privacy issues as well as increasing the efficiency of the database—should be resolved for a more efficient running of the database.

China's NDNAD, Y-STR database, and the NMCDD are the three main DNA databases in China. NDNAD has seen rapid growth aligned with the Party-state goals of achieving "harmonious society" and "peaceful development" of the country, further talked about in the next chapter. For the government to be able to maximize the benefits of the databases, data sharing and integration have been promoted. During the Fifth Plenary Session of the 18th CPC Central Committee in 2015, the largest and most important strategic governmental session in China, a slogan to "Implement the national big data strategy, promote open sharing of data resources" was created to signal the larger direction of database development.

33.5 Forensic Science Governance in China: Background, Institutions, and Technologies

The People's Republic of China (PRC) is a Party-state governed by the Communist Party of China. In the Constitution, which is the main governing document of the country, it is stated that the PRC is a socialist state under the people's democratic dictatorship governed by the principle of democratic centralism (Lin 2004). This means that China's governance is highly centralized where decisions are made by the main governing body and then trickled down for implementation. Thus, it is also the case in criminal justice in general, as well as the management of the China's NDNAD in particular. All of the criminal justice institutions are managed by the Ministry of Public Security, which is the main police and public security authority in

the People's Republic of China that ensures law enforcement and is also the governing body of all of the Public Security Bureaus across the country.

Since the establishment of the DNA database in 2009, it was run by the Ministry of Public Security. After the official establishment of the NDNAD in 2005 and National Missing Children DNA Database in 2009, the governance of these two databases was transferred to the Institute of Forensic Science in 2009. The institute, which is located in the capital city of Beijing, is part of the Ministry of Public Security and is the governing body directly responsible for the operation of DNA databases in the country as well as the independent advancement of proprietary Chinese DNA technologies. The responsibilities that the institute officially assumes are: (1) strategic management of the database, (2) setting of Statement of Requirements for the National DNA Database, (3) monitoring of Forensic Science laboratories, (4) delivering DNA database services to police forces, (5) ensuring integrity in the management and the DNA data held, (6) development of database software, (7) providing management information, such as standards and guidance, and (8) delivering training. For these responsibilities, the IFS reports directly to the Ministry of Public Security.

China is also of the few countries that adjoined a research center to develop own STR typing reagents as well as a genetic analyzer. The First Research Institute of the Ministry of Public Security is the largest and the highest unit of forensic science research in the public security system. Founded in 1960, the First Research Institute is also one of the oldest scientific research institutions in China. It is the research and development base of modern police technology and has long been responsible for the research and development of police science and technology. It plays a strategic role in China forensic science field as it pushes for China's technological independence. Accommodated within the First Research Institute umbrella, the Center for Applied Forensic Sciences and Technologies (as shown in Picture 33.3) develops and utilizes own technologies for reference sample inclusion in the NDNAD. Located on the outskirts of Beijing, the center has three functions: it is a research center, an STR kit manufacturing facility as well as a reference sample processing police DNA laboratory. During the fifteenth period, the Institute of Forensic Science of the Ministry of Public Security successfully developed domestic forensic DNA detection reagents, breaking the monopoly of foreign DNA technology products. During the 11th Five-Year Plan period, under the advocacy and care of the Ministry of Public Security, the First Institute of the Ministry of Public Security jointly developed the GA119-16A Genetic Analyzer and related analysis software and consumables with independent intellectual property rights, realized the complete supporting of domestic DNA testing reagents, domestic instruments, and consumables. While not well known internationally, the analyzer is a groundbreaking invention as prior to that Thermo Fisher Scientific had been the only supplier of this laboratory equipment in the world. The genetic analyzer was developed to ensure the government's self-sufficiency, shortened investigative time, and cost reduction.

Acquiring a software that would be able to assist the management of the DNA database has also been important to China. To be able to efficiently store, manage,

Picture 33.3 A government representative outside of the Center for Applied Forensic Sciences and Technologies in Beijing, China

and share DNA data across the country, the Ministry of Public Security invested in different types of data management software, which have undergone several developmental stages. Representatives from the Ministry of Public Security Material Identification Center have summarized the developmental pathway and main technologies used:

> *In June 2009, in order to completely solve the increasingly apparent incompatibility between the DNA database system and the work requirements of the public security organs at that time, the Ministry of Public Security approved the 'National DNA Database Upgrade and Renovation' [request]. In July 2014, the project passed the acceptance of the Ministry of Public Security. The project has completely solved the problems of the first generation DNA database level: many nodes, weak data flow monitoring, and slow comparison. The new generation of national DNA database structure integrates Management Information System (MIS), Laboratory Information Management System (LIMS), DNA Data Retrieval Comparison System (DIS), and duplicate personnel information search. System, catastrophic accident identification system (DVI) and data conversion tools and a series of standard data interfaces, and design entry, storage, comparison, review, backup and other functions according to 100 million data scale. In December 2015, a new generation of 'National DNA Database System' was used in 33 DNA laboratories in 9 provinces and cities including Beijing, Tianjin and Heilongjiang; in April 2016, 30 DNA laboratories in 7 provinces and cities including Shanxi, Inner Mongolia and Hunan. In June 2017, after nearly one year of actual test operation and continuous optimization and improvement, the Ministry of Public*

> *Security held a DNA Database Upgrade and Reconstruction Project meeting in Beijing, which officially launched the nation-wide construction and application of the second-generation national DNA database. The timeline for the replacement of the first-generation database is set as the end of this year. The second-generation 'National DNA Database' will better support rapid DNA query, classification and other applications of massive data quantities*. (Ge et al. 2017: 6)

The DNA databases are governed by a set of regulations named "Criteria for Forensic DNA Databases"(法庭科学*DNA*数据库建设规范). The set of regulations are fairly simple and even somewhat loose: while they do set the basic understanding of what elements constitute a DNA database, the document maintains ample space for the government to include, preserve, and use DNA data as deemed necessary. While the size of the DNA database is growing exponentially, it is not entirely clear what standards the government is upkeeping. There are two widely internationally accepted loci set standards: the expanded Combined DNA Index System (CODIS) loci standard is the United States set by the Federal Bureau of Investigation (Hares 2015) and the European Standard Set (ESS) loci decided on by the DNA working group of the European Network of Forensic Science Institutes (ENFSI). In China the situation has been different, and it has become a source of criticism both nationally and internationally. Since the beginning of DNA database construction, the loci standards were not set. The difficulties arose in the very beginning of the DNA database building as researchers discovered that due to genetic differences, some loci were close to obsolete in Chinese populations, such as loci typical in Caucasian and Black populations (Wang et al. 2003). As a result, different local and international manufacturers of the DNA technologies have followed own chosen standards, often the CODIS standard applied in the USA. Scientist Bing Liu (2015) outlined that due to lack of regulations, various DNA technology providers have come up with own versions of STR analysis kits causing a fair sense of havoc and advised that despite the rapid growth of the database, too little thought was put into setting standards for the industry. In 2015, the Chinese NDNAD was undergoing a major technological upgrade, and grasping that opportunity, Bing urged the government to put further thought into learning from the experience of foreign scientists and considering various factors that would make a major difference in the application of the DNA database (Bing 2015). Currently, a new standard of 20 loci is being applied, although it has not been officially published.

The "outsourcing" of governance responsibilities concerning DNA databases to one dedicated agency has proved to have increased the efficiency standardization. The IFS has put efforts to increase the efficiency of the DNA database in terms of pushing for the development of DNA analysis technologies locally, upgrading the DNA analysis software across the country, and is aiming to unify the loci standards used by different laboratories. Further efforts should be invested toward the transparency of DNA collection rules and regulations as the current standards are loose.

33.6 Future Directions

The Chinese DNA database has exhibited extreme growth and has shown voracious adoption rates of new technologies and heavy allocation of resources for innovative DNA research. The country has already implemented three types of DNA databases (NDNAD, NMCDD, and Y database), fingerprint database, face image database, and has already planned for an SNP database, according to the announcement made during the tenth AFSN Annual Meeting in Beijing, China (Wang 2018).

Next-generation sequencing (NGS), also referred to as second-generation sequencing and massive parallel sequencing (MPS), has risen in popularity in the Chinese forensic circles. During international presentations as well as published articles, IFS has indeed signaled that they would be ready to consider NGS technologies in the future: "Currently, MPS appears sufficiently robust to type reference samples for uploading DNA profiles into databases. With the technology evolving, it is likely that in the near term MPS will be able to offer the sensitivity of detection to analyze low quantity and quality DNA samples, and will be capable of analysis of forensic casework evidence" (Ge et al. 2014: 13). Furthermore, during 2017 ISFG meeting, IFS representatives noted that the capacity of the Chinese database is ready to hold a large increase of data that using MPS technologies would require. In the 2018 AFSN event, SNP profile database was confirmed as planned for implementation.

The SNP database will be running on NGS equipment: a new type of genetic analysis equipment, also referred to as second-generation sequencers, which is able to analyze large amounts of different types of genetic data at the same time. Once implemented, it will be one of the most technologically advanced databases used around the world. In a recent publication by the IFS, the clear vision of how NGS would be used with the existing NDNAD and the biggest challenges were discussed. NGS would be incorporated into the existing NDNAD instead of replacing it, and the critical challenge was to resolve compatibility issues between NGS-generated data and the currently used STR data (Zhao et al. 2017). The article comments: "From the perspective of its developmental history, the applications of forensic DNA databases have closely followed the development of DNA technologies applied in the field of biology. Thus, following the development of second-generation sequencing technologies and platforms as well as the development of many emerging biological genetic technologies such as SNP, mRNA, and whole genome sequencing, the next step in the development of DNA databases should be gradually combining the new and the currently used technologies to support better data use and retrieval" (Zhao et al. 2017: 257). One of the main considerations prompting the adoption of NGS is the desire to make the most use of the currently existing data through increasing its efficiency. NGS could be a potential answer: as previously mentioned, it can analyze different types of data in the same reaction, thus increasing the accuracy of the profiles. It is recognized that the costs of running NGS tests as well as the need for the technology to mature and be fully compatible with the existing NDNAD as two main challenges to be solved before a large-scale adoption is undertaken (Zhao et al. 2017). The Beijing Institute of Genomics, Chinese

Academy of Sciences, has taken an active role in NGS research in China, and they have identified and called for resolution of the following issues: low-template library preparation, error rate, type estimations, as well as existing problems with NGS data processing and mining (Yang et al. 2014). Despite the existing issues and the high costs, it can be foreseen that the government will adopt the technology once all the issues are resolved.

China has successfully applied various emerging technologies to ensure the wider societal goals to ensure the stable growth and development of the country. With the NDNAD, Y-STR, and NMCDD to manage and an articulated need to obtain China-specific DNA markers with high discrimination, MPS technologies do have the potential to unify all of the present needs of the police on one platform. This is one single large development, which is known to be under serious consideration by the highest governing units in China.

33.7 Issues Pertinent to the DNA Databases

The problem of order is a genuinely transhistorical problem rooted in inescapable conflict between the interests and desires of individuals and the requirements of society.
Dennis Wrong, "The Problem of Order" (Wrong 1994).

