InDel Loci in Forensic DNA Analysis

47

Tugba Unsal Sapan

Contents

Abstract

As a part of forensic science, in the investigations of forensic cases, such as murder, sexual assault, and theft, DNA profile, which is obtained from biological materials found at the crime scene, is used to determine whether there is a connection between the suspect people and the offense. Depending on the type of case, autosomal or gonosomal polymorphism, which is obtained from X and Y chromosome, can be used in solving forensic cases. Although STR (Short Tandem Repeats) loci is still widely used in routine forensic genetic analysis, in many cases, typing problems can occur in highly degraded biological samples collected from the crime scene, and results cannot be obtained. In order to solve these problems, forensic scientists have been looking for alternative genetic markers.

InDels, which have been used in forensic science in recent years, can achieve more successful results in forensic identification when used together with STR and SNP loci. InDels are the most common type of polymorphism in the genome

T. Unsal Sapan (✉)
Institute of Addiction and Forensic Sciences, Uskudar University, Istanbul, Turkey
e-mail: tugba.unsalsapan@uskudar.edu.tr

H. R. Dash et al. (eds.), *Handbook of DNA Profiling*,
https://doi.org/10.1007/978-981-16-4318-7_49

after SNPs. InDel occurs as a result of insertion and/or deletion of one or more nucleotides, and it is a kind of polymorphism which can be used in population studies since insertion/deletion is seen with a frequency of more than 1% of the population. These loci can be used for identification, biogeographic genealogical research analysis, evolutionary research, and revealing kinship relationships since they allow to do multiplexing PCR study and have high heterozygosity rate and also small amplicon lengths (60–200 bp). Research on the use of InDels in forensic science has gained momentum in recent years. The number of ready-made commercial kits for these loci is quite limited. For this reason, researchers create their own InDel multiplex panels to use in their studies.

Keywords

Forensic biology and genetics · Human identification · Insertion/deletion (InDel) polymorphism · InDel loci

Introduction

Advances in molecular genetics in the 1980s allowed the study of polymorphic traits directly at the DNA level. The use of DNA in forensic science developed rapidly after Alec Jeffrey's discovery of polymorphic repeat sequences in the DNA molecule in 1985. Since then, the identification of biological samples collected from both individuals and from the crime scenes in the determination of paternity-kinship relations and other criminal investigations has carried out using DNA analysis techniques. In forensic cases, such as terrorism, murder, sexual assault, and theft, it is possible to link up between the suspect and the crime scene by using DNA profiles, which are isolated from biological materials detected at the crime scene (Jeffreys et al. 1985; Robertson et al. 1990; Chan 1992). Although VNTR (variable number of tandem repeats) loci, which were used in the first period of DNA analysis in forensic sciences, have a high discrimination power, it has been replaced by new technologies on account of the need of good qualified (non-fragmented) and excessive amounts (300–500 ng) of DNA, long and laborious analysis times, exposing radioactive materials, and so on. For the last 20 years, STR loci have been widely used in forensic identification (Lee et al. 1994; Robertson et al. 2002).

Short tandem repeats (STR) loci have been used as ideal genetic markers in forensic science in recent years due to their small amplicon size, successful results in degraded biological samples, capability to do multiplex analysis, and no requirement of expensive equipment (Weber and May 1989; Edward et al. 1992). Although STR loci have been still used today, typing problems are experienced in extremely degraded biological samples, which are collected from the crime scene, and successful results for comparison cannot be obtained many times. In order to solve this problem, mini short tandem repeats (miniSTRs) that allow typing even in degraded samples were put on the market in the early 2000s (Coble and Butler 2005). In recent years, for the same purpose, researchers have been focused on

different DNA polymorphisms, such as single-nucleotide polymorphism (SNP), which takes up a very little space in DNA, and InDel polymorphism (Pereira et al. 2009a, b).

