



Introduction to Forensic DNA Typing and Current Trends

2

Monika Chakravarty and Prateek Pandya

Contents

From Sampling to Identification	30
Improved Software Tools	31
Massively Parallel Sequencing/Next-Generation Sequencing: Advanced Human Identification	32
The NGS Technology	32
Comparing CE and NGS	33
Forensic Applications of NGS Technology	35
STR and SNP Sequencing	36
Mitochondrial DNA Sequencing in Forensics Using NGS	37
DNA Intelligence	38
Forensic DNA Phenotyping (FDP): Visible Phenotype Estimation	38
Ancestry Informative Markers	39
Degraded Samples	39
Identifying Monozygotic Twins	39
Emerging Applications of NGS	40
Utility of NGS/MPS over CE	40
Forensic Genetic Genealogy: The New Means of Genetic Identification	41
The Ekeby Man Case	43
The Golden State Killer	43
The Rapid DNA Instrument	43
Conclusion	45
References	45

M. Chakravarty
Forensic Science Laboratory, Home Department Government of NCT of Delhi, Delhi, India

P. Pandya (✉)
Amity Institute of Forensic Sciences, Amity University, Uttar Pradesh, India
e-mail: ppandya@amity.edu

Abstract

DNA profiling in forensic context has seen tremendous growth over the last almost four decades. Forensic DNA fingerprinting started with Alec Jefferey's single and multilocus probes via autoradiography followed by fluorescence-based PCR-STRs. Lately, there has been a significant change in technology for human identification. Chip-based massively parallel sequencing is the most recent advancement. Drawing investigative leads, disaster victim identification, and unsolved cold cases created a requirement for more efficient ways of extraction, rapid DNA testing, and genetic genealogy. Pooling forensic genetic genealogy, mitochondrial DNA for challenging sample analysis and STR-SNP sequencing has opened more avenues to generate the data leading to identification. This chapter highlights some of the advanced methods adopted in forensic DNA profiling and a few appropriate cases incorporating such methods.

Keywords

InDels · Massively parallel sequencing · Forensic DNA phenotyping · Biogeographic ancestry · Genetic genealogy · GEDmatch · Rapid DNA

From Sampling to Identification

Human identification is an essential component of forensic science. In the earlier times, the emphasis was on the phenotypic aspects of human body such as anthropometric measurements, color, and related features. However, since the advent of DNA technology, the field of human identification has grown significantly in terms of establishing individuality, relationships, and cultural and ethnic correlations. In addition, DNA technology has also provided various interesting and useful avenues such as establishing the genetic basis of various diseases through marker identification, evolutionary biology, gene therapy, mapping, subspecies level identification, habitat correlation, and related areas. The journey of DNA as an identification tool in forensics for the past three decades has been comparatively simple. Generally, a crime scene sample is profiled using existing methods and compared to a suspect or uploaded to a forensic DNA database of convicted offenders. Crime investigating laboratories mostly apply polymerase chain reaction (PCR) and fluorescence-based capillary electrophoresis (CE) to detect length variations in short tandem repeats (STRs). Although current capillary electrophoresis is the gold standard for analysis of forensic samples and is the method of choice for separating STRs for most forensic laboratories, the advent of massively parallel sequencing (MPS) or next-generation sequencing (NGS) has shown new avenues for detailed genetic analysis. NGS technology is evolving rapidly over the last decade and has proved

advantageous in challenging forensic samples, including mixtures, low copy number DNA, and degraded samples. NGS technology offered comprehensive results in case of mismatches observed in disputed paternity cases (Ma et al., 2016).

Improved Software Tools

Software tools for data collection emanating from capillary electrophoresis are an essential component of DNA profiling technology workflow. Advanced data collection and analysis software could affect data analysis in capillary electrophoresis. This could facilitate data processing by reducing off-scale data, thereby increasing laboratory's output. Stochastic effects like allele dropouts, allele drop in, sister allele imbalance, and stutters occur more often in low DNA samples and may lead to stutter amplification. Upgraded 3500 Data Collection Software v4.0.1 (User Bulletin Publication Number 100075298, Thermo Fisher Scientific) provides new features to reduce spectral pull ups. Gene Mapper™ v1.6, automated genotyping software (User Bulletin Publication Number 100073905, Thermo Fisher Scientific), cuts off threshold limit from 33,000 RFU (Relative Fluorescence Unit) to 66,000 RFU and enables accurate evaluation of peak heights to determine minor or low-level contributor. It reduces off-scale data and is useful both for routine casework and database samples. Improved software provides profile comparison feature, improved marker labeling along with stutter filters so that the stutter peaks are not exported. Recently, there have been integrated CE and NGS case management workflows having data concordance (Converge software, Thermo Fisher Scientific) in which the samples can be analyzed both for CE and NGS at the same platform. Improved software tools have population statistics included in the software itself; however, the databases can be customized according to the region. Variants can be accurately detected, and samples that have been run on the same sequencer can be compared. Fast, automated analysis of complex data and multiple data export options have streamlined the analysis. Some laboratories are also in the process of developing probabilistic genotype software to estimate the number of male and female contributors (Coble and Bright 2019).

CE Data Interpretation Improvement

- Pull ups reduced.
- Signal optimization across the capillaries.
- Reduced off-scale data.
- Incorporation of 6-dye chemistry.
- Automatic spatial calibration.
- RFID tracking.
- One reagent cartridge.
- Common array length and single polymer for all applications.
- Small bench top instruments, touch display.

Massively Parallel Sequencing/Next-Generation Sequencing: Advanced Human Identification

DNA sequencing consists of identifying base sequence of certain sections or entire length of DNA molecule. Sequencing technology has evolved tremendously since first-generation technology started in the 1960s (Heather and Chain 2016). Since then, the technology has improved significantly while jumping from first-generation (Maxam and Gilbert 1977; Sanger, Nicklen, and Coulson 1977) to second- (Hyman 1988; Shendure and Ji 2008) and third-generation tools (Niedringhaus et al. 2011). Conventional capillary electrophoresis has been the chosen method for forensic laboratories to identify the perpetrators of crime and to exonerate individuals (Chakravarty et al., 2019; Shrivastava et al., 2012). Molecular biology in association with population genetics principles paved the way for excellent human identification and construction of large number of DNA databases (Dixit et al. 2019; Srivastava et al. 2020). Short tandem repeats (Autosomal STRs, Y-STRs, X-STRs, and mini-STRs,) mitochondrial DNA, and single-nucleotide polymorphisms (SNPs) are the tools used for CE fragment analysis (Fig. 1).

With recent developments in next-generation sequencing (NGS), novel methods have been devised within the forensic community to make way for the investigative proceedings when CE-STR analysis fails to produce results. Several forensic laboratories are in the process of applying the MPS technology for the analysis of conventional STR markers and mitochondrial control DNA region and the possible uses of other DNA markers not frequent in casework such as next-generation STR kits, single-nucleotide polymorphisms (SNPs), insertion/deletion (InDel) markers, and mitochondrial DNA sequence (Phillips et al., 2007). With NGS we can generate a lot of data from a single sample, viz., autosomal, Y-STRs, X-STRs, identity SNPs, as well as phenotype and biogeography ancestry (Phillips, 2015) (Fig. 2).