China has made a strong push to ensure that forensic DNA technologies are an integral part of public safety and social stability. Managing around 600 laboratories across the country—their staff, buildings, technical and governance standards—is not an easy task. It is not surprising that the DNA databases have not been without issues. In a recent publication, the Physical Evidence Identification Center, the Ministry of Public Security, identified six main issues that the NDNAD is currently encountering, the main one being the lack of adequate management skills as well as the infrastructure and quality management capabilities. The database has been ordered to be created so quickly that personnel training has lagged behind and hard skills of DNA laboratory management have not followed (Ge et al. 2017).

The authors also comment on the serious lack of both personnel and laboratory issues: "Some of the newly built, some of the rebuilt and some of the modernized and expanded DNA laboratories still have a low level of modernization and standardization. Some have just started work, and there is an obvious lack of functionality, laboratory room partition confusion, observed high risk for lab processes. National DNA database information quality monitoring can find contaminated data every month. Problems such as the failure of the quality management system to meet standards have seriously reduced the standardization, credibility and authority of inspection and appraisal. Among the 387 provincial and prefecture DNA laboratories participating in the 2016 Ministry of Public Security's proficiency assessment, 15 failed the results (2 provincial), and the rate of failed qualification assessment reached 3.87%." (Ge et al. 2017: 10). The second observed issue refers to the ability to collect physical evidence and handle samples. Ge et al. (2017: 10) comment, "In some places, especially in areas where DNA laboratories

have only recently been built, there is a significant lack of expertise, evidence handling awareness, technical preparation and operational skills in DNA technology. The on-site bio-material evidence extraction rate is far below the national average. Some provinces have more prominent problems, and the rate of on-site bio-materials extraction is less than 5%. In some provinces, the effective collection rate and sample size of DNA samples of eight key personnel such as crimes are very low, and repeated collections are at a high level for a long time. In some places, the problems of long-term backlog, failure to inspect samples or not inspecting samples in time are more prominent." Further, issues with ensuring adequate numbers of qualified specialists are pertinent. To comment in numbers, there were only 112 professional DNA technicians who qualified in 2006 national DNA laboratory assessment, and the number of full-time DNA database administrators and DNA intelligence analysts who have mastered DNA testing techniques and surveyed onsite, familiar investigative work, and proficient in computer technology is extremely scarce (Ge et al. 2017). With only the recent update of the DNA database management software, system security and risk management capabilities have been an issue. Although the DNA database operates on the public security network, there are still many risks in security management: in May 2017, after the outbreak of the "Eternal Blue" blackmail-born virus, DNA database server virus was found in 19 provinces and cities nationwide, causing system paralysis; the most serious consequence of DNA database virus was information loss that occurred in several individual cities (Ge et al. 2017). Failure to upgrade the DNA software system in time was the primary reason for that, which could have been prevented with stricter management of DNA laboratories. Such large-scale database server virus infection is a serious red flag as it poses a serious security threat. The final issue has to do with the staff skills to handle incoming information and their ability to successfully use the information at hand while solving cases.

All six of the technical issues mentioned here are serious problems. It can be seen that only half of the laboratories that are being run employ staff who is fully capable to work in a forensic genetics laboratory. Consequentially, without strong support in a genetic laboratory, new staff without prior experience in the field would find the work very challenging. In 2017, the Ministry of Public Security launched "Three Strikes and One Remediation" campaign to resolve some of the most crucial of these issues; the results of this campaign are yet to be measured and evaluated.

Technical issues of the database is only one side of the coin. Ethical issues have also been pertinent to the DNA database: the steps that the country has taken around the building of its DNA database could be called somewhat radical. The country has taken big steps toward large-scale monitoring of its population without putting much effort toward transparency. Some of the measures taken to build the forensic DNA database are implemented against well-researched experience from other countries. For example, in the UK, the number of subject, or reference, profiles loaded to the DNA database has been decreasing from year to year, but the efficiency measures by the match rate have increased. The percentile ratio of match rates as compared to the overall size of the Chinese DNA databases have not seen stable annual increases—in the year 2015, they had even decreased. Therefore, a possibility that the goal of

"growing" the database by adding as many reference sample profile inclusions as possible does not result in increased efficiency. The reason of why so many reference samples are "pumped" in the name of the growth of the DNA database are unclear. Common points of critique are often also directed toward the over-simplistic governance standards that barely account for physical integrity of citizens, over-policy regular citizens without criminal records, and infringe on their right to privacy.

33.8 Running the World's Biggest DNA Database: Conclusion

The Ministry of Public Security has pushed the NDNAD to grow and become the world's largest DNA database in just over 10 years. Additionally, China's Y-STR database is now also the world's largest database of such kind. To achieve this goal, the Chinese government has taken some radical decisions in terms of fueling the expansion of DNA laboratories without a set supply of specialists to work in them, lagging creation and implementation of standards and not publicly disclosing comprehensive uses of DNA databases. In some cases, work was started first and the relevant regulations followed later, creating concern around ethics of the sample collection, storage, and use of the DNA profiles. Since there is a lack of transparency, it can be argued that an authoritarian country with so much power over its people would not necessarily only use all the data collected for the purpose of solving crime. The implications of owning and managing this size of a database are varied and extending. China must pay close attention to maintaining a balance between the legitimate goal of maintaining a safe and stable society, identified as one of the major building blocks of the present political system, smart use of government funds, as well as the ethics of forensics in the country.

In the future, the size of the NDNAD will grow not only in size but also in complexity with the use of next-generation sequencing technologies, as well as the addition of the Y-STR database. With the rise of facial recognition technologies, it can be speculated that at some point in the near future, a crime could be solved by CCTV recognizing an offender and any DNA evidence conclusively confirming a crime. Overall, China is preparing to meet yet another Five-Year Plan in 2020 with its forensic technologies, laboratories, staff, and standards mature for another challenge to come. In a way, much of the rushed growth can be explained with the goals of planned governance and central management. We are yet to fully understand how such planning would work out in the context of forensic DNA.

References

Bing L (2015) A preliminary evaluation of DNA evidence in criminal investigations based on National DNA database. Foren Sci Technol 40(3):199–203. [Chinese Mandarin]

Chen Z (2013) Application and perfection of DNA database in the fight against child trafficking crime. PhD Dissertation, Southwest University of Political Science and Law. [Chinese Mandarin]

CODIS—NDIS Statistics (2018). https://www.fbi.gov/services/laboratory/biometric-analysis/codis/ndis-statistics. Accessed 14 Dec 2018

Cole D (1999) Race, policing, and the future of the criminal law. Hum Rts 26:2

Deray ES (2011) The double-helix double-edged sword: comparing DNA retention policies of the United States and the United Kingdom. Vand J Transnat'l L 44:745

Doleac JL (2017) The effects of DNA databases on crime. Am Econ J Appl Econ 9(1):165–201

Ge J, Yan J, Xie Q, Sun H, Zhou H, Li B (2013) Development of Chinese forensic Y-StrDNA database. Fayixuezazhi 29(3):212–215. [Chinese Mandarin]

Ge J, Sun H, Li H, Liu C, Yan J, Budowle B (2014) Future directions of forensic DNA databases. Croat Med J 55(2):163

Ge J, Liu F, Xu J (2017) History, current situation and future of DNA database construction in public security organs across the country. In: Ge J (ed) DNA database construction, application results and prospects V. Qunzhongchubanshe, Beijing. [Chinese Mandarin]

Hares (2015) Selection and implementation of expanded CODIS core loci in the United States. Forensic Sci Int Genet 17:33–34

Home Office (2019) National DNA database annual report, 2017 to 2018 [Government website]. GOV.UK. https://www.gov.uk/government/publications/national-dna-database-annual-report-2017-to-2018

Hu L, Chen S, Zhang G (2003) Value and important issues about national DNA database. Foren Sci Technol 6:3–5. [Chinese Mandarin]

Human Rights Watch (2018) China: Minority region collects DNA from MILLIONS. https://www.hrw.org/news/2017/12/13/china-minority-region-collects-dna-millions#. Accessed 2 Oct 2018

Institute of Forensic Science (2017) Working proceedings of the fifth National Congress of forensic DNA database. Ha'erbin, 12–14 Sept 2017

Interpol Global DNA Profiling Survey (2008). http://www.interpol.int/content/download/8994/66950/version/2/file/GlobalDNASurvey.pdf. Full report on: http://www.dnaresource.com/documents/2008INTERPOLGLOBALDNASURVEYREPORTV2.pdf. Accessed 19 Sept 2018

Lin G (2004) Leadership transition, intra-party democracy, and institution building in China. Asian Surv 44(2):255–275

McCartney C (2004) Forensic DNA sampling and the England and Wales National DNA Database: a sceptical approach. Crit Criminol 12(2):157–178

National DNA (2018) Database strategy board annual report 2016/17. National Police Chiefs' Council and Home Office. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/724596/040718_new_CCS0518718592_National_DNA_Database_Strategy_Board_AR_2016-17_updates_NEW.pdf. Accessed 13 Sept 2018

NDNAD (2013) Home Office National DNA database strategy board annual report. 2012–13. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/252885/NDNAD_Annual_Report_2012-13.pdf. Accessed 13 Sept 2018

Randerson J (2006) DNA of 37% of black men held by police. The Guardian 5. https://www.theguardian.com/world/2006/jan/05/race.ukcrime. Accessed 3 Oct 2018

Sapio F, Trevaskes S, Biddulph S, Nesossi E (eds) (2017) Justice: the China experience. Cambridge University Press, Cambridge

Skinner D (2013) 'The NDNAD has no ability in itself to be discriminatory': ethnicity and the governance of the UK National DNA database. Sociology 47(5):976–992

The Forensic Genetics Policy Initiative (2018). http://dnapolicyinitiative.org/. Accessed 13 Sept 2018

Tracy PE, Morgan V (1999) Big brother and his science kit: DNA databases for 21st century crime control. J Crim Law Criminol 90:635

Wallace HM, Jackson AR, Gruber J, Thibedeau AD (2014) Forensic DNA databases–ethical and legal standards: a global review. Egypt J Forensic Sci 4(3):57–63

Wang G (2018) Forensics solving cold cases in China. In: Tenth Asian forensic science network annual meeting, 4–8 Sept 2018, Beijing, China

Wang W, Jia H, Wang Q, Cai Z, Wei L, Wang D, Bittles AH (2003) STR polymorphisms of "forensic loci" in the northern Han Chinese population. J Hum Genet 48(7):337–341
Wrong D (1994) Problem of order. Simon and Schuster, New York
Yang Y, Xie B, Yan J (2014) Application of next-generation sequencing technology in forensic science. Genomics Proteomics Bioinformatics 12(5):190–197
Zhao Z et al (2017) The application of next generation sequencing in forensic DNA database. Sci Discov 5(4):257–261. [Chinese Mandarin]

DNA Databases: Risks, Benefits, Privacy, and Human Rights

34

Sachil Kumar

Abstract

Latest developments in DNA technologies and the DNA polymorphisms discovery have led to the formation of DNA databases for the purpose of forensic investigation. The aim of establishing forensic DNA databases was to assist the policeman with information on who might have been there at the crime scene, especially where the identity of those involved is unknown. If there is no arrests until the end credits, DNA profiles generated from samples collected from crime site will be retained in a repository which can be accessed later to track the "matching" perpetrator. Forensic DNA databases are currently in operation in approximately 69 countries, although others are being extended or developed in at minimum 34 new nations. These types of databases are needed in a populated country like India. The Government of India is working on a revised edition of the law aimed to set up a central DNA database for perpetrators. New issues are emerging, as predicted with the tremendous progress of the usage of DNA profile repositories.