Insertion

It is a type of mutation that disrupts the natural sequence as a result of addition of one or more bases to the DNA base sequence. This addition can be a base or as much as a whole chromosome (Fig. 1). As a result of DNA polymerase shift, it is usually thought to be formed by addition of base sequences, which are not adjacent to microsatellite sites, to the main sequence (Gelbart et al. 2002; Kondrashov and Rogozin 2004; Rodriguez-Murillo and Salem 2013).

Deletion

It is a type of mutation that occurs as a result of the deletion of one or more bases in the DNA sequence. Deletion, as well as insertion, can occur on one or more bases, as well as in chromosome size (Fig. 2). If a deletion occurs when part of the

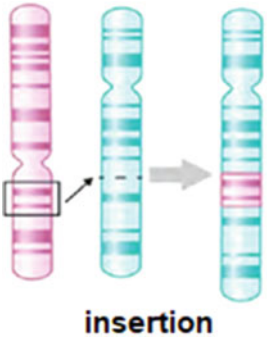

Fig. 1 Insertion on a chromosome level (Gordon and Egner 2013)

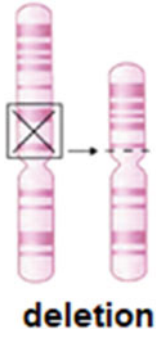

Fig. 2 Deletion on a chromosome level. (Gordon and Egner 2013)

chromosome breaks off and disappears, it can lead to serious genetic diseases (Gelbart et al. 2002; Kondrashov and Rogozin 2004; Rodriguez-Murillo and Salem 2013).

InDel (Insertion/Deletion)

Point or gene mutations are defined as changes that occur in the DNA sequence and can be passed on to subsequent generations. These mutations usually occur in one or more nucleotides and cause changes in the structure of the genome. Point mutations can also occur in the form of insertion or deletion (Campbell et al. 2006) (Fig. 3). However, insertion or deletion mutations are much more important changes than point mutations. Because insertion or deletion of one or more bases in the DNA chain usually leads to the shift of the genetic code that starts from the point where the insertion or deletion occurs, this causes polymorphism by creating significant changes in the structure of the gene (Pereira et al. 2009a, b).

InDels and somatic/gonosomal chromosome mutations are terms used to express mutation combinations, which include deletion or insertion separately or together, in the studies of forensic molecular genetics, evolution, and population genetics (Gelbart et al. 2002; Kondrashov and Rogozin 2004; Gregory 2004).

InDels, which account for 16–25% of all genetic variations in the genome, are the most commonly seen DNA polymorphism after SNPs, and 1.6–2.5 million InDel polymorphisms have been identified in human population studies. This, in turn, suggests that hereditary changes that occur as a result of insertion or deletion (InDel) mutations can be used as a genetic marker since they are seen in the human genome frequently. Even though they are so common, studies with InDels are limited (Hongbao 2005). However, since InDels have power of discrimination and heterozygosity, it is possible to use InDels in studies of human identification, ancestry, and evolutionary and molecular anthropology. For this reason, today, it

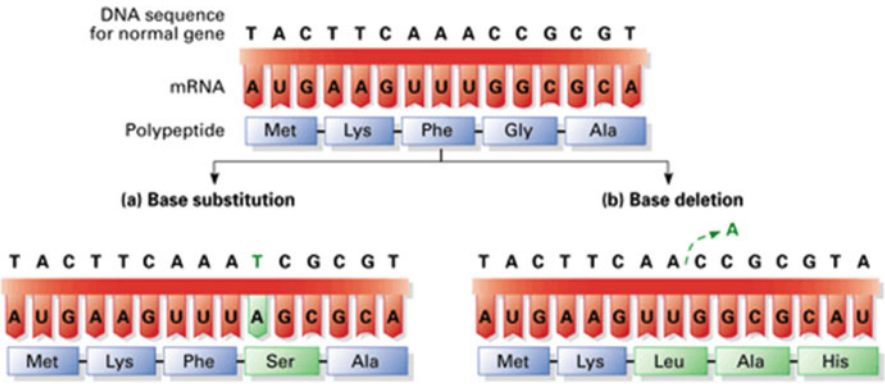

Fig. 3 Schematic representation of insertion and deletion on the DNA sequencing. (Gordon and Egner 2013)

is an alternative polymorphism to SNPs and STRs, which are used in DNA profiling in forensic identification (Reiner et al. 2005; Pereira et al. 2009a, b, 2012; Martínez-Cortés et al. 2015).