The NGS Technology

With the initiation of Sanger sequencing method in the 1970s (Sanger et al. 1977), DNA sequencing technology has come a long way. Several genome projects have been completed using Sanger technology. However low throughput and high cost pose a limitation to its use in more complex genome analyses (Fullwood et al., 2009). The recently introduced NGS technology has superseded these problems and

Fig. 1 Current tools for routine, degraded, kinship, and sexual assault samples

Autosomal-STRs	Y-STRs
X-STRs	Mini-STRs
Mitochondrial DNA	SNPs

Fig. 2 NGS: Multiple markers, one amplification

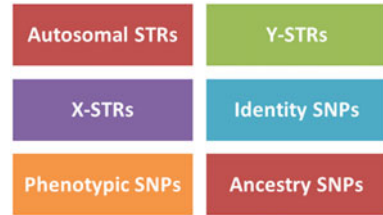
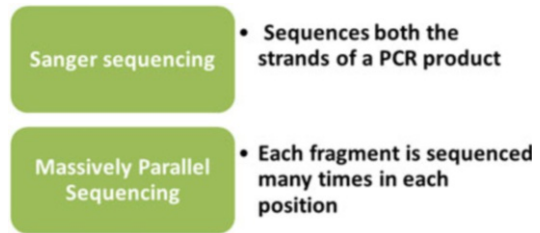


Fig. 3 Sanger sequencing versus massively parallel sequencing



is being used in forensics (Weber-Lehmann et al. 2014), diagnosing diseases (McCarthy et al., 2013), and ancient DNA analysis (Poinar et al. 2006). With Sanger sequencing, both the strands of a PCR product are sequenced, i.e., we read the sequence in one direction followed by reading in other direction, eventually sequencing each base twice. In MPS/NGS each base is sequenced many times yielding much more information about a sequence (Yang 2014). Billions of molecules can be sequenced in parallel, hence the name massively parallel sequencing (Fig. 3). Sequencing multiple reads at the same time leads to reduction in time as well as cost. Roche introduced the world's first high-throughput sequencing system utilizing pyrosequencing based on sequencing by synthesis in 2005 (Margulies et al. 2005) followed by the technologies offered by Thermo Fisher Scientific, Illumina, Ion Torrent Inc., and Pacific Biosciences (PacBio) to name a few. Since then, the NGS technology has been offering new horizons for forensic genetics.

Comparing CE and NGS

Extracted DNA in CE technology is PCR amplified using commercially available kits with fluorescent dye labeled primers to multiplex STRs with overlapping size ranges. PCR products are then separated through CE according to their molecular weight. The result is in the form of electropherogram comprising peaks with their sizes expressed as base pairs and height expressed in relative fluorescent units (Fig. 4) (Riman et al., 2020). NGS workflows (Fig. 5) also have PCR amplification to enrich STR markers. The target-specific primers contrarily are not fluorescently labeled but are taken for library construction. PCR products have adapters attached

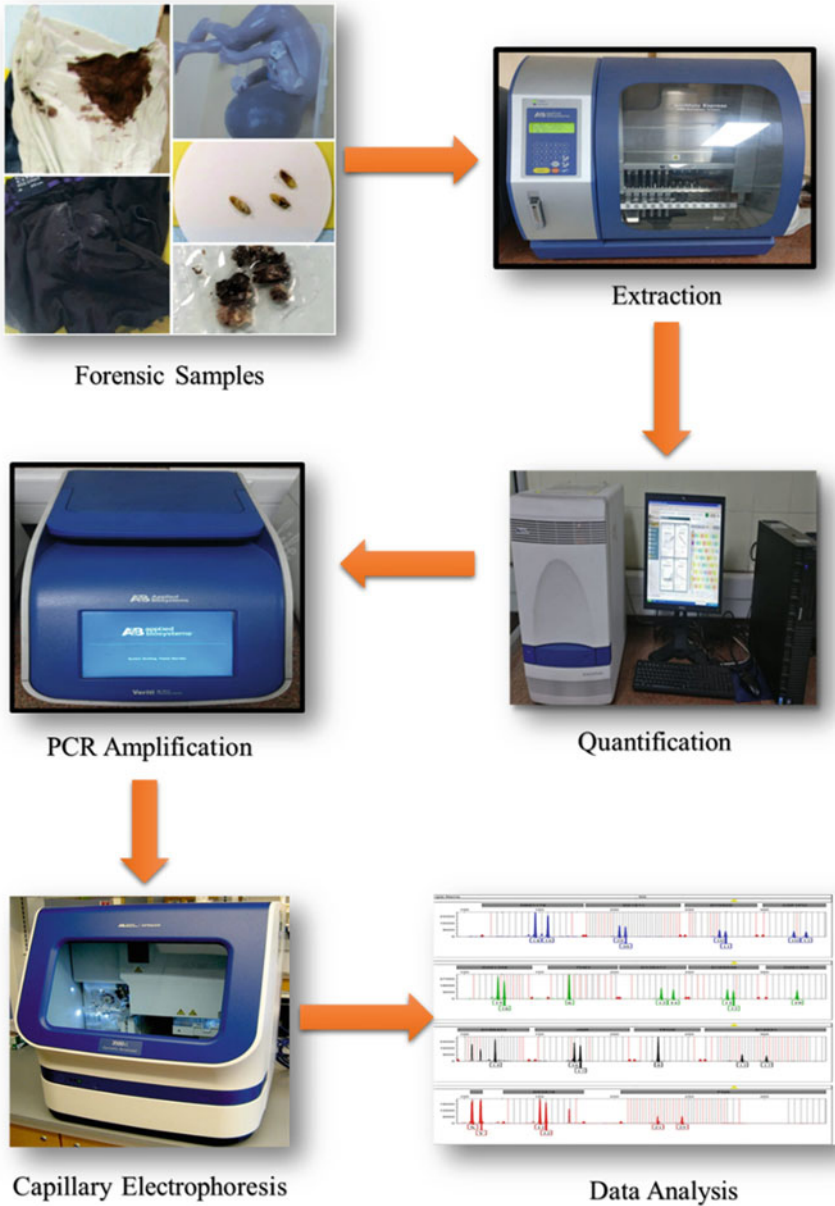


Fig. 4 Workflow of a traditional DNA Analysis

at both the ends producing DNA libraries which can be sequenced (Müller et al. 2018). Raw digital data of sequencing reads is obtained as a result of sequenced DNA libraries (Mardis 2017). Multiple libraries can be pooled into one reaction as

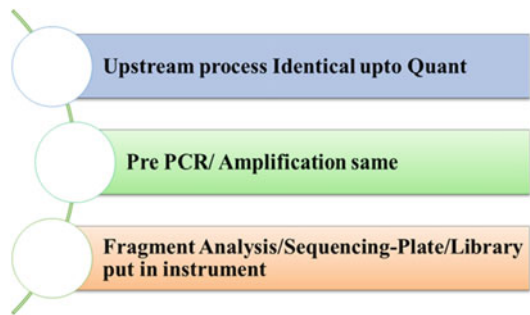
Fig. 5 NGS Workflow



Table 1 CE versus NGS

Parameters	Capillary electrophoresis	NGS/MPS
Variation	Length based	Sequence based
Primers	Fluorescent	Non-fluorescent
Random matching probability	Length-based discrimination	Sequence -based discrimination
Population studies	Uses allele frequencies	Uses frequencies for sequence-based alleles

Fig. 6 Similarities between CE and NGS/MPS methodology



the samples are barcoded. Available algorithms or software can determine the sequence and length-based polymorphisms (Woerner et al., 2017). Table 1 shows differences between CE and NGS technologies, whereas Figs. 6 and 7 depict the similarities between methodologies adopted in the processes.