The continuous growth in the size of repositories poses concerns on the criterion of inclusion and retention and suspicions on the usefulness, reliability, and privacy violation of such vast personal data. In view of its wide scope, the database posed questions regarding privacy, government monitoring, and human rights. The preservation of the DNA of an innocent citizen may be seen as an infringement into personal privacy and a breach into civil liberties.

Keywords

DNA database · CODIS · DNA profiles · STRs · Criminal investigation · Risks · Bio-surveillance · Benefits · Privacy · Familial searching · Human rights

S. Kumar (✉)
Department of Forensic Sciences, College of Criminal Justice, Naif Arab University for Security Sciences, Riyadh, Saudi Arabia
e-mail: SKumar@nauss.edu.sa

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4_34

34.1 Introduction

Recent developments in DNA technologies and the DNA polymorphisms discovery have led to the formation of DNA databases or DNA data banks for forensic enforcement purposes (Schneider 1998). As a result, a wide variety of options have been opened up for forensic investigation, because if we tend to equate the DNA analysis of the evidence collected at the scene of the crime with the collection of samples that make up the database, we may easily determine who the perpetrator might be and connect them to any other offenses that they might have perpetrated. If there is no arrests until the end credits, DNA profiles generated from samples collected from scenes of crime will be retained in a database which can be used later to find a "matching" perpetrator (Albujja et al. 2018).

Logically, as the number of people who have had their DNA analyzed and included in a repository grows, there is often a greater likelihood of identifying suspects. The permanent retention of DNA from scenes of crime on a repository contributed significantly to prosecutions in high-profile incidents, like the sexual assault and murder of Melanie Road, a teenage girl (17-year-old), A' level student, in June 1984. In 2006, Christopher Hampton, 67 years, a painter and decorator, pleaded guilty to Melanie Road's murder (Hayhurst 2016). The aim of establishing forensic DNA databases was to assist the policeman with information on who might have been there at the crime scene, especially where the identity of those involved is unknown (Wallace 2006). A data bank of DNA profiles from crime site scan also offer insight on the existence of links between various crime scenes, potentially identifying serial perpetrators and assisting in the analysis for identifying and analyzing patterns and trends in crime and disorder (McCartney 2006; Butler 2012).

Forensic DNA databases are currently in operation in approximately 69 countries, although others are being extended or developed in at minimum 34 new nations (Interpol 2016). The biggest DNA databases are in China (approximately more than eight million person DNA profiles, <1% of the populace), preceded by the United States (14.3 million DNA profiles, 4.5% of the populace) and the United Kingdom (5.7 million DNA profiles, 9% of the populace). The UK database is one that has decreased in size, deleting records from innocent persons hasn't had any significant adverse impact on crime clearance rates (DNA Policy Initiative n.d.-a).

34.2 United Kingdom National DNA Database

The UK became the first nation to proceed on the so-called forensic DNA revolution (Jeffreys et al. 1985a; Gill et al. 1985), with the DNA fingerprinting test being first applied for the resolution of the immigration dispute as a result of disputed family relationships in early 1985 (Jeffreys et al. 1985b; Aronson 2005). Soon afterwards, DNA evidence in a paternity dispute was submitted to a civil court in the UK. In October 1986, DNA typing in criminal investigations started with Enderby's murder trial, an investigation that contributed to the release of a key suspect following subsequent DNA testing proved his innocence. In 1987, DNA typing tests had

been submitted as proof in criminal courts in the United Kingdom and the United States, and then in 1988, the United Kingdom Home Office and the Foreign and Commonwealth Office had granted approval for the usage of DNA typing to resolve immigration issues linked to the contested family relationships (Home Office 1989).

It was soon realized that this method was theoretically far more commonly used in the justice system. As a consequence, the application of DNA in the area of law enforcement has expanded, enabled by rapid technological developments, followed by incremental legal reforms, enabling the acquisition and the usage of DNA profiles to become an integral part of the criminal justice cycle. Enthusiast use of the modern methodology has fueled by both political and media hyperboles regarding the effectiveness of DNA profiling in identifying, stopping, and regulating crime. DNA profiling has been celebrated as an uncomplicated success tale, with the National DNA Database set up without any explicit laws, thus evading any focused political or public debate (Wallace 2006).

The UK's National DNA Database (NDNAD) was established in 1995 utilizing the SGM (second-generation multiplex) profiling method. Since 1998, the SGM*Plus* methodology are used for DNA profiling comprising of eight CODIS loci (TH01, D8S1179, FGA, D16S539, D3S1358, D18S51, VWA, or D21S11), together with two additional markers, D19S433 and D2S1338, and also the AMELX locus (Butler 2006). However, as of July 24, 2014, samples are profiled using the DNA-17 profiling technique [kits of 17 STR loci (along with gender identifier)]. Only patterns of STRs are stored in NDNAD instead of an individual's entire genomic sequence.

All information kept on the NDNAD is regulated by a Tripartite Board involving the Home Office, the Association of Police Authorities, and the Association of Chief Police Officers.

More than a year after its launch, the NDNAD has expanded to include DNA samples from 2.7 million people—around 5.2% of the people of the UK—several of whom have never been charged with or convicted of any offence. It is the world's foremost and largest forensic DNA database of its kind. Under existing legislation, it could be extended to 25% of male population and nearly 7% of female population (Williams and Johnson 2005). It had 3.1 million profiles in 2005, 5.77 million by 2015 and 5.86 million by 2016 (National DNA Database Statistics 2015). The number of DNA profile records kept in NDNAD rose by more than 172,000 in the year up to March 2018, the highest annual rise since more than one million records were removed from the database after the "Protection of Freedoms Act" came into effect. As of March 2018, the NDNAD held 6,196,278 DNA profile, a rise of 2.86% on the total for March 2017. This follows a rise of 2.79% between March 2016 and March 2017 and a rise of 1.63% (to March 2016) and 0.88% (to March 2015) over the preceding 2 years (Statewatch n.d.).

34.3 New Zealand DNA Data Bank

New Zealand was the second nation across the globe to set up the DNA Profile Data Bank (DPD) in 1995 and this constructive approach to crime culminated in a high rate of success in providing useful information for unsolved offences (Harbison et al. 2001). In New Zealand, "the Criminal Investigations (Blood Samples) Act 1995" ("CI (BS) Act") authorizes the setup of a DNA data bank. In addition to the sex test, 15 STR loci are used in New Zealand to create DNA profiles (Institute of Environmental Science and Research Website 2013). This included the compilation of DNA profiles of registered criminals and volunteers in a national database, which is managed on behalf of the police by the ESR (Environmental Science and Research).

Project DPD includes two repositories: the National DNA Database (individual's profiles) and the Crime Sample Database (profiles from unresolved crimes). By making comparisons, potential perpetrators can be classified and offences linked.

There is also a stand-alone temporary database ("TD"), comprising DNA profiles from individuals who have been detained by the police or expect to be prosecuted. If such an individual is thereafter convicted, their profile would be passed from the TD to the DPD for permanent retention.

DPD utilizing the SGM DNA profiling system until July 2009 (Sparkes et al. 1996). Since that time, the AmpFlSTR® SGM Plus® system (the six core SGM loci, amelogenin, and four additional loci) has been used (Cotton et al. 2000).

The DPD currently maintains some 189,000 individual DNA profiles and more than 40,000 crime sample DNA profiles. New Zealand dominates the world in DNA matching with almost 70% of all unresolved offences loaded to the crime sample data banks effectively connecting to offenders and 30% to another incident.

34.4 France National DNA Database

French created the national DNA database named FNAEG (Fichier National Automatisé des Empreintes Génétiques) in 1998, which was used by the local gendarmerie and national police. The Guigou Legislation on the Prevention of Sexually Oriented Offences, enacted by the Plural Left Lionel Jospin Administration, established a forensic DNA data bank in 1998. The deployment, initially scheduled for 1999, was eventually accomplished in 2001, with the repository itself housed at Écully, controlled by a sub-directorate of the French police's technological and scientific divisions.

After the 9/11 attacks, the government of France extended the context of the repository to incorporate DNA linked to other indictable offenses. Yet another "law for interior safety" adopted on March 18, 2003 broadened the spectrum even more to include nearly all acts of abuse, but not traffic violations or crimes perpetrated overseas. Samples are taken from convicted persons as well as simple suspects. The law does not specify a minimum age (Le Monde 2007).

Matthieu Bonduelle, the Secretary General, Syndicat de la Magistrature stated that “nobody defends a universal database, but, in fact, it is being done” in 2009 (Le Canard enchaîné 2009; Ouvrez la bouche 2009).

FNAEG maintained DNA profiles from 2.6 million persons and 2.3 million unknowns remain from crimes by 2014 (DNA Policy Initiative n.d.-b).

34.5 Combined DNA Index System (CODIS)

The CODIS (Combined DNA Index System) is a system of DNA profile indexes established by the FBI (Federal Bureau of Investigation). The CODIS combines DNA and computer technologies into a powerful tool for linking crimes. This allows local and state forensic laboratories to digitally share and match profiles, linking violent crimes with each other and with known perpetrators. CODIS consists of three tiers (or levels); National DNA Index System (NDIS) enables states to compare DNA information with each other; State DNA Index Systems (SDIS) allows labs between states to exchange information; and the Local DNA Index Systems (LDIS) where DNA profiles originate.

In 1989, TWGDAM (Technical Working Group on DNA Analysis Methods) first announced the creation of a national DNA database (Budowle et al. 1998). FBI initiated a pilot DNA databased program involving 14 local and state laboratories in 1990 (Combined DNA Index System (CODIS) 2018). Congress enacted the DNA Identification Act in the year 1994 (42 U.S.C. § 14132) requiring the FBI to set up a nationwide DNA repository for incarcerated criminals and also specific repositories for missing people and forensic samples retrieved from crimes (42 U.S.C. § 14132 n.d.). The National DNA Index System or the NDIS become functional in 1998. Presently, all 50 states, the Federal Police, the Army Laboratories, Puerto Rico, and the District of Columbia are involved in NDIS (Federal Bureau of Investigation 2017).

The CODIS contains a number of specific indices for storing DNA profile information. There are three indexes for assisting in criminal investigations: the offender index, which comprises DNA profiles of people accused of violent offences; the arrestee index, which includes profiles of those arrested of crimes pursuit to the laws of the particular state; and the forensic index, which includes profiles collected from crime scene evidence (Data and Communication Flow in CODIS 2017). Other indexes, like the biological relatives of missing people’s index, the missing persons index, and the unidentified human remains index, are used to help identify missing individuals (Procedure for CODIS 2017).

