

Review

Genetic variation and DNA markers in forensic analysis

Imad Hadi Hameed^{1*}, Mohammed Abdullah Jebor², Aamera Jaber Ommer³, Cheah Yoke⁴,
Haider K. Zadian¹, Ali H. Al-Saadi¹ and Muataz A. Abdulazeez⁵

¹Department of Molecular Biology, Babylon University, Hilla City, Iraq.

²Centre of DNA Research, Babylon University, Hilla City, Iraq.

³Institute of medico-legal in Baghdad, Ministry of Health of Iraq, Iraq.

⁴Department of Biomedical Science, University Putra Malaysia, Selangor, Malaysia.

⁵College of Medicine, University of Al- Nahrain, Iraq.

Received 15 August, 2013; Accepted 20 June, 2014

In this review, we highlighted the DNA markers and their applications in forensic medicine which included the study of genetic variation or polymorphism, markers for genetic variation, types of variation, diallelic polymorphisms (SNPs and Indels), haplotypes, autosomal short tandem repeat (STRs), Y-chromosome short tandem repeat (Y-STRs) and Y chromosome. This is a useful tool for tracing human evolution, analysis of Y-chromosomal microsatellite haplotypes in globally human populations, organization of the human mitochondrial genome, and mitochondrial coding region as a source for variability and interpretation of sequence data. The light has been focused and directed in this study to establish the basic forensic genetic information, knowledge, data and statistics which might be so ultimately helpful practically in forensic science and criminology and to let evaluate and present the DNA weight evidences in medico-legal institute and courts of law.

Key word: Applications, DNA markers, forensic, genetic variation, review.

INTRODUCTION

Every organism stores its information on identity in its DNA (Dobbs et al., 1994). Separate chromosomes which are available in the nucleus of cells of humans have DNA. Serological and DNA typing are being used in forensic biology to analyze evidence found at crime

scenes. The sources biological evidences may be bodily fluids or tissues (Pierce and Wangh, 2007).

Microsatellites are a group of molecular markers chosen for a number of purposes which include forensic individual identification and relatedness testing

*Corresponding author. E-mail: imad_dna@yahoo.com. Tel: 009647716150716.

Author(s) agree that this article remain permanently open access under the terms of the [Creative Commons Attribution License 4.0 International License](#)

Abbreviations: AMH, Anatomically modern human; bp/kbp, base pair/thousand (kilo) base pairs; CRS, Cambridge reference sequence; D-loop, displacement loop/control region of mtDNA; DNA, deoxyribonucleic acid; hg(s), haplogroup(s); HVS-I/HVS-II, the first/second hypervariable segment; LD, linkage disequilibrium; LGM, last glacial maximum; MRCA, most recent common ancestor; mtDNA, mitochondrial DNA; np(s), nucleotide position(s); NRY, non-recombinant Y chromosomes; OXPHOS, oxidative phosphorylation; PCR, polymerase chain reaction; RFLP, restriction fragment length polymorphism; GD, gene diversity; HWE, Hardy-Weinberg equilibrium; ISFG, International Society of Forensic Genetics; PD, power of discrimination; PE, power of exclusion; PI, paternity index; PIC, polymorphism informative content; PM, match probability; Pr, probability; SNP, single nucleotide polymorphism; STR, short tandem repeat.

(Yamamoto et al., 1999; Nakamura et al., 2009). Low quantities of template DNA are required (10-100 ng) (Markoulatos et al., 2002), when using microsatellites. There is a high genomic abundance of random distribution throughout the genome. There is also an abundance of polymorphism. A nuclear DNA present in one copy per cell and only in males is called the Y chromosome. It includes the sex determining region and known as a paternal lineage marker (Butler et al., 2002; Carolina et al., 2010; Kuppareddi et al., 2010). The genetic information is inherited from the father to the son, and this information does not change except for mutational events (Hanson and Ballantyne, 2007; Mohammad and Imad, 2013). The individual short tandem repeat (STRs) are inherited as a single unit because of the lack of recombination which is called a haplotype and behaves as single allele per individual (Parson et al., 2003; Kwak et al., 2005).

The Y - chromosome is specific to the male portion of a male-female DNA mixed such as is common in sexual assault cases (Park et al., 2007). These STRs can also be useful in missing persons investigations, historical investigations, some paternity testing scenarios, and genetic genealogy (Park et al., 2007; Andrea et al., 2008). Although they are often used to suggest which haplogroup an individual matches, STR analysis typically provides a person haplotype. Most tests on the Y chromosome examine between 12 and 67 STR markers (Kayser et al., 1997; Imad et al., 2013a). The Y chromosome is less variable than the other chromosomes. Many markers are thus needed to obtain a high degree of discrimination between unrelated males (Kuppareddi et al., 2010; Hanson and Ballantyne, 2007).