Forensic Applications of NGS Technology

The application of massively parallel sequencing is increasing in forensic DNA analysis as the crime investigation laboratories are looking for methodologies to obtain maximum information from a trace or degraded forensic sample (Fig. 8) (Montano et al. 2018). Presently it is possible to map whole genomes with constantly increasing speed and decreasing costs (Børsting and Morling 2015). The true diversity in core forensic loci has been explored, thereby adding statistical weightage to the evidence.

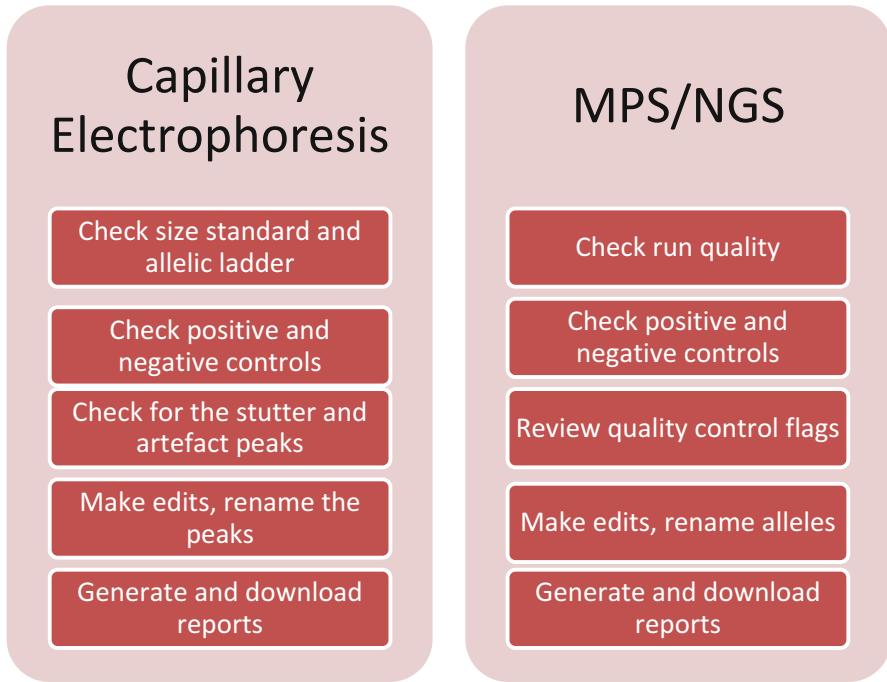


Fig. 7 Data analysis similarities between CE and NGS

STR and SNP Sequencing

CE explores length based, whereas NGS is sequence-based variation (Gettings et al. 2015). Multi-application sequencing has helped in achieving new heights. With NGS, it is possible to derive full sequence information (SNPs and InDels) within the STR loci to derive investigative leads and to estimate mutational events in kinship testing (Dalsgaard et al., 2013; Gettings et al., 2015). Insertion/deletion polymorphisms (InDels) which comprise the characteristics of both STRs and SNPs (Kidd et al. 2012) have now been used for forensic case work examination, databasing, and anthropological studies (Liu et al., 2020). SNPs and mitochondrial DNA (mtDNA) provide an effective accompaniment to traditional CE-STR analysis by increasing the amount and kind of genetic information that a single sample may yield. By putting together this information, a powerful investigative profile can be generated for use with missing persons and mass disaster victim identification (DVI) and to determine the number of contributors in a mixture (Petrovick et al. 2020) and cold cases. It is helpful particularly in those cases where traditional CE-based methods do not provide a full profile. Isometric alleles having identical size, but different sequence, could be identified. MPS offers high resolution as well as high power of discrimination, and degraded sample can still be used. Lots of sequence variation can be analyzed within an

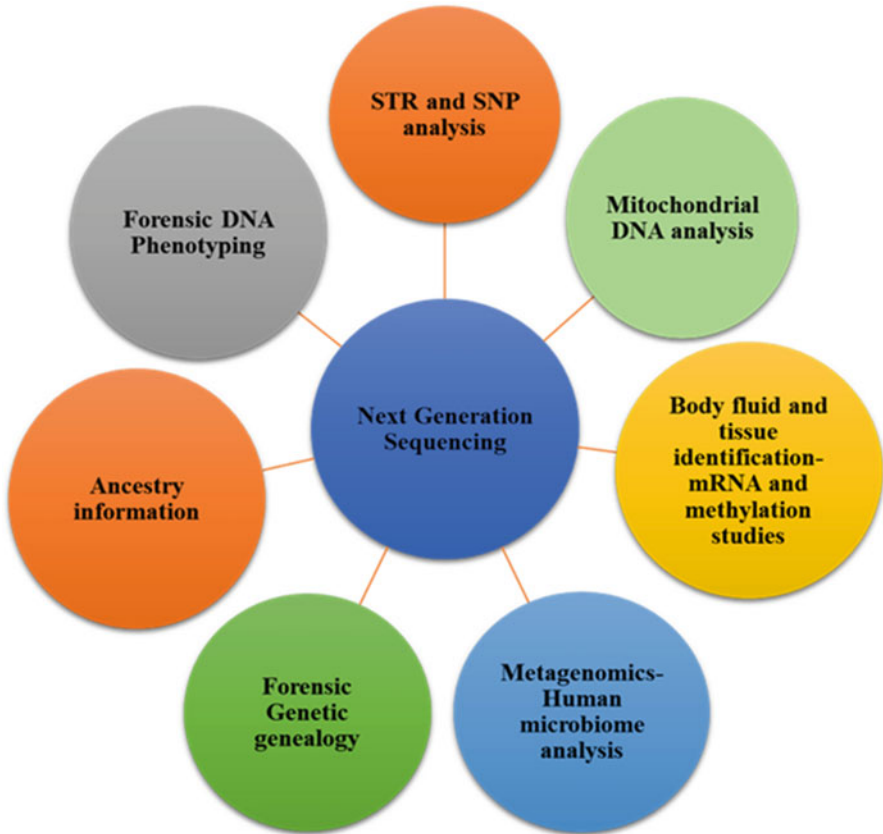


Fig. 8 NGS facilitated forensic tools

allele. NGS technology can target just a few genes to all the nucleotides in a whole genome (Guo et al., 2017). A large number of panels have been developed for major as well as admixed populations. Successfully reliable and responsible sequencing of autosomal as well as Y-STR loci leads to conviction of a rape suspect by a court in Amsterdam, Netherlands. This case worldwide was the first case where conviction was based on MPS.