CODIS-based identifications rely on STRs that are distributed in the human genome and on statistics that are used to measure the prevalence of that particular profile in the population (The Biology Project 2000; Forensics 2017). Similarly, CODIS requires data on mitochondrial DNA (mtDNA) to be added to missing person related indexes. Because mtDNA passes from a mother to all of her children, this concept can be used to link people over decades.

CODIS core is mentioned below; the most recent loci with asterisks have been introduced to the CODIS core in year 2017 (FBI CODIS Core STR Loci 2015; Butler 2006).

CSF1PO, D5S818, D18S51, D8S1179, D7S820, D16S539, D3S1358, D21S11, D13S317, TH01, FGA, TPOX, vWA, D2S1338*, D2S441*, D22S1045*, D10S1248*, D1S1656*, D19S433*, D12S391*.

Although the US database is not explicitly connected to any other nation, many organizations across the globe use the underlying CODIS program. As of 2016, 90 foreign laboratories in 50 countries have been using CODIS software (Forensics 2017).

34.6 CODIS Success Stories (CODIS n.d.)

The following are some CODIS success tales.

34.6.1 St. Paul, MN, November 1994

A person covering his face using nylon stocking and carrying a knife leap out from underneath the bushes and sexually harassed a female walking around. Semen retrieved from the sufferer's clothing and saliva was tested using DNA technology. The resultant profile was looked up against Minnesota's CODIS repository. The hunt revealed the offender "Terry Lee Anderson," who admitted and now is behind bars.

34.6.2 Tallahassee, FL, February 1995

The Florida Department of Law Enforcement correlated semen spotted on a Jane Doe murder victim to a convicted perpetrator's DNA profile. The perpetrator's DNA was obtained, assessed, and retained in the CODIS data bank while he was jailed for another sexual assault. The match was informative; it stopped the perpetrators' from being released on parole, slated 8 days later.

34.7 Australia

The National Criminal Investigation DNA Database (NCIDD), managed by CrimTrac agency, was initiated in 2001. The NCIDD offers police and forensic experts with access to DNA profiles of alleged offenders, which can then be compared with evidence collected from crime scenes and persons. The repository is intra-jurisdictional, implying that all Australian state and territory repositories are interconnected, and DNA profiles may be exchanged (CrimTrac 2014).

The Australian Criminal Intelligence Commission (ACIC) is now running the NCIDD following the integration between the CrimTrac and the Australian Crime Commission in 2016 (Minister for Justice Michael Keenan 2015). ACIC is established under the ACC Act (Cwlth) (*Australian Crime Commission Act 2002*) (Starling 2018).

Through July 2018, the NCIDD has 837,000+ DNA profiles and is increasing continuously (Commission, Australian Criminal Intelligence 2018; Mobbs 2001). The NCIDD was using nine STR loci and a gender identifier (Amelogenin locus), but in 2013, it was expanded to 18 core STR loci (Curtis and Hereward 2017). The NCIDD was updated in 2015 to include modern technologies. This repository was among the most advanced throughout the globe, with enhanced kinship matching, familial searching, as well as advanced direct matching.

34.8 Canada

The National DNA Data Bank of Canada (NDDB) was founded in 1998 but started operations in the year 2000 (Milot et al. 2003). Federal courts have concluded that the legislation enacted by Parliament to control the usage of this technology in the justice system is mindful of the constitutional and privacy rights of offenders, particularly of individuals alleged to be convicted of serious crimes (National DNA Data Bank 2013).

After the UK's launch of its data bank (NDNAD) in year 1995, the acquittal of two false convicts (Milgaard case, 1969 and Morin case, 1985) and the adoption of the C-104 bill (to amend the Criminal Code and the Young Offenders Act) acknowledged the need for such a specific data bank in Canada and prompted the establishment of the NDDB via the Identification Act (Law C-37 of December 10, 1998) (Curran 1997). The Government of Canada adopted the DNA Identification Act on December 11, 1999. This would allow the creation and modification of a Canadian DNA data bank for the purposes of the Criminal Code. This law became legal on June 29, 2000. On the basis of 13 CODIS core loci, profiles are managed and compared using CODIS (Bank, Government of Canada 2001; Police 2012).

NDDB comprised DNA profiles in two separated indexes: Convicted Offender Index (COI), biological samples resulting in DNA profiles can only be collected from perpetrators against whom post-conviction DNA data bank orders have been made; and Crime Scene Index (CSI), DNA profiles derived from crime scene investigations.

As of July 15, 2013, the COI had over 237,000 DNA profiles and the CSI had over 87,000 DNA profiles (Milot et al. 2013). Moreover, there were 21,563 matches between crime scene and convicted offender DNA profiles and 2741 "forensic matches" per crime scene.

34.9 Dubai

In 2006, the United Arab Emirates declared that it would create a DNA database for its whole populace. The UK Forensic Science Service (FSS) officially signed contracts to build a database, but the FSS has since shut down due to recession.

In the year 2017, Dubai unveiled an initiative named Dubai 10X, designed to generate "disruptive innovation" in the region (Sutton 2017). DNA database was one of the programs in this initiative that would gather genomes from all three million nationals over a 10-year span. This was expected to use the database to determine the hereditary origins of diseases and to create customized medical therapies (Treviño 2018).

34.10 Germany

Germany established its DNA data bank for the Federal Criminal Police (in German: Bundeskriminalamt, abbreviated BKA) in 1998 (GeneWatch UK 2016; Germany's DNA Database 2016; National DNA Intelligence Databases in Europe 2016; Peerenboom 1998). At the end of 2010, the database held more than 700,000 DNA profiles, and 1,162,304 DNA profiles in 2017 (Käppner 2016).

In the year 2011, "Stop the DNA Collection Frenzy!" movement and numerous human rights and data security groups submitted a public letter (Ope Letter Stop 2016) to the Minister of Justice of Germany urging her to take steps to avoid the "preventive expansion of DNA data-collection" and the "preemptive use of mere suspicions and of the state apparatus against individuals" and the cancelation of foreign DNA data sharing initiatives at European and transatlantic stage (Schultz 2016).

34.11 Israel

Israel Police DNA Index System (IPDIS) is the Israeli national DNA data bank (Zamir et al. 2012) which was created in the year 2007 and has a database of more than 135,000 DNA profiles and more than 2000 confirmed hits reported. The IPDIS contains DNA profiles of accused, suspect, and convicted criminals.

In order to manage the high-throughput processing and DNA analysis with FTA cards, IPDIS has built a semi-automatic LIMS system that enables a limited number of police officers to complete the collection of a vast number of samples in a fairly short time period and are also liable for the future follow-up of assays.

34.12 Kuwait

Kuwait created its DNA database in 2002 without any relevant legislation on DNA databases. There were renewed calls in 2012 for expansion. In 2015, the Kuwaiti government passed a law (Act No. 78/2015) requiring not only of citizens living in the Persian Gulf state but also of other residents and even country visitors to have their DNA onto a national repository (Visser 2015; Ley 78/2015 de Kuwait 2016). Such a database would have been the first of its kind in the world.

The explanation for this legislation was safety issues following the ISIS suicide bomb attack (Imam Sadiq Mosque), which killed 27 people and wounded several more (ISIL 2015). They planned to finish collecting the DNA from approximately 4.2 million people in a security crackdown to help officers detain suspects by September 2016, which outside analysts felt was unrealistic (Field 2015).

The Ministry of the Interior has allocated $400 million to the project, that is likely to be criticized by human rights defenders for violating the person's right to privacy. Many that fail to give samples could face a penalty of up to $33,000 or $29, EUR 700, and a 1-year prison term, and those found supplying false samples will be imprisoned for 7 years.

Nevertheless, as the legislation is now scheduled to be implemented soon, several researchers and advocacy groups have voiced their doubts and pleaded with the Government of Kuwait to reconsider its plans. In the New Scientist, geneticist Olaf Riess claimed that the legislation was a "huge attack on genetic confidentiality" which severely jeopardized Kuwait's global image. The Kuwait Constitutional Court, in 2017, ruled against the statute that there was a breach of personal privacy and that the program had been terminated (Coghlan 2017).

34.13 Brazil

The Forensic DNA Research Institute of Federal District Civil Police in 1998 set up a DNA data bank designed to help resolve serious offences (Ferreira et al. 2015). In 2012, Brazil approved a national legislation creating DNA repositories at national and state levels for the DNA typing of people with criminal convictions (Ferreira et al. 2015).

In 2013, the Brazilian National DNA Database (BNPG) was created. In the first half of 2019, the number of genetic profiles that entered BNPG increased by 70%. In addition to this rise, the number of matches in the same time have risen by 46%.

Following the 2013 declaration of the President of the Republic of Brazil, which governs the 2012 Act, Brazil has started utilizing CODIS as well the DNA database of sexual assault evidence to resolve abuse cases in Brazil (Ferreira et al. 2015).

34.14 Other European Countries

The Netherlands is the biggest gatherer of its people's DNA profiles compared to other European countries. The DNA database, containing DNA profiles, was established at the National Forensic Institute (NFI) in 1997. The concept "database" in Dutch law involves the collection of autosomal STR profiles that are held in this repository (Beekman et al. 2004). Right now, the DNA repository at the Netherlands Forensic Institute holds DNA profiles of more than 316,000 Dutch people (Veiligheid 2013). In comparison with the situation in most other European nations, the Dutch authorities have wide-ranging rights to collect and preserve DNA samples if an individual is guilty of a recordable crime, except when the conviction only involves paying a penalty.

Throughout Sweden, only the profiles of offenders who have served more than 2 years in jail are kept. In Norway, court orders are necessary. Austria introduced their own DNA database in 1997, which is run by the Ministry of Interior (Hindmash and Prainsack 2010), and Italy also set up a criminal DNA database in 2016 (Negri 2016; Jamaica Observer 2016). Switzerland launched a provisional criminal DNA registry in 2000 and passed the law in 2005 (Haas et al. 2006).

The Government of Portugal suggested the development of a DNA data bank for the whole populace of Portugal in the year 2005 (Newropeans Magazine n.d.). But, after an informed debate, including the view of the Portuguese Ethics Council (CNECV—ConselhoNacional de Éticapara as Ciências da Vida n.d.), only the criminal's DNA profile were kept in the database (Skinner 2010).

GMI (Genomics Medicine Ireland), a life sciences company, is establishing a genetic registry that will include information from almost 10% of the Irish populace, involving patients with different disorders and normal individuals (Genomics Medicine n.d.; Irish Times n.d.-a). The likelihood of a private company retaining public DNA data posed questions. At present, it appears that Ireland has adopted a fully commercial approach to genomics. This policy jeopardizes the unrestricted access of genomic data for scientific researches which could aid the patient (Irish Times n.d.-b).