Use of Insertion/Deletion Polymorphism in Forensic Genetics

Insertion/deletion polymorphism (InDel) is known to be as a length polymorphism, which formed as a result of insertion and/or deletion of one or more nucleotides on the genome. Differences are used to discrimination of two people from each other in forensic identification, which makes polymorphisms the basis of identification (Rodriguez-Murillo and Salem 2013).

Polymorphism occurs as a result of successive mutations and is passed down from generation to generation according to Mendelian laws. In 2002, James Weber and colleagues identified over 2000 biallelic insertion/deletion polymorphisms on the human genome. Thus, studies of InDel polymorphism began for the first time in forensic sciences (Weber and May 1989). InDel loci have been used in forensic identification since they are able to study with multiplex PCR and have high heterozygosity and small amplicon lengths (60–200 bp) (Fig. 4). They can also be used to identify disaster victims in mass fatality cases, such as aircraft accidents, terrorist attacks or natural disasters, ancestry determination, evolutionary research, and molecular anthropology (Manta et al. 2012; Martínez-Cortés et al. 2015). More successful results can be achieved in identification by using InDel polymorphism, which are called next-generation genetic variations together with STR and SNP loci (Sanchez et al. 2006; Pereira et al. 2012).

InDel loci have been used in forensic sciences, especially in recent years. Commercial kit production for InDel loci is highly restricted. So forensic scientists often create their own InDel panels and try to popularize them (Guangyao et al. 2015).

Commercial Kits Based on InDel Analysis

InDel loci have been used in forensic science in recent years. However, InDel kit production is also quite limited. Currently, there are three commercial InDel kits made for the use in forensic science.

Investigator® DIPplex Kit

It is a kit containing 30 InDel loci found in somatic chromosomes that have been marketed by QIAGEN firm in the last 2–3 years. The kit also includes the locus of amelogenin. Selected InDel loci are smaller than 160 bp and are specifically designed for the use in the identification or anthropological research. In addition, they are analyzed by ABI 310, 3130, 3130XL, 3500, and 3500XL capillary electrophoresis (Investigator® DIPplex Handbook 2014).