Mitochondrial DNA Sequencing in Forensics Using NGS

Mitochondrial genome has higher copy numbers as compared to nuclear DNA and is less prone to degradation. Degraded samples could yield good results using mt DNA. Challenging samples are difficult to analyze with nuclear DNA; forensically fruitful information can be obtained by means of mitochondrial (mt) DNA (Holland & Parsons, 1999). It is the only source of information from samples like hair shafts

where nuclear DNA is generally depleted (Higuchi et al., 1988). Mt. DNA analysis is an indispensable tool in forensic DNA examination (Holland et al., 2019); Mt. DNA is important for lineage, ancient DNA applications, and maternal inheritance for bio-ancestry. Capillary electrophoresis-based Sanger sequencing has been the gold standard for the last few decades (Berglund et al., 2011). But the technique is time-consuming and laborious and can focus only a portion of mitochondrial genome, i.e., HVI and HVII. Detection of heteroplasmy and occurrence of two or more mitochondrial genotypes in a cell can increase the discrimination power of analysis to a great extent. Analyzing heteroplasmy with Sanger sequencing is exceedingly difficult. Introduction of NGS technologies has revolutionized the arena of genomics (Kircher et al., 2012;). Last few decades have witnessed the utility of varied NGS technologies in assessing mtDNA, viz., Roche's 454 (Payne et al. 2013; Illumina's GAII (Li et al. 2010), Illumina's HiSeq 2000 (Tang et al. 2013), and Ion Torrent's Personal Genome Machine (PGM) (Parson et al. 2013). MPS addresses the limitations of traditional sequencing and sequences the entire mtDNA in one reaction without consuming much sample. MPS has large multiplex panels which makes it feasible to sequence the entire mitochondrial genome. With MPS complete, mtDNA sequence can be deduced from hair samples that come across in forensic case work (Parson et al. 2015). For aged and ancient genetic material, quantification and estimation of the level of degradation is the correct approach. In the highly degraded sample where nuclear DNA profile is not possible, NGS/SNP and mitochondrial DNA is the best alternate method. Tsunami victims case where high temperature and moisture resulted in degradation, the Russian Czar family missing children, and missing Mexican students of 2014 where apparently the individuals were burnt and killed and the remains were not amenable to DNA typing were some of the cases solved using mitochondrial analysis.

DNA Intelligence

DNA intelligence or forensic DNA phenotyping (FDP) facilitates the prediction of bio-geographic ancestry and external visible characteristics of the donors of forensic samples (Fig. 9). This sort of intelligence administers priceless investigative leads in cases where database matches do not yield results from STR analysis (Phillips et al., 2019).

Forensic DNA Phenotyping (FDP): Visible Phenotype Estimation

STR analysis fails to identify a person whose profile is unknown to the analysts. The drawbacks of comparative STR analysis lead to the birth of a new field in forensic genetics, i.e., forensic DNA phenotyping (FDP) (Kayser and De Knijff; 2011). Forensic DNA Phenotyping implies to the prediction of human appearances from traits such as hair color, eye color, and skin color from unknown crime scene samples

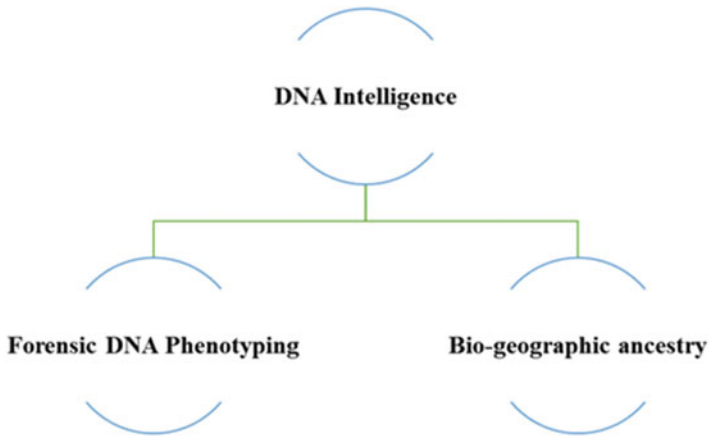


Fig. 9 DNA intelligence

(Kayser 2015). Investigative leads can be obtained from unknown sample donors, unidentified with the current CE technology. The technology can be of great help in disaster victim identification and missing person identification cases.

Ancestry Informative Markers

Bio-geographic ancestry (BGA) can be predicted using ancestry informative markers (AIM) (Xavier et al., 2020). During the course of the past few years, a large number of panels of the ancestry informative markers have been recommended for analyzing population genetic structure (Jiang et al. 2018). Also, ancestry can be established from the components of mixtures.

Degraded Samples

Conventional forensic analysis has a limitation of allele size with respect to degraded samples. To overcome this, shorter markers (MiniSTRs) were adopted (Martín et al., 2006). Currently SNPs have been incorporated in sequencing panels (Gettings et al., 2015). In CE, as a ski slope is obtained for a degraded sample, but NGS is not size based, so interpretation of results is easier, and also it also works with less amount of sample.

Identifying Monozygotic Twins

Monozygotic twins, having the same genetic structure, cannot be differentiated by conventional techniques like STR, SNP, and mitochondrial DNA analysis.

Identification of extremely rare mutations by NGS can differentiate between monozygotic twins (Weber-Lehmann et al., 2014).

Emerging Applications of NGS

Species origin of a sample, age range of contributors, metagenomics or human microbiome analysis, and methylation analysis are some of the emerging applications of NGS. Microbiome gives unique identity to an individual and can be used to determine the site of origin of the sample (Tozzo et al., 2020). NGS has a wide variety of application in body fluid and tissue identification via mRNA and methylation studies (Ingold et al. 2018). Studies have reported the use of miRNAs (micro-RNAs), a group of small noncoding RNAs having applicability in age prediction for body fluids or crime scene stains using MPS (Fang et al. 2020). NGS is also useful in wildlife forensics.

Utility of NGS/MPS over CE

The analysis of conventional STR markers using MPS provide a number of benefits over standard CE analysis, namely, particularly increased number of loci, higher discrimination power, and shorter amplicon length for a conclusive analysis of degraded and trace DNA evidence:

- Low DNA samples.
- Male minor contributor.
- Degraded samples.
- Unknown tissue origin.
- Cold cases.
- Suspect untraceable cases.
- Failed CE cases.
- No STR profile match in DNA database.
- Cold homicide cases.
- Mixture interpretation.
- Complex kinship cases or familial search.