34.15 South Africa

The National Forensic DNA Database (NFDD) is the South Africa national DNA data bank used by law enforcement agencies. Sexual abuse was a major burden on the nation, and the country was reported with the maximum rape rate, with a female sexually assaulted every 17 seconds posing massive challenge to the nation and highlighting the severity of the issue. Thus, in 2014, President Zuma approved the legislation "*DNA Act*" to protect the innocent (The Evolution of Forensic DNA Database Policy 2017; Global Offender Data Base 2014).

34.16 China

In 2004, China set up its DNA database to track abducted and missing children. Since 2001, Hong Kong within China has had a separate database, utilizing the United States CODIS, the FBI's forensic DNA database system (DNA Policy Initiative n.d.-c).

The 2011 Interpol survey reported that China's DNA data bank has expanded to 461,513 crime scene DNA profiles, 7,701,745 individual's profiles, 22,718 missing person's profiles, 65,115 unidentified human remain profiles and 279,999 other profiles (total: 8,531,099 DNA profiles). Chinese data bank is projected to have 20 million profiles by 2013 and 40 million by 2016, rendering this the biggest in the globe (DNA Policy Initiative n.d.-c).

The Chinese authorities intend to increase their current DNA archive to 100 million records by 2020, according to the *Wall Street Journal* examination of reports from police departments throughout China (WSJ n.d.). There has been criticism that China can use DNA repositories not only to solve offence but also to monitor dissidents, particularly Uyghurs (The New York Times n.d.).

34.17 DNA Database in India

In India, also, this type of DNA profile database is needed which may help Indian law enforcement agencies. DNA fingerprinting has been a useful tool for law enforcement, as it works both ways to secure correct convictions and also to exonerate the innocent. India's latest compilation of crime statistics for 2016 from the National Crime Records Bureau (NCRB) indicates that the rate of conviction for rape is 25%. The 2016 NCRB statistics documented more than 300,000 crimes that directly affect people or property, but the conviction rate remains at around 30%. Only a small proportion of these criminal offences are currently being investigated through DNA testing (Kumar et al. 2016).

The Union Government is focusing on a revised draft of the legislation which aimed at creating a national DNA database for offenders. The proposal to draft a bill authorizing the usage of DNA profiles for crime-related purposes started in 2003 when the Department of Biotechnology (DBT) set up a council named the DNA Profiling Advisory Committee to make guidelines for the drafting of the "*2006 DNA Profiling Bill*" that ultimately became the "*2007 Human DNA Profiling Law*" (International Forensic Sciences n.d.; Maguire et al. 2013; Jeffreys 2005; Machado and Silva 2014). The proposed "*Human DNA Profiling Bill*" was disclosed in 2007, but never presented in Parliament.

There are 17 sites or locations along the DNA strand that we used to identify an individual. However, neither of these sites ever disclosed nothing about an individual's medical records. They are almost as neutral as fingerprints (Livemint n.d.). Rights activists criticized the law as a potential infringement of the rights of residents and questioned it on legal and technical grounds. Dr. Helen Wallace from GeneWatch, a UK-based not-for-profit organization that advocates against DNA

data banks, believes that India needs to learn from foreign experience, especially from the United Kingdom, which in 1995 became the first nation to create a data bank that also prompted retention of DNA records of innocents (The Hindu 2020). The United Kingdom in 2011 adopted the "*Protection of Freedoms Act*" which would delete an additional one million records from the DNA database. Jeremy Gruber, Executive Director, Council for Responsible Genetics, said that the Indian data bank was troublesome as it would be a Big Brother surveillance tool rather than a crime-solving platform (Livemint n.d.).

The revised 2012 Profiling Bill proposes just incremental changes to its counterpart, the 2007 Draft "Human DNA Profiling Bill." In a contentious move in 2012 to increase state intervention in the lives of common citizens, the UPA government introduced the DNA Profiling Bill at the Parliament's Winter Session. If it becomes law, the bill would give the right to obtain vast amounts of confidential DNA information from people even if they are "suspects" in a criminal trial. The information will be retained until the person has been cleared by the tribunal (Sanhati n.d.).

The bill advocates the creation of a National DNA Data Bank to be supervised by an officer who is of the rank of Joint Secretary to the Government of India. For protestors, this will help the government to play the role of an alarming Big Brother gathering a vast amount of sensitive data of citizens. The proposed amendment to the bill states that "DNA analysis offers sensitive information which, if misused, can cause harm to a person or society." It is not obvious under what conditions "volunteers" will disclose their personal details with the government.

The data, according to the bill, will also be used for the "creation and maintenance" of population statistics and can be used for identification, research, protocols development or quality control. Surprisingly enough, the punishment for "misusing" the DNA profiles is a jail sentence of a few months or a fee of a paltry, Rupees 50,000. In reality, law enforcement organizations like the CBI have lobbied the government to enact the bill early. They referenced the outcomes of a British parliamentary report published by the Department of Science and Technology in February 2006 claiming that prosecutions in crime trials rose significantly after the government decided to retain the details on DNA profiling in future years.

In the meantime, senior civil servants who are associated with the bill and have made detailed briefings to the DoB are unhappy that the bill calls for the deletion of DNA profile data after an individual has been discharged by the courts. They believe that preserving and growing data gradually and progressively would go a fair way in the detection and settlement of offences. Though this is a valid point, the absence of a strict privacy legislation poses questions regarding the obtrusive existence of the proposed DNA Profiling Bill. The proposal has been questioned for not resolving privacy issues. The Citizens Forum for Civil Liberties criticized the bill on privacy and in 2012 submitted a petition to the National Human Rights Commission (The Hindu 2012).

The Bill was further developed by an advisory committee because of concerns about privacy and lack of safeguards. Nonetheless, a new draft bill was presented to Parliament in the summer of 2015 but was opposed by a number of committee members who said their concerns had not been fully addressed. A public

consultation was then announced (deadline August 20, 2015). An expert committee was set up by the DoB to address these issues.

A revised version of the bill was sent to the Indian Parliament in August 2018, but it was not implemented before the end of the parliamentary session. As a result, after the 2019 Lok Sabha General Election, the bill was reintroduced again. The Act was approved in the Lok Sabha in January but lapsed because it could not be passed in the Rajya Sabha owing to lack of support from the opposition parties. Soon, after the Cabinet's approval, a new bill would be presented in Parliament, sources aware of the developments stated.

The Act needs informed consent from individuals to obtain DNA samples from them. Permission shall not be needed for crimes punished by more than 7 years in jail or by a death penalty. The genomic details of 10,000 Indians will be collected in the first stage of "Genome India." The proposed National DNA Database, or the Indian Database as it would be named, would merge profiles from multiple state-level repositories and retain at no less than six indices: an offenders index, suspects index, missing people's index, a crime scene index, unknown deceased person's index, and a volunteer's index. Those guilty of an offense would have their profiles retained in the data bank. Missing persons eventually identified or accused who have been acquitted in an incident should be removed from the database when a court directs the DNA database bank manager to do so (Livemint n.d.).

34.18 Benefits and Risks

Expansion of forensic DNA data bank poses concerns about the criterion for inclusion and retention and doubts about the effectiveness, affordability, and privacy infringement of such large collections of personal data. Unlike in the past, not just violent crimes but all offences are subject to DNA analysis creating thousands and thousands of DNA profiles, all of which are stored and continually searched in national DNA repositories. As always when big datasets are gathered, new mining procedures based on correlation became feasible. For example, "*Familial Searching in which a profile from crime site is intentionally run through the criminal database in the hopes of getting a list of profiles that are genetically identical to the DNA evidence and using this knowledge as an investigation leads to questioning families of the partial matches profiles*".

Finally, the very first familial searching was carried out effectively in the United Kingdom in 2004, which led to the prosecution of Craig Harman for assassination. Craig Harman from Frimley, Surrey, was prosecuted and imprisoned for 6 years on the grounds of a familial DNA search that connected him to the murder scene via a near relative's DNA profile. Another family search that hit the headlines in the UK was the case of Lynette Deborah White, who was killed on Valentine's Day in Cardiff, Wales, in 1988. Three locals were falsely jailed for life in 1990 and declared innocent and released after appeal in 1992. In 2009, 13 police officers involved in the case were charged with a plot to pervert the course of justice by jailing three innocents.

In 2000, a DNA profile of the alleged murderer was retrieved, but no match was made in the NDNAD. A family search was undertaken in 2003 and a close resemblance to a 14-year-old child with an identical genetic profile was revealed in the database. Since the child was not yet born when the offense was done, the law enforcement officials searched to his nearest associates or immediate family. This guided the investigators to Jeffrey Gafoor, a 38-year-old paternal uncle of the boy. Gafoor has admitted to White's assassination and has been sentenced to life in prison. Subsequently, the technique has been adopted in some US states, but it is not enforced at country level.

German police were often also active with Family Search strategies. In Northern Germany, the cops detained a man convicted of raping a woman after they had examined the DNA of his two brothers who had taken part in the dragnet. Owing to partial matches between the DNA profiles of the crime scene and these brothers, the perpetrator was confirmed. In contrast to other jurisdictions, the Federal Constitutional Court of Germany ruled against the potential usage of this sort of evidence by the court in December 2012.

Familial searches are certainly a useful tool in police investigations that can help to solve previously unsolved crimes. The biggest criticism is about how minorities are going to be affected. Ms. Simoncelli claimed that the usage of familial searches would harm a specific group of people—minorities. Simoncelli claimed that African Americans and Latinos should be increasingly targeted by law enforcement; that family members would be put under "genetic surveillance" for offenses they did not undertake and that family members would become "genetic informants". Simoncelli also stated that if familial searches were carried out on a regular basis, hundreds of thousands of innocent citizens who happen to be relatives of persons in the DNA database will be exposed to lifetime monitoring. Defense Attorney Stephen Mercer claimed that the risk is in the usage of the DNA database to decide if individuals who perform violent actions have genes that tend to lead them to abuse or other disruptive behavior (DNA Forensics n.d.).

Earlier, Alec Jeffreys questioned how the DNA profiles collected by UK police held not only individuals who were wrongfully convicted of crimes but also people arrested without prosecution, suspects acquitted in the inquiry, or also innocent citizens who had never been charged with an unlawful act. He also questioned the fact that large national DNA repositories, such as the NDNAD in the UK, are likely to be socioeconomically biased. It has been reported that most matches are for petty offences; as per GeneWatch in Germany, 63% of the data bank matches are for robbery, and fewer than 3% are for murder and rape. The amendments to the UK database came in 2012s "*Protection of Freedoms Act*," after a significant setback at the European Court of Human Rights in 2008. Since 2013, 1.1 million profiles (of about seven million) have been deleted to eradicate profiles of innocents from the data bank.

Despite the risks that such a universal system poses to people's freedoms, the government may not seem overly worried. A new survey of public opinion on issues related to DNA databases has found that a more censorious approach against wider national data banks is connected to the age and schooling of participants. There is a

need to raise awareness of the risks and benefits of very large-scale DNA collections and to follow specific ethical and privacy principles for the creation and regulation of DNA repositories where citizens' views are taken into account.