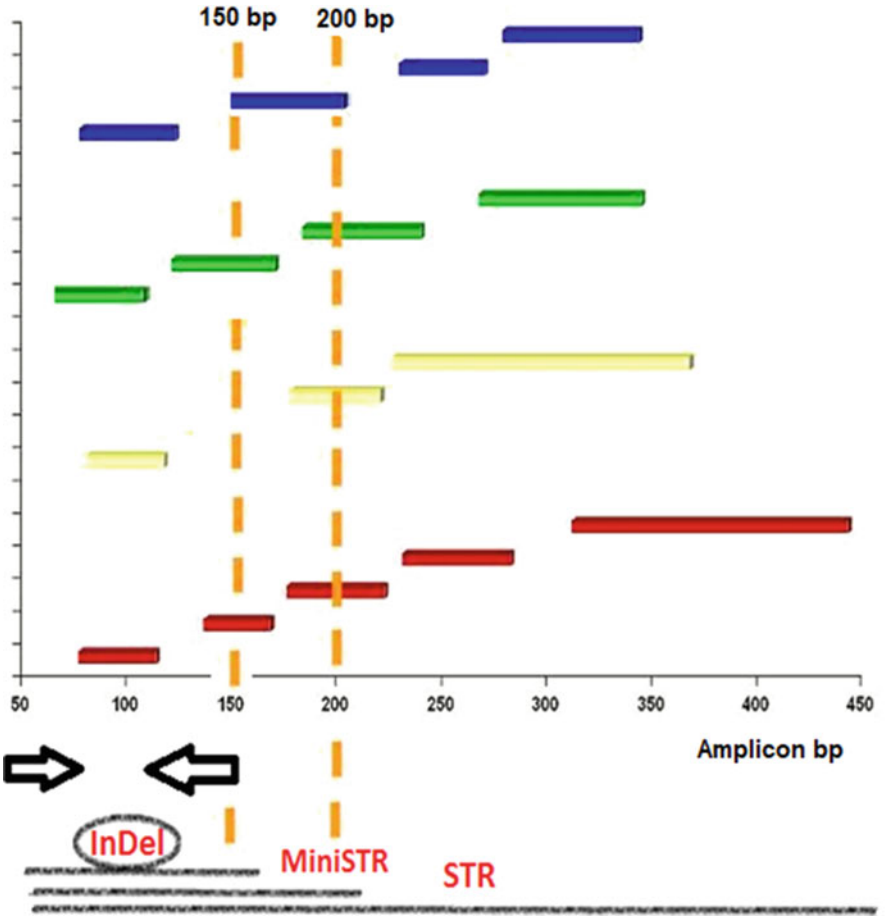

Fig. 4 Comparison of PCR product sizes of InDels with STRs

InDelPlex INDEL Polymorphism Detection Kit

It is a kit developed by Pereira R. et al. in partnership with the Institute of Molecular Immunology and Pathology of Oporto University and the University of Santiago de Compostela and commercialized with Genomica firm over the past few years. The kit allows multiplex PCR amplification of 38 InDel region and produced for the use of forensic identification and clinical diagnostic purposes. These analyses can be done by using ABI 310, 3130, 3130XL, 3500, and 3500XL capillary electrophoresis (Pereira et al. 2009a, b, 2012).

Mentype® DIPplex PCR Amplification Kit

It is a kit containing 30 InDel loci found on somatic chromosomes that have been marketed in recent years by the Biotype firm and also includes the locus of amelogenin. This kit restricted amplicon length to ~150 bp, which makes the kit

perfectly suitable for analyzing critical stains. The kit can be analyzed by using ABI 310, 3130, 3130XL, 3500, and 3500XL capillary electrophoresis (Mentype® DIPplex PCR Amplification Kit Handbook 2009).

Use of Gonosomal InDel Loci in Forensic Genetics

InDel loci exist on both somatic and sex chromosomes. Compared to somatic chromosomes, there are InDel variations in similar ratios on the X and Y chromosomes too. It can be used in addition to autosomal genetic markers in father/daughter relationship in X-linked polymorphism analysis. In the cases of paternity of two relative men (such as father/son), paternity can be determined by the use of polymorphism on the X chromosome if the child is a girl. It is because suspicious fathers will have different X chromosomes since they have different mothers (Szibor et al. 2005; Prinz and Sansone 2001).

Since the Y chromosome is transferred unchanged from father to son, it is observed in male members of the same family in the same form, except for mutations and genetic abnormalities. For this reason, population genetics, biogeographic lineage determination, male kinship evaluation or link analysis, and forensic genetic identification studies can be done by utilizing Y chromosome-related polymorphism analysis. In paternity cases, if the child is a boy, it is possible to get results by typing Y chromosome loci of any man (grandfather, uncle, cousin, etc.), who is in the family tree of the father candidate, especially when DNA cannot be obtained for various reasons (such as the father candidate cannot be found, DNA cannot be obtained from his biological material or he is dead, etc.) (Roewer et al. 2000; Corach et al. 2001; Prinz and Sansone 2001). Similarly, in paternity cases, where the father candidate is dead or could not be found, it is possible to get results by analyzing the X-linked polymorphism between the grandmother candidate and child (if the child is a girl) because the child gets one of her X chromosomes that her father gets from her grandmother. As a result, the paternal X chromosome in a girl will necessarily be coming from one of her grandmother's X chromosomes (Szibor et al. 2005, Prinz and Sansone 2001).