The technique has certain drawbacks also, such as being expensive, lacking standardization, requiring huge resources for bioinformatics, and data storage besides accumulating sensitive personal data which is difficult to protect as entire genome including the coding regions can be sequenced. Although CE has a limitation of panels being limited by size ranges and fluorescent labels, still CE-based STR typing will remain the standard casework and database application as it is cheap, fast, and reliable and can be undertaken on regular basis, whereas MPS-based STR typing represents a specialized tool in forensics and can be used for specific cases. It involves time taking, tedious steps in library and DNA template preparation,

effective bioinformatics tools for sequence alignment, efficient data storage servers for the storage of sequence data files, and lack of standardization. The said shortcomings are being worked upon and do not seem to be a major long-term issue (Alonso et al. 2018). The technology guarantees endless advancement and may soon become the standard benchmark of quality. Today NGS platforms are being used not only in forensics (Fordyce et al. 2011) but also in microbial (Caporaso et al. 2012) and cancer research (Kirsch and Klein 2012).

Forensic Genetic Genealogy: The New Means of Genetic Identification

DNA analysis is being widely used in human identification, missing person/disaster victim identification and kinship testing (Fig. 10). A CE-STR based analysis involves uploading a crime scene profile to a database to obtain a hit or match (Fig. 11), whereas forensic genealogy is based on SNP testing and uploading the results to genealogy database to measure the genetic relatedness (Fig. 12).

Genetic genealogy has helped solve dozens of cold cases that would not have been cracked otherwise (Greytak et al., 2019). Matching is done by searching for DNA segments which are shared among the relatives. Genome-wide SNP data is used to measure the genetic relatedness (Fig. 13) (Kling and Tillmar 2019). Genetic genealogy though is not linked to CODIS but cases end the same way. In its evolutionary stage, genetic genealogy started with Y STR analysis. For the last 2 years, over 100 cases have been solved that have not been solved over decades. It has proved to be extremely accurate. Family trees have been constructed to find birth

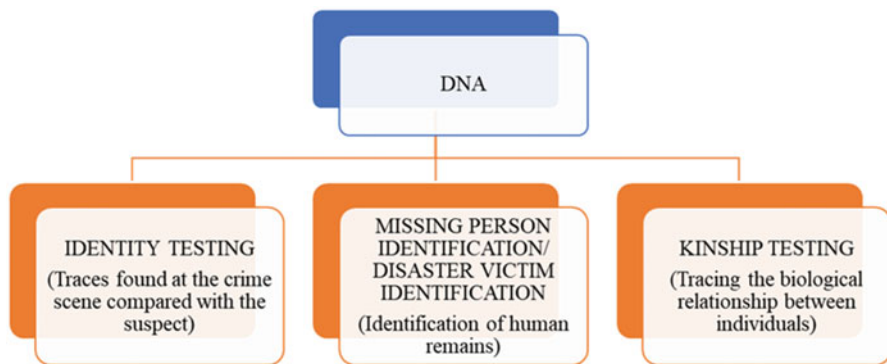


Fig. 10 Applications of DNA in legal field

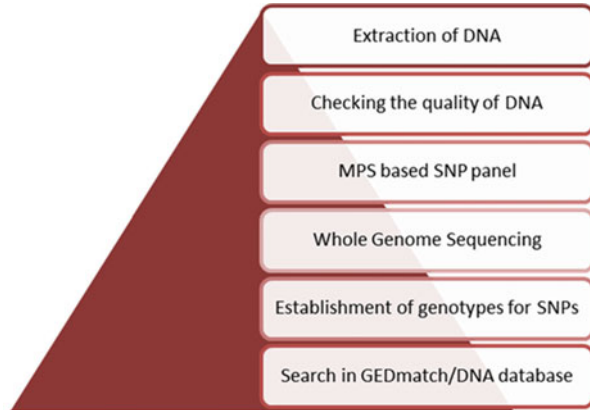
Fig. 11 Flow chart for CE-STR-based analysis



Fig. 12 Flow chart for forensic genetic genealogy



Fig. 13 From remains to database search: protocol for FGG



families by genealogists. Reference samples are compared with evidence material to confirm identity.

Cold cases, missing person identification, unsolved heinous crimes, and exoneration of the innocent can be solved by forensic genetic genealogy. Voluntarily submitted DNA data for forensic purposes is being used by “Genealogy Data Matching” (GEDmatch, Verogen, Inc.) software to generate investigative leads. GEDmatch allows the user to upload and compare SNP data in database. Autosomal profiles are uploaded to GEDmatch, to find out a potential match, and most recent common ancestors of a suspect were found on family tree, although there still exist issues in genetic genealogy relating to privacy notwithstanding the advancements (Court S. Denise, 2018; Wickenheiser, 2019). Moreover, the technique is expensive, time intensive, and resource intensive as is outsourced by forensic agencies and requires specialized knowledge. Low-quality uploads result in false genealogy connections. A cold case, a 16-year-old double homicide case, was solved by FGG in Sweden. SNP profiles were generated from the crime scene samples using whole genome sequencing and advanced bioinformatics tools. Relatives were searched in GEDmatch, and family tree DNA databases were constructed ultimately giving leads. With the arrival of IGG, alternate route for DNA analysis based on SNPs has been described. The following are the technical requirements for SNP testing:

- At least 1 ng of DNA.
- Single source DNA samples preferred.
- Degradation of samples could be a problem.

- DNA quality.
- Partial SNP profile.
- Success is genealogy database dependent.

Two most famous cases solved by this path breaking technique have been discussed below:

The Ekeby Man Case

The technique puts forward a great challenge for low copy number degraded DNA as in the case of forensic samples. Progression in DNA sequencing methodology has enabled the probability to process low copy number degraded biological samples (Prüfer et al., 2014). One such example is to identify an unknown male remains (“the Ekeby man case”) found murdered in Sweden in 2003, where whole genome sequencing was done on a bone sample along with bioinformatics tools which generated around 1.4 million SNPs (Tillmar et al. 2020). The SNP genotypes were searched for relatives on DNA database GED match. A list of relatives was prepared to identify the unknown remains, and investigative leads were obtained.

The Golden State Killer

The man known as the Golden State killer (De Angelo serial killer, 74 years now) was sentenced to multiple life sentences for dozens of crimes he committed. He was convicted for the rapes and killings which he committed from 1975 to 1986 covering a wide geographical area that was thought initially because of multiple people. He began with home burglaries before committing many rapes and killings across southern California. He often broke into people’s homes at midnight and carry out rapes and killings. The crimes mysteriously ended in 1986. The heinous crimes committed might have gone unsolved without the innovative genetic technique that was developed to reunite adoptees with their biological parents. DNA recovered from one of the crime scenes was put into the DNA database to find the relatives of the killer. A common ancestor among them was found, and family trees down to the present day were created (Phillips, 2018). De Angelo appeared as a possible suspect. Eventually an item containing De Angelo’s DNA was found by the investigators and compared with the DNA recovered from the crime scenes, and a match was obtained. He was arrested in April 2018. Dozens of victims testified before the court. He was sentenced to life imprisonment without parole on August 21, 2020.