34.19 Privacy and Human Rights

The usage of DNA to locate individuals accused of perpetrating an offence has been a big move forward in the world of policing. If profiling is used cautiously, it can serve to prosecute people who have committed atrocious crimes or to absolve others those who are innocence. Nevertheless, issues occur if individual's tissue samples, computerized DNA profiles, and personal records are retained permanently in a DNA database. It is worried that this knowledge may be used in ways that jeopardize the privacy and rights of individuals and their families. The retention of an individual's DNA profiles and other information on a computer database thus allows a form of "bio-surveillance" or biological tagging, which can be applied to ascertain where they have been. This assures that DNA repositories may be used to identify people who have not perpetrated a crime or whose "crime" is a nonviolent protest or confrontation. For instance, in a nation where freedom of speech and democratic freedoms are limited, the government or intelligence services might attempt to collect DNA samples from a civil gathering to ascertain whether or not certain people have been involved. DNA repositories link searchable computer records of personal demographics, with the potential to biotag an entity and monitor their position using their DNA profile. Family members of an individual can often be identified through a partial matching of their DNA. There is concern that the Chinese government are using Genetic tests to censor the country's overwhelmingly Muslim Uyghur population. Human rights organizations claim that China has set up a massive DNA list of Uighurs residing in Xinjiang province, gathering samples without approval from approximately 36 million people as part of a free medical check-up scheme. Chinese scientists in 2014 released a report explaining how to separate Uighurs from other ethnic groups by utilizing DNA testing (Independent n.d.-a). DNA databases therefore dramatically change the balance-of-power from individual to state.

"Bio-surveillance" concerns stretch outside the state to anybody who may breach the system and gain access to the person's DNA profile. That may involve organized criminal or terrorist organizations, or someone trying to locate a person. People on witness protection schemes, for example, can have their image changed but cannot modify their DNA. When someone is aware of them and gathers their DNA, their identification may be disclosed by comparing it to the DNA profile contained in the database, whether it is available and connected to their old identity. Their families may also be identified through "familial searching" (seeking partial matches with other people's DNA profiles in the database).

These issues do not apply exclusively to the collection, retention, access, and usage of DNA samples that form the basis of the DNA profile, but rather to other information that may be stored. For example, as DNA is collected on arrest and

retained permanently, extra information is preserved in police arrest reports and in samples that might be kept in the DNA research laboratories. Individuals are worried that potential employers, other government departments, or even insurance providers will have access to their genetic records. Medicare providers will have a significant interest in verifying the medical integrity of those needing health care; employers would even have an interest in collecting statistics about the physical fitness or race and heritage of a prospective employee. Access to private details can have an effect on the employability of the individual applying for a position.

Kuwait introduced the world's first legislation in 2015 allowing both citizens and visitors to submit samples of their DNA, although it was revoked prior to its entry into effect. It was introduced with the intention of acknowledging extremists, but in 2017 the Constitutional Court ruled that the bill breached Kuwait's fundamental protection of personal liberty. Rwanda also introduced the first country-wide DNA database in the world, a scheme requiring the gathering of samples from all 12 million residents of the region in an attempt to crack down on violence. The program generated questions from human rights activists or human rights campaigners who fear that the database might be misused by the regime to breach international human rights law (Independent n.d.-b).

34.20 Conclusion

DNA databases are well established in several nations across the globe. These databases have proven to be a valuable resource in the battle against offence (fast and equal conviction of crimes, exoneration of accused individuals, and diminishing ethnic differences in law enforcement), which, from the perspective of public protection and criminal justice, make such a database extremely appealing. A number of people have a strong, adverse response to the concept of a DNA repository. Their main fear is that having a DNA database would massively intrude on their privacy and human rights. Safeguards are necessary because:

- DNA databases may be used to track individuals or their families (biological tagging or "bio-surveillance," so that the DNA database can be misused by regimes or someone else who may breach the system).
- DNA databases are connected to certain electronic documents, such as detention reports, which could influence the career prospects of the individual applying for a position or visa or to discriminate against them.
- Familial searching may contribute to major violations by enabling investigators or anyone who takes down the data bank to track down or harass political dissidents' relatives, or identify paternity and non-paternity for private, business, or illegal reasons.

Maintaining national DNA databases ought to pay heed over the adequacy of safeguards to secure privacy rights and prevent adverse social impacts.

References

42 U.S.C. § 14132 (n.d.) Index to facilitate law enforcement exchange of DNA identification information

Albujja MH, Bin Dukhyil AA, Chaudhary AR, Kassab AC, Refaat AM, Babu SR, Kumar S (2018) Evaluation of skin surface as an alternative source of reference DNA samples: a pilot study. J Forensic Sci 63(1):227–233

Aronson JD (2005) DNA fingerprinting on trial: the dramatic early history of a new forensic technique. Endeavour 29:126e131. https://doi.org/10.1016/j.endeavour.2005.04.006

Bank, Government of Canada, Royal Canadian Mounted Police, National DNA Data (22 Apr 2001) National DNA Data Bank

Beekman M, Posthuma D, Heijmans BT, Lakenberg N, Suchiman HED, Snieder H, de Knijff P et al (2004) Combined association and linkage analysis applied to the APOE locus. Genet Epidemiol 26(4):328–337. https://doi.org/10.1002/gepi.10318

Budowle B, Moretti TR, Niezgoda SJ, Brown BL (1998) CODIS and PCR-based short tandem repeat loci: law enforcement tools. In: Second European symposium on human identification, vol 7388. Promega Corporation, Madison, WI

Butler JM (2006) Genetics and genomics of core short tandem repeat loci used in human identity testing. J Forensic Sci 51:253–265

Butler J (2012) Advanced topics in forensic DNA typing: methodology. Elsevier, London

CNECV—ConselhoNacional de Éticapara as Ciências da Vida (n.d.)

CODIS—combined DNA index system (n.d.). http://www.fbi.gov/hq/lab/codis/index1.html

Coghlan A (9 Oct 2017) Kuwait's plans for mandatory DNA database have been cancelled. New Scientist, Accessed 2 Mar 2018

Combined DNA Index System (CODIS) (2018) Federal Bureau of Investigation. Accessed 8 May 2018

Commission, Australian Criminal Intelligence (18 July 2018) National Criminal Investigation DNA Database

Cotton EA, Allsop RF, Guest JL, Frazier RRE, Koumi P, Callow IP et al (2000) Validation of the AMPFlSTR® SGM Plus™ system for use in forensic casework. Forensic Sci Int 112 (2–3):151–161

CrimTrac (2014) Annual report. Commonwealth of Australia, Canberra. http://www.crimtrac.gov.au/about_us/CorporateDocuments.html

Curran T (1997) Forensic DNA analysis: technology and applications. In: Background paper BP-443E. edBibliothèque du parlement (Ottawa, ON: Parliamentary Research Branch, Library of Parliament), 43. Available from: http://www.parl.gc.ca/Content/LOP/researchpublications/bp443-e.htm

Curtis C, Hereward J (29 Aug 2017) From the crime scene to the courtroom: the journey of a DNA sample. The Conversation. Accessed 14 Oct 2017

Data and Communication Flow in CODIS (PDF) (2017) Sexual assault kit initiative. Accessed 9 Apr 2017

DNA Forensics (n.d.). http://www.dnaforensics.com/familialsearches.aspx

DNA Policy Initiative (n.d.-a). http://dnapolicyinitiative.org/wiki/index.php?title=Global_summary

DNA Policy Initiative (n.d.-b) France. http://dnapolicyinitiative.org/wiki/index.php?title=France. Accessed 26 Apr 2020

DNA Policy Initiative (n.d.-c). http://dnapolicyinitiative.org/wiki/index.php?title=China

FBI CODIS Core STR Loci (2015) National Institute of Standards and Technology, 26 Aug 2015. Accessed 29 Dec 2017

Federal Bureau of Investigation (2017) Frequently asked questions on CODIS and NDIS. Federal Bureau of Investigation. Accessed 6 Mar 2017

Ferreira ST, Paula KA, Maia FA, Svidizinski AE, Amaral MR, Diniz SA, Siqueira ME, Moraes AV (2015) The use of DNA database of biological evidence from sexual assaults in criminal

investigations: a successful experience in Brasília, Brazil. Forensic Sci Int Genet Suppl Ser 5: e595–e597. https://doi.org/10.1016/j.fsigss.2015.09.235

Field, D (3 Sept 2015) Kuwait's war on ISIS and DNA. Oxford University Press Blog. Accessed 10 Oct 2015

Forensics: Fingering Criminals Using DNA (2017) For dummies. Accessed 6 Mar 2017

GeneWatch UK—Germany (2016). genewatch.org. Accessed 29 Dec 2016

Genomics Medicine (n.d.). https://genomicsmed.ie/genomics-medicine-ireland-raises-40-million-series-funding/

Germany's DNA Database (2016) Accessed 29 Dec 2016

Gill P, Jeffreys AJ, Werrett DJ (1985) Forensic application of DNA 'fingerprints'. Nature 318:577e579. https://doi.org/10.1038/318577a0

Global Offender Data Base (2014) Tim Schellberg Gordon Thomas Honeywell Government Affairs Tbilisi, Georgia. Available from: http://www.slideplayer.com/slide/6660048/. Last cited 10 Oct 2017

Haas C, Voegeli P, Hess M, Kratzer A, Bär W (1 Apr 2006). A new legal basis and communication platform for the Swiss DNA database. In: International congress series. Progress in forensic genetics 11 proceedings of the 21st international ISFG congress held in Ponta Delgada, The Azores, Portugal between 13 and 16 Sept 2005, vol 1288, pp 734–736. https://doi.org/10.1016/j.ics.2005.11.040

Harbison SA, Hamilton JF, Walsh SJ (2001) The New Zealand DNA databank: its development and significance as a crime solving tool. Sci Justice 41(1):33–37. https://doi.org/10.1016/S1355-0306(01)71846-1

Hayhurst C (2016) Melanie road murder: how a remarkable 32-year-old murder cold case was solved. The Independent, 10 May 2016. Available from: http://www.independent.co.uk/news/uk/crime/melanie-road-murder-how-dna-collected-in-1984-solved-the-32-year-old-case-with-christopher-hampton-a7022056.html. Cited 20 June 2018

Hindmash R, Prainsack B (eds) (2010) Genetic suspects: global governance of forensic DNA profiling and databasing. Cambridge University Press, Cambridge, p 154. ISBN: 978-0521519434

Home Office (1989) DNA profiling in immigration casework. Report of a pilot trial by the Home Office and Foreign and Commonwealth Office. Home Office, London

Independent (n.d.-a). https://www.independent.co.uk/news/world/asia/china-expands-dna-testing-xinjiang-muslims-security-crackdown-human-rights-watch-a7739791.html

Independent (n.d.-b). https://www.independent.co.uk/news/world/africa/rwanda-dna-database-human-rights-privacy-a8832051.html

Institute of Environmental Science and Research Website (2013). www.esr.cri.nz

International Forensic Sciences (n.d.) DNA Labs India (DLI), Truth Labs and Bio-Axis DNA Research Centre (P) Limited