Pregnancies that occur after sexual assault crimes can result in abortion. In 6–8 weeks of abortions, the tissue and maternal blood materials are found together with miscarriage, and it is difficult to separate them microscopically. In this case, if the fetus is female, the presence of the X chromosome, which was inherited from the suspected father, can be determined by X-linked polymorphism analysis (Szibor et al. 2005; Prinz and Sansone 2001).

The Y chromosome is also used in illuminating sexual assault crimes. In cases where the victim is a woman, the victim's vaginal swab sample contains a mixture of DNA belonging to the victim and the perpetrator. Since there is no Y chromosome in women, only the profile of the perpetrator is obtained when identification is performed using Y chromosome loci in the mixture DNA (Roewer et al. 2000; Corach et al. 2001; Prinz and Sansone 2001).

By analyzing the polymorphism associated with X chromosome, it can also be determined whether two girls' father is the same, regardless of whether their mothers are different or the same. Since the father has one X chromosome, he will pass on the same X chromosome to all girls. But if their father is not the same, the X chromosome loci transferred from their father will also be different (Szibor et al. 2005; Prinz and Sansone 2001).

By using gonosomal and autosomal InDels together with STR and SNP loci, more successful results can be achieved in such cases or in identification of samples, which were collected from the crime scene (Szibor et al. 2005; Prinz and Sansone 2001).

Analysis of InDel Loci

Although InDel loci are biallelic, it is analyzed by the method of fragment analysis, such as STRs. These steps are followed in the identification of InDels as in STRs: first, extraction and quantification of DNA sample, then amplification of InDel loci of DNA by using fluorescence-marked InDel primers in multiplex PCR, and, finally, obtaining the profile of person by separating these loci in capillary electrophoresis with ABI 310, 3130, 3130XL, 3500, and 3500XL devices.

As a result of the analysis of InDel loci in capillary electrophoresis, three types of alleles can be seen in electropherogram because either deletion, insertion, or a combination of both can be transferred from the mother and father. These are expressed as deletion (Del – minor allele), insertion (Ins – major allele), and insertion-deletion (InDel), where both deletion and insertion are transferred together. The base pair size of the identified InDel loci on the DNA is defined by researches. Since it is defined how many base deletions or insertions they undergo, determining only the alleles as insertion or deletion will be sufficient to say the number of base pairs.

Before the analysis in capillary electrophoresis, this base pair data is defined in panel manager. It is created with a panel for each dye color (primer) with bins. In this way, the data of all alleles that exist in the population, just like in analysis of STR loci, is defined into the device. In capillary electrophoresis of DNA, if a person's peak was observed in the first bin, as shown in the figure, this shows the person's allele is "Del" because it has an incomplete base sequence. If the insertion has been transferred from the parent, the peak will be observed in the second bin, as shown in Fig. 5, which means that the person has the insertion as a result of addition of base, and the allele is "Ins." If a peak is observed in both bins, it means that the person has received both an insertion and deletion polymorphism from the parent. This, in turn, refers to the third allele, which we call "InDel" (Fig. 5).

In this way, the profile of people analyzed with InDel multiplex panel kits in all regions can be determined just like in STR analysis.

The number of loci that will reach sufficient (99.9%) discrimination power for human identification of InDel loci differs from STR loci number. In order to use InDels for identification purposes, as in STR analysis, a sufficient number of loci to