The Rapid DNA Instrument

Generating DNA profile from capillary electrophoresis involves a large number of tedious steps. The setup required for performing STR analysis includes centrifuge machines, thermal cyclers, and capillary electrophoresis instrumentation in a

centralized laboratory consuming at least 10 hours. Advancements in instrumentation have led to automation during the last few years (Hopwood et al. 2010). Robotic platforms (Frégeau et al., 2010) have reduced hands-on time. The rapid DNA instruments combine the steps of isolation, faster amplification, denaturation, sizing, and genotyping (Fig. 14). RapidHIT[®] 200 and the RapidHIT ID Systems have been developed by IntegenX for generating STR profiles from reference samples in nearly 90 minutes. This also reduces the chance of contamination between the samples (Pleasanton, CA) (Shackleton et al., 2019). One could employ such kind of a unified instrument for reference samples with the least amount of time required by the analyst (Dash et al., 2020). This kind of technique not only reduces the risk of contamination but has the added advantage of swab reusability for conventional DNA methods. Such kind of instruments can be used by the law enforcement personnel in the police stations, in the crime investigation labs for increased lab productivity, and in the field as well according to the need.

The RapidHIT[™] ID system (Thermo Fisher Scientific) is a fully automated system for human identification. Run information is transferred to Rapidlink[™] software to process. Reagents used for each instrument can be traced at a central location. All the data is fed into computer, and any DNA profile can be matched. Quality flags/colored flags can tell about the run quality. One person at a central position can see all the data coming from all the instruments and can address the problem. Data quality is monitored at one workstation by the software. The system is cartridge based with two kinds of cartridges, viz., primary cartridge and sample cartridge:

1. Primary cartridge contains polymers, capillary, and buffers.
2. Sample cartridge are of two types the Ace GlobalFiler Express sample cartridge for single source reference samples and the Rapid INTEL[™] sample cartridge for forensic samples such as bones (Buscaino et al. 2018).

The developmental validation studies for this system were performed with mock casework samples according to the Scientific Working Group on DNA Analysis Methods (SWGDM) guidelines (Scientific Working Group on DNA Analysis

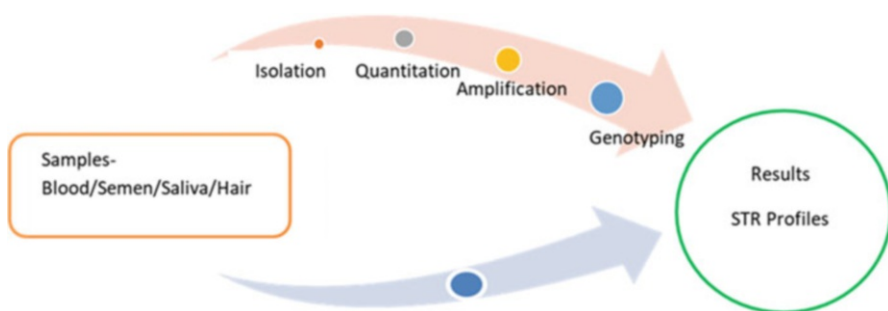


Fig. 14 Rapid hit system

Methods Validation Guidelines for DNA Analysis Methods. December 2016) (User Bulletin, “RapidINTEL™ Sample Cartridge for blood and saliva samples, Thermo Fisher Scientific). The samples are analyzed on the GeneMarker®HID STR Human Identity software on the instrument (Holland & Parson, 2011). Samples that require fast results may benefit from the rapid platform. Numerous Rapid Hit systems linked with RapidLink software can prove to be useful in criminal investigations.

Conclusion

Rapid DNA examination will soon facilitate new applications (Butler 2015). Massively parallel sequencing (MPS) provides the potential to multiplex diversified forensically relevant markers and multiple samples together in a single run in comparison with traditional capillary electrophoresis method (Churchill et al., 2016). Sequence based allele frequency data with the establishment of within STR allele sequence variants, will aid forensic community to enhance the power of discrimination for human identification and mixture deconvolution by increasing the effective allele number (Gaag et al. 2016) and kinship analysis. Forensic DNA phenotyping (FDP) and bio-geographic ancestry (BGA) from an unknown crime scene sample are gaining interest of the forensic community (Xavier et al., 2020). Whole genome sequencing has been successfully used to constitute genealogy DNA databases for generating investigative leads in cold cases (Tillmar et al. 2020). DNA intelligence marks a considerable peculiar application of genetic evidence unlike the one presented in the courtroom (Kayser 2015). NGS is an up-and-coming technology and is increasingly being implemented in forensic case work examination.

References

- Alonso A, Barrio PA, Müller P, Köcher S, Berger B, Martin P, Bodner M, Willuweit S, Parson W, Roewer L, Budowle B (2018) Current state-of-art of STR sequencing in forensic genetics. *Electrophoresis* 39;(21):2655–2668
- Berglund EC, Kiialainen A, Syvänen AC (2011) Next-generation sequencing technologies and applications for human genetic history and forensics. *Investig Genet* 2(1):23
- Børsting C, Morling N (2015) Next generation sequencing and its applications in forensic genetics. *Forensic Sci Int Genet* 18:78–89
- Buscaino J, Barican A, Farrales L, Goldman B, Klevenberg J, Kuhn M, Lin F, Nguyen P, Salceda S, Schuereen R, Smith C, Troup C, Tsou D, Vangbo M, King D (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
- Butler JM (2015) The future of forensic DNA analysis. *Philos Trans R Soc B Biol Sci* 370 (1674):20140252
- Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *Int Soc Microb Ecol* 6(8):1621–1624
- Chakravarty M, Pandya P, Raina A, Priyanka (2019) Resourcefulness of Str markers in forensic casework analysis. *Int J Med Toxicol Legal Med* 22(3–4):110–113