Interpol. Global DNA Profiling Survey Results (2016). https://www.interpol.int/INTERPOL-expertise/Forensics/DNA/Media-library/INTERPOL-Global-DNASurvey-Results-2016. Accessed 22 Feb 2019

Irish Times (n.d.-a). https://www.irishtimes.com/special-reports/r-d-ireland/genomics-exploring-new-horizons-1.3985051

Irish Times (n.d.-b). https://www.irishtimes.com/opinion/ireland-putting-profit-before-people-with-genomic-medicine-strategy-1.3944438

ISIL claims responsibility for Kuwait Shia mosque blast (2015) Al Jazeera, 27 June 2015. Accessed 10 Oct 2015

Jamaica Observer (2016) Italy creates national DNA database to enhance anti-terror fight. Jamaica Observer, 26 Mar 2016. Accessed 11 May 2017

Jeffreys A (2005) Genetic fingerprinting. Nat Med 11:1035–1039. https://doi.org/10.1038/nm1005-1035

Jeffreys AJ, Wilson V, Thein SL (1985a) Individual-specific "fingerprints" of human DNA. Nature 316:76e79. https://doi.org/10.1038/316076a0

Jeffreys AJ, Brookfield JFY, Semeonoff R (1985b) Positive identification of an immigration test-case using human DNA fingerprints. Nature 317:818e819. https://doi.org/10.1038/317818a0
Kumar S, Verma AK, Singh P, Singh R (2016). Current scenario of forensic DNA databases in or outside India and their relative risk. Egyptian J Forensic Sci 6(1):1–5.
Käppner J (2016) Justiz: VerräterischeProben (in German). SüddeutscheZeitung. Accessed 29 Dec 2016
Le Canard enchaîné (2009) "L'ADN fiche la trouille", article of Dominique Simonnot, mercredi 16 septembre 2009, p 4 (in French)
Le Monde (2007) A 8 et 11 ans, ilssontmenacés de fichagegénétique pour vol de jouetsdans Le Monde daté du 05 mai 2007 (in French)
Ley 78/2015 de Kuwait (consulted 5 Dec 2016). Available from: http://kuwaitalyawm.media.gov.kw/content/%D9%82%D8%A7%D9%86%D9%88%D9%86-78
Livemint (n.d.). https://www.livemint.com/Politics/hFVRzai11wlgloIlr6zuZP/Proposed-DNA-database-greatly-expands-scope-of-surveillance.html
Machado H, Silva S (2014) Would you accept having your DNA profile inserted in the National Forensic DNA database? Why? Results of a questionnaire applied in Portugal. Forensic Sci Int Genet 8:132–136. https://doi.org/10.1016/j.fsigen.2013.08.014
Maguire CN, McCallum LA, Storey C, Whitaker JP (2013) Familial searching: a specialist forensic DNA profiling service utilising the National DNA Database Òto identify unknown offenders via their relatives—the UK experience. Forensic Sci Int Genet 8:1–9. https://doi.org/10.1016/j.fsigen.2013.07.004
McCartney C (2006) The DNA expansion programme and criminal investigation. Br J Criminol 46:175e192. https://doi.org/10.1093/bjc/azi094
Milot E, Lecomte MM, Germain H, Crispino (2003) The National DNA Data Bank of Canada: a Quebecer perspective. Front Genet 4:249
Milot E, Marie L, Hugo G, Frank C (2013) The national DNA data bank of Canada: a Quebecer perspective. Front Genet 4:249
Minister for Justice Michael Keenan (5 Nov 2015) New super agency to tackle emerging threats (Press release). Accessed 9 Dec 2016
Mobbs JD (2001) Crimtrac-technology and detection. In: Fourth national outlook symposium on crime in Australia, New crimes or new responses, Canberra
National DNA Data Bank (2013) Archived 2013-06-25 at the Wayback Machine
National DNA Database Statistics, Q1 2015 to 2016 (2015) National DNA Database statistics. UK Government Home Office. Accessed 11 Oct 2015
National DNA Intelligence Databases in Europe (2016) Report on the current situation (PDF). Accessed 29 Dec 2016
Negri G (26 Mar 2016) Italy approves DNA database to fight crime. Il Sole 24 Ore, English edition. Accessed 11 May 2017
Newropeans Magazine (n.d.) The European perspective. Preparing for the world of tomorrow
Ope Letter Stop the police's DNA collection frenzy! (PDF) (2016) Accessed 29 Dec 2016
Ouvrez la bouche (2009) vousêtesfiché, article in Le Monde, of 18 Sept 2009
Peerenboom E (1998) Central criminal DNA database created in Germany. Nat Biotechnol 16:510–511. https://doi.org/10.1038/nbt0698-510
Police RCM (2012) The national DNA data bank of Canada. In: Annual Report (ed) Forensic science and identification services. Royal Canadian Mounted Police, Ottawa, ON, pp 1–30
Procedure for CODIS (2017) 10. North Carolina Department of Justice, 28 Dec 2015. Accessed 13 May 2017
Sanhati (n.d.). http://sanhati.com/excerpted/5753/
Schneider P (1998) Datenbankenmitgenetischen Merkmalen von Straftätern. Werdendurch die forensische DNA-Analyse persönlichkeitsrelevante Merkmale aufgedeckt, pp 330–333
Schultz S (2016) Stop the DNA collection frenzy!: expansion of Germany's DNA database. Forensic Genetics Policy Initiative. Accessed 29 Dec 2016
Skinner D (14 July 2010) Sociology 52: 13: Machado and Silva: Forensic DNA in Portugal

Sparkes R, Kimpton C, Gilbard S, Carne P, Andersen J, Oldroyd N et al (1996) The validation of a 7-locus multiplex STR test for use in forensic casework. II. Artefacts, casework studies and success rates. Int J Legal Med 109(4):195–204

Starling E (19 Mar 2018) Australian Criminal Intelligence Commission annual report 2016–17. Australian Criminal Intelligence Commission. Accessed 26 Nov 2018

Statewatch (n.d.). https://www.statewatch.org/news/2019/feb/uk-ndnad-growth.htm

Sutton M (14 Feb 2017) HH Sheikh Mohammed launches 10× initiative. ITP.net. Accessed 2 Mar 2018

The Biology Project (2000) What are the 13 core CODIS loci?, 27 Oct 2000. Accessed 6 Mar 2017—via The University of Arizona

The Evolution of Forensic DNA Database Policy (2017). Available from: https://www.thermofisher.com/content/dam/LifeTech/Documents/PDFs/3-Tim-Schellberg.pdf. Last cited 10 Oct 2017

The Hindu (2012) Bill to create bank for DNA profiling of accused coming. The Hindu, 21 Oct 2012. Accessed 5 Jan 2015

The Hindu (2020) DNA databases: 'India must learn from international experiences'. [online] available from: https://www.thehindu.com/news/cities/bangalore/dna-databases-india-must-learn-from-international-experiences/article4055382.ece. Accessed 25 Apr 2020

The New York Times (n.d.). https://www.nytimes.com/2019/02/21/business/china-xinjiang-uighur-dna-thermo-fisher.html

Treviño J (20 Mar 2018) Dubai wants to DNA test its millions of residents to prevent genetic disease. Smithsonian. Accessed 2 Mar 2018

Veiligheid MvJ (14 May 2013) Home—Nederlandse DNA-databank. dnadatabank.forensischinstituut.nl (in Dutch) Accessed 22 Sept 2019

Visser N (14 July 2015) Kuwait to Institute mandatory DNA testing for all residents. Huffington Post. Accessed 10 Oct 2015

Wallace H (2006) The UK national DNA database. EMBO Rep 7(S1):S26–S30. https://doi.org/10.1038/sj.embor.7400727

Williams R, Johnson P (2005) Inclusiveness, effectiveness and intrusiveness: issues in the developing uses of DNA profiling in support of criminal investigations. J Law Med Ethics 33:545–558. https://doi.org/10.1111/j.1748-720X.2005.tb00517.x

WSJ (n.d.). https://www.wsj.com/articles/china-snares-innocent-and-guilty-alike-to-build-worlds-biggest-dna-database-1514310353

Zamir A, Dell'Ariccia-Carmon A, Zaken N, Oz C (2012) The Israel DNA database—the establishment of a rapid, semi-automated analysis system. Forensic Sci Int Genet 6(2):286–289. https://doi.org/10.1016/j.fsigen.2011.06.003

Appendix A: SWGDAM Guidelines (Web Link)

1. Scientific Working Group on DNA Analysis Methods (SWGDAM) Guidelines:
   https://www.swgdam.org/
2. SWGDAM Interpretation Guidelines for Autosomal STR:
   https://1ecb9588-ea6f-4feb-971a-73265dbf079c.filesusr.com/ugd/4344b0_50e2749756a242528e6285a5bb478f4c.pdf
3. SWGDAM Contamination Prevention and Detection Guidelines for Forensic DNA Laboratories:
   https://1ecb9588-ea6f-4feb-971a-73265dbf079c.filesusr.com/ugd/4344b0_91f2b89538844575a9f51867def7be85.pdf
4. SWGDAM Guidelines for the Validation of Probabilistic Genotyping Systems:
   https://1ecb9588-ea6f-4feb-971a-73265dbf079c.filesusr.com/ugd/4344b0_22776006b67c4a32a5ffc04fe3b56515.pdf
5. SWGDAM Guidelines for the Collection and Serological Examination of Biological Evidence:
   https://1ecb9588-ea6f-4feb-971a-73265dbf079c.filesusr.com/ugd/4344b0_b3deba7a272b4b268d7f522840607410.pdf
6. SWGDAM Guidelines for STR Enhanced Detection Methods:
   http://media.wix.com/ugd/4344b0_29feed748e3742a5a7112467cccec8dd.pdf
7. SWGDAM Interpretation Guidelines for Y-Chromosome STR Testing:
   http://media.wix.com/ugd/4344b0_da25419ba2dd4363bc4e5e8fe7025882.pdf
8. SWGDAM Guidelines for Missing Persons Casework:
   http://media.wix.com/ugd/4344b0_da25419ba2dd4363bc4e5e8fe7025882.pdf
9. SWGDAM Guidelines for Missing Persons Casework:
   http://media.wix.com/ugd/4344b0_2ba78a46a2664b29948c60bc0aebc902.pdf
10. SWGDAM Training Guidelines
    SWGDAM Validation Guidelines for Forensic DNA Analysis Methods:
    https://1ecb9588-ea6f-4feb-971a-73265dbf079c.filesusr.com/ugd/4344b0_813b241e8944497e99b9c45b163b76bd.pdf
11. SWGDAM QAS Clarification Document:
    https://1ecb9588-ea6f-4feb-971a-73265dbf079c.filesusr.com/ugd/4344b0_c96fecf022bf4278b014941278d51ebc.pdf

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4

12. SWGDAM Mitochondrial DNA Analysis Interpretation Guidelines:
https://1ecb9588-ea6f-4feb-971a-73265dbf079c.filesusr.com/ugd/4344b0_f61de6abf3b94c52b28139bff600ae98.pdf
13. SWGDAM Mitochondrial DNA Nomenclature Examples:
https://1ecb9588-ea6f-4feb-971a-73265dbf079c.filesusr.com/ugd/4344b0_2044739c57574dbea97f2f85b6f73c9d.pdf
14. SWGDAM Recommendations for the Efficient DNA Processing of Sexual Assault Evidence Kits:
https://1ecb9588-ea6f-4feb-971a-73265dbf079c.filesusr.com/ugd/4344b0_4daf2bb5512b4e2582f895c4a133a0ed.pdf
15. SWGDAM Guidelines for the Validation of Probabilistic Genotyping Systems:
https://1ecb9588-ea6f-4feb-971a-73265dbf079c.filesusr.com/ugd/4344b0_22776006b67c4a32a5ffc04fe3b56515.pdf

Appendix B: International Society for Forensic Genetics (ISFG)

The International Society for Forensic Genetics is an international association promoting scientific knowledge in the field of genetic markers analyzed for forensic purposes. The ISFG has been founded in 1968 and represents more than 1100 members from over 60 countries. Regular meetings are held at regional and international level. Scientific recommendations on relevant forensic genetic issues are developed and published by expert commissions of the ISFG.