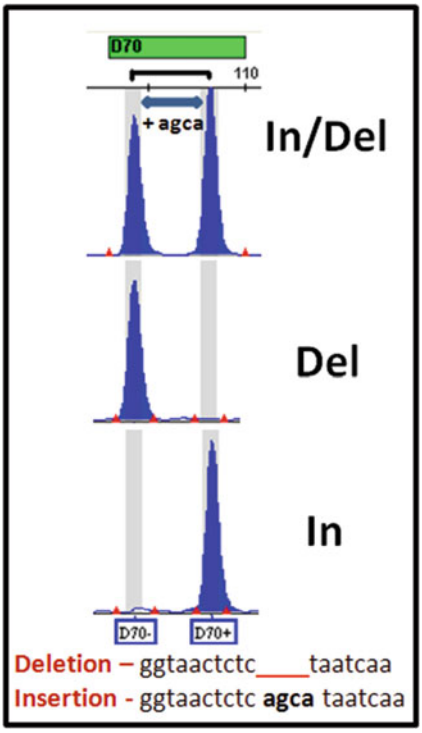

Fig. 5 Electropherogram image of insertion/deletion (InDel) polymorphism. (Fondevila Álvarez et al. 2011)

use are determined, depending on the heterozygosity rates and discrimination power of the locus contained in the multiplex kit or panel. The number of loci contained in InDel kits or panels developed by researchers, which exist on the market today, has a 99.9% discrimination power, and this locus number ranges from 20 to 40 (Ünsal et al. 2017).

While expressing D70+ as the insertion allele in the figure, D70- refers to the deletion allele.

Conclusion

Currently, STR systems are widely used for human identification in criminal laboratories. Forensic scientists are developing new polymorphic systems that can obtain results from all kinds of biological samples as an alternative to polymorphic systems (Coble and Butler 2005).

Insertion/deletion polymorphism occurs in the form of the addition or loss of one or more bases in the human genome, which is causing polymorphism and can be used in human identification and genealogical determination in illuminating forensic cases. Especially in cases where DNA obtained from biological evidence from the

crime scene is degraded or trace amount; As an alternative to STRs, small-sized systems on DNA are preferred for analysis in forensic sciences. In such cases, successful identification results can be obtained by using InDel polymorphism (Pereira et al. 2009a, b, 2012).

The loci that these systems occupy on DNA are quite small (60–200 bp), and the potential for a successful DNA profile is high, even if the DNA is degraded. The amount of DNA that can be obtained from a degraded or trace biological sample is either non-existent or ranges from 100 pg to 1 ng. Since it is known that the amount of DNA required for multiplex PCR used for identification studies is 0.5–1 ng, it has been determined by studies that identification with InDel loci gives successful results.Therefore InDels have been used in forensic science in recent years, and studies continue to develop InDel panels as an alternative or complement to STR and SNP loci in identification (Ünsal et al. 2017).

References

Campbell N.A, Williamson, B., Heyden, R.J. (2006) Biology: exploring life. Boston: Pearson Prentice Hall, 11(6): 242–243

Chan L (1992) Advances in molecular biology with applications in clinical medicine. Klin Lab 38: 2–4

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

Corach D, Risso FL, Marino M, Penacino G, Sala A (2001) Routine YSTR typing in forensic casework. Forensic Sci Int 118:131–135

Edward SA, Hammond HA, Jin L, Caskey T, Chakraborty R (1992) Genetic variation at five trimeric and tetrameric tandem repeats. Genomics 12:241–253

Fondevila Álvarez M, Pereira R, Gusma L, Phillips C, Lareu MV, Carracedo A, Butler JM, Vallone PM (2011) Forensic performance of short amplicon insertion-deletion (InDel) markers. Forensic Sci Int Genet Suppl Series 3:e443–e444

Gelbart WM, Lewontin RC, Griffith JF, Miller JH (2002) Modern genetic analysis: integrating genes and genomes. W.H. Freeman and CO., New York

Gordon A, Egner S (2013) Dynamic biology for high school. Sapling Learning, Austin

Gregory TR (2004) Insertion-deletion biases and the evolution of genome size. Gene 324:15–34

Guangyao F, Yi Y, Haibo L, Yiping H (2015) Screening of Multi-InDel markers on X-chromosome for forensic purpose. Forensic Sci Int Genet 5:42–44

Hongbao M (2005) Development application of polymerase chain reaction (PCR). J Am Sci 1(3):1–47