- Churchill JD, Schmedes SE, King JL, Budowle B (2016) Evaluation of the Illumina® Beta version ForenSeq™ DNA signature prep kit for use in genetic profiling. *Forensic Sci Int Genet* 20:20–29
- Coble MD, Bright JA (2019) Probabilistic genotyping software: an overview. *Forensic Sci Int Genet* 38:219–224
- Court S, Denise (2018) Forensic genealogy: some serious concerns. *Forensic Sci Int Genet* 36:203–204
- Dalsgaard S, Rockenbauer E, Gelardi C, Børsting C, Fordyce SL, Morling N (2013) Characterization of mutations and sequence variations in complex STR loci by second generation sequencing. *Genet Suppl Serie Forensic Sci Int* 4(1):e218–e219
- Dash HR, Pankaj S, Surajit D (2020) RAPID DNA technology: a boon to forensic DNA. *Typing* 313–316
- Dixit S, Pankaj S, Kumawat RK, Kaitholia K, Dash HR, Sharma H, Choubey G (2019) Forensic genetic analysis of population of Madhya Pradesh with PowerPlex fusion 6C™ multiplex system. *Int J Legal Med* 133(3):803–805
- Fang C, Xu L, Jing Z, Bingbing X, Jialin Q, Wenli I, Baoming L, Xiaoli Z, Huijuan W, Jiangwei Y (2020) Age estimation using bloodstain MiRNAs based on massive parallel sequencing and machine learning: a pilot study. *Forensic Sci Int Genet* 47:102300
- Fordyce SL, Ávila-Arcos MC, Rockenbauer E, Børsting C, Frank-Hansen R, Petersen FT, Willerslev E, Hansen AJ, Morling N, Gilbert TP (2011) High-throughput sequencing of Core STR loci for forensic genetic investigations using the Roche genome sequencer FLX platform. *BioTechnique* 51(2):127–133
- Frégeau CJ, Marc Lett C, Fournay RM (2010) Validation of a DNA IQ™-based extraction method for TECAN robotic liquid handling workstations for processing casework. *Forensic Sci Int Genet* 4(5):292–304
- Fullwood MJ, Wei CL, Liu ET, Ruan Y (2009) Next-generation DNA sequencing of paired-end tags (PET) for transcriptome and genome analyses. *Genome Res* 19(4):521–532
- Gaag VD, Kristiaan J, De Leeuw RH, Hoogenboom J, Patel J, Storts DR, Laros JFJ, De Knijff P (2016) Massively parallel sequencing of short tandem repeats – population data and mixture analysis results for the PowerSeq™ system. *Forensic Sci Int Genet* 24:86–96
- Gettings KB, Aponte RA, Kiesler KM, Vallone PM (2015a) The next dimension in STR sequencing: polymorphisms in flanking regions and their allelic associations. *Genet Suppl Serie Forensic Sci Int* 5:e121–e123
- Gettings KB, Kiesler KM, Vallone PM (2015b) Performance of a next generation sequencing SNP assay on degraded DNA. *Forensic Sci Int Genet* 19:1–9
- Greytak EM, Moore CC, Armentrout SL (2019) Genetic genealogy for cold case and active investigations. *Forensic Sci Int* 299:103–113
- Guo F, Yu J, Lu Z, Li J (2017) Massively parallel sequencing of forensic STRs and SNPs using the Illumina® ForenSeq™ DNA signature prep kit on the MiSeq FGx™ forensic genomics system. *Forensic Sci Int Genet* 31:135–141
- Heather JM, Chain B (2016) The sequence of sequencers: the history of sequencing DNA. *Genomics* 107(1):1–8
- Higuchi R, Von Beroldingen CH, Sensabaugh GF, Erlich HA (1988) DNA typing from single hairs. *Nature* 332(6164):543–546
- Holland MM, Parsons TJ (1999) Mitochondrial DNA sequence analysis – validation and use for forensic casework. *Forensic Sci Rev* 11(1):21–50
- Holland MM, Bonds RM, Holland CA, McElhoe JA (2019) Recovery of MtDNA from unfired metallic ammunition components with an assessment of sequence profile quality and DNA damage through MPS analysis. *Forensic Sci Int Genet* 39:86–96
- Holland MM, Parson W (2011) GeneMarker® HID: a reliable software tool for the analysis of forensic STR data. *J Forensic Sci* 56(1):29–35
- Hopwood AJ, Hurth C, Yang J, Cai Z, Moran N, Lee-Edghill JG, Nordquist A, Lenigk R, Estes MD, Haley JP, McAlister CR, Chen X, Brooks C, Smith S, Elliott K, Koumi P, Zenhausern F, Tully G (2010) Integrated microfluidic system for rapid forensic DNA analysis: sample collection to DNA profile. *Anal Chem* 82(16):6991–6999
- Hyman ED (1988) A new method of sequencing DNA. *Anal Biochem* 174(2):423–436

- Ingold S, Dørum G, Hanson E, Berti A, Branicki W, Brito P, Elsmore P, Gettings KB, Giangasparo F, Gross TE, Hansen S, Hanssen EN, Kampmann ML, Kayser M, Laurent FX, Morling N, Mosquera-Miguel A, Parson W, Phillips C, Porto MJ, Pośpiech E, Roeder AD, Schneider PM, Schulze Johann K, Steffen CR, Syndercombe-Court D, Trautmann M, van den Berge M, van der Gaag KJ, Vannier J, Verdoliva V, Vidaki A, Xavier C, Ballantyne J, Haas C (2018) Body fluid identification using a targeted mRNA massively parallel sequencing approach – results of a EUROFORGEN/EDNAP collaborative exercise. *Forensic Sci Int Genet* 34:105–115
- Jiang L, Wei YL, Zhao L, Li N, Liu T, Liu HB, Ren LJ, Li JL, Hao HF, Li Q, Li CX (2018) Global analysis of population stratification using a smart panel of 27 continental ancestry-informative SNPs. *Forensic Sci Int Genet* 35:e10–e12
- 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 Knijff P (2011) Improving human forensics through advances in genetics, genomics and molecular biology. *Nature Rev Genet* 12(3):179–192
- Kidd KK, Kidd JR, Speed WC, Fang R, Furtado MR, Hyland FCL, Pakstis AJ (2012) Expanding data and resources for forensic use of SNPs in individual identification. *Forensic Sci Int Genet* 6(5):646–652
- Kircher M, Sawyer S, Meyer M (2012) Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res* 40(1):e3–e3
- Kirsch S, Klein CA (2012) Sequence error storms and the landscape of mutations in Cancer. *Proc Nat Acad Sci U S A* 109(36):14289–14290
- Kling D, Tillmar A (2019) Forensic genealogy—a comparison of methods to infer distant relationships based on dense SNP data. *Forensic Sci Int Genet* 42:113–124
- Li M, Schönberg A, Schaefer M, Schroeder R, Nasidze I, Stoneking M (2010) Detecting Heteroplasmy from high-throughput sequencing of complete human mitochondrial DNA genomes. *Am J Hum Genet* 87(2):237–249
- Liu J, Weian D, Wang M, Liu C, Wang S, He G, Zheng W (2020) Forensic features, genetic diversity and structure analysis of three Chinese populations using 47 autosomal InDels. *Forensic Sci Int Genet* 45:102227
- Ma Y, Kuang JZ, Nie TG, Zhu W, Yang Z (2016) Next generation sequencing: improved resolution for paternal/maternal duos analysis. *Forensic Sci Int Genet* 24:83–85
- Mardis ER (2017) DNA sequencing technologies: 2006–2016. *Nat Protoc* 12(2):213–218
- Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Lei D, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer MLI, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM (2005) Genome sequencing in microfabricated high-density Picolitre reactors. *Nature* 437(7057):376–380
- Martín P, García O, Albarrán C, García P, Alonso A (2006) Application of mini-STR loci to severely degraded casework samples. *Int Cong Series* 1288:522–525
- Maxam AM, Gilbert W (1977) A new method for sequencing DNA. *Proc Nat Acad Sci* 74(2):560–564
- McCarthy JJ, McLeod HL, Ginsburg GS (2013) Genomic medicine: a decade of successes, challenges, and opportunities. *Sci Trans Med* 5(189):189sr4–189sr4
- Montano EA, Bush JM, Garver AM, Larijani MM, Wiechman SM, Baker CH, Wilson MR, Guerrieri RA, Benzinger EA, Gehres DN, Dickens ML (2018) Optimization of the Promega PowerSeq™ auto/Y system for efficient integration within a forensic DNA laboratory. *Forensic Sci Int Genet* 32:26–32
- Müller P, Alonso A, Barrio PA, Berger B, Bodner M, Martin P, Parson W (2018) Systematic evaluation of the early access applied Biosystems precision ID Globalfiler mixture ID and Globalfiler NGS STR panels for the ion S5 system. *Forensic Sci Int Genet* 36:95–103