1. Web link of International *Society for Forensic Genetics (ISFG)*
   https://www.isfg.org/
2. Web link of ISFG *Publications*
   https://www.isfg.org/Publications
      https://www.isfg.org/Publications/FSI+Genetics
      https://www.isfg.org/Publications/DNA+Commission
      https://www.isfg.org/Publications/Paternity+Testing+Commission
      https://www.isfg.org/Publications/Congress+Proceedings
3. Web link of *ISFG software development site*
   https://www.isfg.org/Software
4. Web link of *ISFG working groups*
   https://www.isfg.org/Working%20Groups
5. Web link of *ISFG database links*
   https://www.isfg.org/Links
6. Web link of online database with a compilation of references on DNA transfer and persistence
   https://www.dropbox.com/sh/jf286hcollyerlf/AABQHfhGGiB03YdcJgpO4NUga?%20dl=0
7. *ISFG Recommendations*
   Time to time various recommendations have been framed by the International Society of Forensic Genetics (ISFG), earlier named as the International Society for Forensic Haemogenetics (ISFH)
      Recommendations of the DNA Commission (Source: www.isfg.org)
      Gill P, Hicks T, Butler JM, Connolly E, Gusmão L, Kokshoorn B, Morling N, Van O, Parson W, Prinz M, Schneider PM, Sijen T, Taylor D (2018) DNA commission of the ISFG: assessing the value of forensic biological evidence—

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4

guidelines highlighting the importance of propositions: Part I. Evaluation of DNA profiling comparisons given (sub-)source propositions. Forensic Sci Int Genet 36:189–202 (Gill et al. 2018)

Tillmar AO, Kling D, Butler JM, Parson W, Prinz M, Schneider PM, Egeland T, Gusmão L (2017) DNA Commission of the International Society for Forensic Genetics (ISFG): guidelines on the use of X-STRs in kinship analysis. Forensic Sci Int Genet 29:269–275 (Tillmar et al. 2017)

Coble MD, Buckleton J, Butler JM, Egeland T, Fimmers R, Gill P, Gusmão L, Guttman B, Krawczak M, Morling N, Parson W, Pinto N, Schneider PM, Sherry ST, Willuweit S, Prinz M (2016) DNA Commission of the International Society for Forensic Genetics: recommendations on the validation of software programs performing biostatistical calculations for forensic genetics applications. Forensic Sci Int Genet 25:191–197 (Coble et al. 2016)

Bodner M, Bastisch I, Butler JM, Fimmers R, Gill P, Gusmão L, Morling N, Phillips C, Prinz M, Schneider PM, Parson W (2016) Recommendations of the DNA Commission of the International Society for Forensic Genetics (ISFG) on quality control of autosomal Short Tandem Repeat allele frequency databasing (STRidER). Forensic Sci Int Genet 24:97–102 (Bodner et al. 2016)

Parson W, Ballard D, Budowle B, Butler JM, Gettings KB, Gill P, Gusmão L, Hares DR, Irwin JA, King JL, Knijff P, Morling N, Prinz M, Schneider PM, Neste C, Willuweit S, Phillips C (2016), Massively parallel sequencing of forensic STRs: Considerations of the DNA commission of the International Society for Forensic Genetics (ISFG) on minimal nomenclature requirements. Forensic Sci Int Genet 22:54–63 (Parson et al. 2016). Supplementary data available: Excel spreadsheet with complete STR locus information

Parson W, Gusmao L, Hares DR, Irwin JA, Mayr WR, Morling N, Pokorak E, Prinz M, Salas A, Schneider PM, Parsons TJ (2014) DNA Commission of the International Society for Forensic Genetics: revised and extended guidelines for mitochondrial DNA typing. Forensic Sci Int Genet 13:134–142 (Parson et al. 2014)

Gill P, Gusmão L, Haned H, Mayr WR, Morling N, Parson W, Prieto L, Prinz M, Schneider H, Schneider PM, Weir BS (2012) DNA commission of the International Society of Forensic Genetics: Recommendations on the evaluation of STR typing results that may include drop-out and/or drop-in using probabilistic methods. Forensic Sci Int Genet 6(6):679–688 (Gill et al. 2012). Supplementary data available: Excel spreadsheet for LR calculation including dropout and dropin events

Linacre A, Gusmão L, Hecht W, Hellmann AP, Mayr WR, Parson W, Prinz M, Schneider PM, Morling N (2011) ISFG: recommendations regarding the use of non-human (animal) DNA in forensic genetic investigations. Forensic Sci Int Genet 5(5):501–505 (Linacre et al. 2011)

Prinz M, Carracedo A, Mayr WR, Morling N, Parsons TJ, Sajantila A, Scheithauer R, Schmitter H, Schneider PM (2007) DNA Commission of the

International Society for Forensic Genetics (ISFG): recommendations regarding the role of forensic genetics for disaster victim identification (DVI). Forensic Sci Int Genet 1(1):3–12 (Prinz et al. 2007)

Gill P, Brenner CH, Buckleton JS, Carracedo A, Krawczak M, Mayr WR, Morling N, Prinz M, Schneider PM, Weir BS (2006) DNA commission of the International Society of Forensic Genetics: recommendations on the interpretation of mixtures. Forensic Sci Int 160:90–101 (Gill et al. 2006)

Gusmao L, Butler JM, Carracedo A, Gill P, Kayser M, Mayr WR, Morling N, Prinz M, Roewer L, Tyler-smith C, Schneider PM (2006). DNA Commission of the International Society of Forensic Genetics. DNA Commission of the International Society of Forensic Genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis. Forensic Sci Int 157:187–97 (Gusmao et al. 2006)

Gill P, Brenner C, Brinkmann B, Budowle B, Carracedo A, Jobling MA, De K, Kayser M, Krawczak M, Mayr WR, Morling N, Olaisen B, Pascali V, Prinz M, Roewer L, Schneider PM, Sajantila A, Tyler-smith C (2001). DNA Commission of the International Society of Forensic Genetics: recommendations on forensic analysis using Y-chromosome STRs. Forensic Sci Int124:5–10 (Gill et al. 2001)

Carracedo A, Bär W, Lincoln P, Mayr W, Morling N, Olaisen B, Schneider P, Budowle B, Brinkmann B, Gill P, Holland M, Tully G, Wilson M (2000) DNA Commission of the International Society for Forensic Genetics: guidelines for mitochondrial DNA typing. Forensic Sci Int 110(2):79–85 (Carracedo et al. 2000)

Bär W, Brinkmann B, Budowle B, Carracedo A, Gill P, Lincoln P, Mayr W, Olaisen B (1997) DNA recommendations. Further report of the DNA Commission of the ISFG regarding the use of short tandem repeat systems. Forensic Sci Int 87(3):179–174 (Bär et al. 1997)

Bär W, Brinkmann B, Lincoln P, Mayr W, Rossi U, Budowle B, Fourney R, Gill P, Rand S (1993). Editorial: Statement by DNA Commission of the International Society for Forensic Haemogenetics concerning the National Academy of Sciences report on DNA Technology in Forensic Science in the USA. Forensic Sci Int 59(1):1–2 (Bär et al. 1993)

Bär W, Brinkmann B, Lincoln P, Mayr W, Rossi U, Budowle B, Eisenberg A, Fourney R, Gill P, Rand S (1992) Editorial: Recommendations of the DNA Commission of the International Society for Forensic Haemogenetics relating to the use of PCR-based polymorphisms. Forensic Sci Int 55:1–3 (Bär et al. 1992)

Brinkmann B, Bütler R, Lincoln P, Mayr WR, Rossi U, Bär W, Baur M, Budowle B, Fimmers R, Gill P, Morris J, Rand S, Rittner C, Sensabaugh G (1992). Editorial: 1991 Report concerning recommendations of the DNA Commission of the International Society for Forensic Haemogenetics relating to the use of DNA polymorphisms. Forensic Sci Int 52:125–130 (Brinkmann et al. 1992)

Brinkmann B, Bütler R, Lincoln P, Mayr WR, Rossi U, Baur MP, Blin N, Driesel A, Epplen JT, Van-eyden P, Fimmers R, Henke J, Kömpf J, De-lange G, Martin W, Polesky H, Rand S, Rittner C, Schmitter H, Schneider PM, Werrett DJ, Zang KD (1989). Editorial: Recommendations of the Society for Forensic Haemogenetics concerning DNA polymorphisms. Forensic Sci Int 43:109–111 (Brinkmann et al. 1989)

Appendix C: Forensic Science Journals and Publications

1. Journal of Forensic Sciences:
   https://onlinelibrary.wiley.com/journal/15564029
2. Forensic Science International:
   https://www.journals.elsevier.com/forensic-science-international
3. Forensic Science International: Genetics:
   https://www.journals.elsevier.com/forensic-science-international-genetics
4. Canadian Society of Forensic Science Journal:
   https://www.tandfonline.com/loi/tcsf20
5. The American Journal of Forensic Medicine and Pathology:
   https://journals.lww.com/amjforensicmedicine/pages/default.aspx
6. Journal of Forensic and Legal Medicine:
   https://www.journals.elsevier.com/journal-of-forensic-and-legal-medicine
7. Egyptian Journal of Forensic Sciences:
   https://ejfs.springeropen.com/
8. Science & Justice Journal:
   https://www.journals.elsevier.com/science-and-justice
9. Medicine, Science and the Law:
   https://journals.sagepub.com/home/msl
10. IEEE Transactions on Information Forensics and Security:
    https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=10206
11. Australian Journal of Forensic Sciences:
    https://www.tandfonline.com/toc/tajf20/current
12. Journal of Forensic Sciences & Criminal Investigation (JFSCI):
    https://juniperpublishers.com/jfsci/
13. Forensic Science International Supplement Series:
    https://www.journals.elsevier.com/forensic-science-international-genetics-supplement-series

© Springer Nature Singapore Pte Ltd. 2020

P. Shrivastava et al. (eds.), *Forensic DNA Typing: Principles, Applications and Advancements*, https://doi.org/10.1007/978-981-15-6655-4