Investigator® DIPplex Handbook (2014) For multiplex amplification of 30 deletion/insertion polymorphisms, plus Amelogenin. Qiagen. https://www.qiagen.com/us/resources/download.aspx?id=2e1b66cf-e497-4054-82fc-3e1d2e14faba&lang=en. Accessed 12 Sep 2020

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

Kondrashov AS, Rogozin IB (2004) Context of deletions and insertions in human coding sequences. Hum Mutat 23(2):177–185

Lee HC, Ladd C, Bourke MT, Pagliaro E, Tirnady F (1994) DNA typing in forensic science I. Theory and background. Am J Forensic Med Pathol 15:269–282

Manta F, Caiafa A, Pereira R, Silva D, Amorim A, Carvalho EF, Gusmão L (2012) INDEL markers: genetic diversity of 38 polymorphisms in Brazilian populations and application in a paternity investigation with post mortem material. Forensic Sci Int Genet 6(5):658–661

Martínez-Cortés G, Gusmão L, Pereira R, Salcido VH, Favela-Mendoza AF, Muñoz-Valle JF, Inclán-Sánchez A, López-Hernández LB, Rangel-Villalobos H (2015) Genetic structure and forensic parameters of 38 Indels for human identification purposes in eight Mexican populations. Forensic Sci Int Genet 17:149–152

Mentype® DIPplex PCR Amplification Kit Handbook (2009) Biotype Diagnostic GmbH

Pereira R, Phillips C, Cíntia AA, Amorim Carracedo A, Gusma L (2009a) A new multiplex for human identification using insertion/deletion polymorphisms. Electrophoresis 30:3682–3690

Pereira R, Phillips C, Cíntia AA, Amorim Carracedo A, Gusma L (2009b) Insertion/deletion polymorphisms: a multiplex assay and forensic applications. Forensic Sci Int Genet Suppl Series 2:513–515

Pereira R, Phillips C, Pinto N, Santos C, Batista dos Santos SE, Amorim A, Carracedo A, Gusma L (2012) Straightforward inference of ancestry and admixture proportions through ancestry-informative insertion deletion multiplexing. PLoS One 7:e29684

Prinz M, Sansone MY (2001) Chromosome-spesific short tandem repeats in forensic casework. Croat Med J 42(3):288–291

Reiner AP, Ziv E, Lind DL, Nievergelt CM, Schork NJ, Cummings SR, Phong A, Burchard EG, Harris TB, Psaty BM, Kwok PY (2005) Population structure, admixture, and aging-related phenotypes in African American adults: the Cardiovascular Health Study. Am J Hum Genet 76: 463–477

Robertson J, Ross AM, Burgayne LA (1990) DNA in forensic science: theory. Techniques and applications. Ellis Horwood, London

Robertson J, Ross AM, Burgayne LA (2002) Properties of hypervariable single locus polymorphism and their application to identity testing. In: Balazs I (ed) DNA in forensic science: theory, techniques and applications. Taylor& Francis e-library, London

Rodriguez-Murillo L, Salem RM (2013) Insertion/deletion polymorphism. In Gellman, Marc D., Turner, J. Rick (eds) Encyclopedia of Behavioral Medicine. Springer-Verlag New York 1:1076.

Roewer L, Kayser M, Knijff P, Anslinger K, Corach D, Füredi S, Geserick G, Henke L, Hidding M, Kärgel HJ, Lessig R, Nagy M, Pascali VL, Parson W, Rolf B, Schmitt C, Szibor R, Teifel-Greding J, Krawczak M (2000) A new method for the evaluation of matches in recombining genomes: application to Y-chromosomal short tandem repeat (STR) haplotypes in European males. Forensic Sci Int 114:31–43

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 9:1713–1724

Szibor R, Hering S, Edelmann J (2005) A new web site compiling forensic chromosome X research is now online. Int J Legal Med 25:1–3

Ünsal T, Filoğlu G, Aşıcıoğlu F, Bülbül Ö (2017) Population data of new 21 mini-InDels from Turkey. Forensic Sci Int Genet Suppl Series 6:e189–e191

Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet 44:388–396