- Niedringhaus TP, Milanova D, Kerby MB, Snyder MP, Barron AE (2011) Landscape of next-generation sequencing technologies. *Anal Chem* 83(12):4327–4341
- 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
- Parson W, Strobl C, Huber G, Zimmermann B, Gomes SM, Souto L, Fendt L, Delpont R, Langit R, Wootton S, Lagacé 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
- Payne BAI, Wilson IJ, Yu-Wai-Man P, Coxhead J, Deehan D, Horvath R, Taylor RW, Samuels DC, Santibanez-Koref M, Chinnery PF (2013) Universal Heteroplasmy of human mitochondrial DNA. *Hum Mol Genet* 22(2):384–390
- Petrovick MS, Boettcher T, Fremont-Smith P, Peragallo C, Ricke DO, Watkins J, Schwoebel E (2020) Analysis of complex DNA mixtures using massively parallel sequencing of SNPs with low minor allele frequencies. *Forensic Sci Int Genet* 46:102234
- Phillips C, McNevin D, Kidd KK, Lagacé R, Wootton S, de la Puente M, Freire-Aradas A, Mosquera-Miguel A, Eduardoff M, Gross T, Dagostino L, Power D, Olson S, Hashiyada M, Oz C, Parson W, Schneider PM, Lareu MV, Daniel R (2019) MAPlex – A massively parallel sequencing ancestry analysis multiplex for Asia-Pacific populations. *Forensic Sci Int Genet* 42:213–226
- Phillips C, Salas A, Sánchez JJ, Fondevila M, Gómez-Tato A, Álvarez-Dios J, Calaza M, Casares de Cal M, Ballard D, Lareu MV, Carracedo Á (2007) Inferring ancestral origin using a single multiplex assay of ancestry-informative marker SNPs. *Forensic Sci Int Genet* 1(3–4):273–280
- Phillips C (2015) Forensic genetic analysis of bio-geographical ancestry. *Forensic Sci Int Genet* 18:49–65
- Phillips C (2018) The Golden state killer investigation and the nascent field of forensic genealogy. *Forensic Sci Int Genet* 36:186–188
- Poinar HN, Schwarz C, Qi J, Shapiro B, MacPhee RDE, Buigues B, Tikhonov A, Huson DM, Tomsho LP, Auch A, Rampp M, Miller M, Schuster SC (2006) Metagenomics to Paleogenomics: large-scale sequencing of mammoth DNA. *Science* 311(5759):392–394
- Prüfer K, Racimo F, Patterson N, Jay F, Sankararaman S, Sawyer S, Heinze A, Renaud G, Sudmant PH, De Filippo C, Li H, Mallick S, Dannemann M, Qiaomei F, Kircher M, Kuhlwilm M, Lachmann M, Meyer M, Ongyerth M, Siebauer M, Theunert C, Tandon A, Moorjani P, Pickrell J, Mullikin JC, Vohr SH, Green RE, Hellmann I, Johnson PLF, Blanche H, Cann H, Kitzman JO, Shendure J, Eichler EE, Lein ES, Bakken TE, Golovanova LV, Doronichev VB, Shunkov MV, Derevianko AP, Viola B, Slatkin M, Reich D, Kelso J, Pääbo S (2014) The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* 505(7481):43–49
- Riman S, Iyer H, Borsuk LA, Vallone PM (2020) Understanding the characteristics of sequence-based single-source DNA profiles. *Forensic Sci Int Genet* 44:102192
- Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with Chain-terminating inhibitors. *Proc Nat Acad Sci U S A* 74(12):5463–5467
- Shackleton D, Pagram J, Andrews N, Malsom S, Ives L, Vanhinsbergh D (2019) Development of enhanced sensitivity protocols on the RapidHIT™ 200 with a view to processing casework material. *Sci Justice* 59(4):411–417
- Shendure J, Ji H (2008) Next-generation DNA sequencing. *Nat Biotechnol* 26(10):1135–1145
- Shrivastava P, Trivedi VB, Singh AK, Mishra N (2012) Application of DNA fingerprinting technology in Forensic Investigation. *Int J Scient Res Publ* 2(1):2250–3153
- Srivastava V, Prabhakar Rao K, Surekha Rani H, Kumawat RK, Mishra A, Shrivastava P (2020) Genomic diversity in the Goud population of Telangana, India inferred using twenty-three autosomal marker PowerPlex® fusion 6C system. *Meta Gene* 25:100718
- 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 LJC (2013) Transition to next generation analysis of the whole mitochondrial genome: a summary of molecular defects. *Hum Mutat* 34(6):882–893

- Tillmar A, Sjölund P, Lundqvist B, Klippmark T, Älgenäs C, Green H (2020) Whole-genome sequencing of human remains to enable genealogy DNA database searches – a case report. *Forensic Sci Int Genet* 46:102233
- Tozzo P, D'angiolella G, Brun P, Castagliuolo I, Gino S, Caenazzo L (2020) Skin microbiome analysis for forensic human identification: what do we know so far? *Microorganisms* 8(6):873
- Weber-Lehmann J, Schilling E, Gradl G, Richter DC, Wiehler J, Rolf B (2014) Finding the needle in the haystack: differentiating 'identical' twins in paternity testing and forensics by ultra-deep next generation sequencing. *Forensic Sci Int Genet* 9:42–46
- Wickenheiser RA (2019) Forensic genealogy, bioethics and the Golden state killer case. *Forensic Sci Int Synergy* 1:114–125
- Woerner AE, King JL, Budowle B (2017) Fast STR allele identification with STRait razor 3.0. *Forensic Sci Int Genet* 30:18–23
- Xavier C, de la Puente M, Phillips C, Eduardoff M, Heidegger A, Mosquera-Miguel A, Freire-Aradas A, Lagace R, Wootton S, Power D, Parson W, Lareu MV, Daniel R (2020b) Forensic evaluation of the Asia Pacific ancestry-informative MAPlex assay. *Forensic Sci Int Genet* 48:102344
- Xavier C, de la Puente M, Mosquera-Miguel A, Freire-Aradas A, Kalamara V, Vidaki A, Gross TE, Revoir A, Pośpiech E, Kartasińska E, Spólnicka M, Branicki W, Ames CE, Schneider PM, Hohoff C, Kayser M, Phillips C, Parson W (2020a) Development and validation of the VISAGE AmpliSeq basic tool to predict appearance and ancestry from DNA. *Forensic Sci Int Genet* 48:102336
- Yang Y, Xie B, Yan J (2014) Application of next-generation sequencing Technology in Forensic Science. *Genom Proteom Bioinform* 12(5):190–197