The mitochondrial DNA (mtDNA) is a small circular genome located within the mitochondria in the cytoplasm of the cell. The mitochondrial genome can be divided into two sections: a large coding region, which is responsible for the production of various biological molecules involved in the process of energy production in the cell, and a smaller 1.2 kb pair fragment, called the control region. It is found to be highly polymorphic and harbors three hypervariable regions (HV): HV1, HV2 and HV3 (Kraytsberg et al., 2004; Imad et al., 2013b). Mitochondrial DNA Comprising of about 37 genes coding for 22 tRNAs, two rRNAs and 13 mRNAs are a small circle of DNA (Helgason et al., 2003).

Mitochondrial DNA does not recombine and thus there is no change between parent and child, unlike nuclear DNA. MtDNA is only passed on from mother to child and this is an important fact (Ingman et al., 2003; Ukhee et al., 2005). There is more sequence divergence in mitochondrial than in nuclear DNA (Brown et al., 1993; Giulietta et al., 2000). This may be caused by a faster mutation rate in mtDNA that may result from a lack of repair mechanisms. Sequencing of highly polymorphic segments of the control region of mitochondrial DNA (mtDNA) is today a routine method of analysis of biological

traces which are not suitable for STR analysis due to insufficient concentration of nuclear DNA or heavy degradation processes (Holland, Parsons, 1999; Tzen et al., 2001). A promising approach in this context seems to be analysis of selected single nucleotide polymorphisms (SNPs) that are useful for identification purposes.

GENETIC VARIATION OR POLYMORPHISM

The following: effective population size, population history (migration, bottleneck, and recent expansion), population structure and location of disease genes are determined by the amount and nature of genetic variation in a population. There are three items: (1) changes in nucleotides which could be transition or transversion. In the transition mutation, a pyrimidine (C or T) is substituted by another pyrimidine, or a purine (A or G) is substituted by another purine. The transversion mutation involves the change from a pyrimidine to a purine, or vice versa; (2) insertion or deletion of single nucleotides (indel) and (3) variation in number of repeat of tandemly repeated sequences (microsatellite, minisatellite and satellite) are genetic variation or polymorphism (Roach, 2010; Conrad, 2011).

The study of human genetic variation has both evolutionary significance and medical applications. It can help scientists understand ancient human population migrations as well as how different human groups are biologically related to one another. For medicine, study of human genetic variation may be important because some disease-causing alleles occur more often in people from specific geographic regions (Tishkoff and Verrelli, 2003; Conrad, 2011).

Diallelic polymorphisms (SNPs and INDELS)

Single-nucleotide polymorphisms (SNPs), the occurrence of different nucleotides at a specific place in the genome, are the most common type of human DNA sequence variation, occurring on average 1 per 500 to 1000 bp on a randomly selected chromosome. As noted later, this number varies somewhat depending on the population studied. Insertion/deletion polymorphisms (Indels), the occurrence of some more or fewer nucleotides at one position of the genome, are less common but do occur frequently.

The inserted/deleted sequence can be a few nucleotides long (Weber et al., 2002) up to several hundred nucleotides long, as is the case with the transposable Alu elements (Carter et al., 2004). Some very large polymorphic duplications, hundreds of kilobase pairs long, have also been identified (Sebat et al., 2004; lafrate et al., 2004). Because these polymorphisms are diallelic, they are less informative than microsatellites; however, SNPs are more common in the genome and more

amenable to automation and DNA chip technology. Most of the RFLPs (restriction fragment length polymorphisms) defined between the late 1970s and the late 1990s were SNPs in a restriction site but some Indels were occurring between restriction sites. That is, they are caused by a single mutational event that occurred once in the history of a species. We do not expect to see recurrent mutations at the same site except as extremely rare events. For the polymorphic Alu elements, we know that the ancestral state is absence and the insertion of the Alu is the derived state. For the other markers, we cannot tell from the human polymorphism which allele is ancestral and which is derived. However, we can determine the ancestral state in almost all of those cases simply by determining the genotype of our nearest relatives, the other great apes, following the logic in most cases. Humans share a single allele with the other apes and, by inference this is the ancestral or original allele in humans and the other allele is the derived allele (Hacia et al., 1999).

Multiallelic polymorphisms (STRPs and VNTRs)

Microsatellites consist of approximately 10-50 tandemly repeated copies of particular DNA sequence motifs ranging from 1 to 10 (most commonly 2-4) nucleotide base pairs. These repeat sequences, discovered in 1989 (Litt and Luty, 1989), occur frequently and randomly across the human genome. When the repeat number is polymorphic, microsatellites are also called short tandem repeat polymorphisms (STRPs) (other acronyms have also been used, for example, SSLP). STRP loci usually have multiple alleles and can have high levels of variation, that is, high heterozygosity. They rapidly replaced RFLP markers in gene mapping (especially disease genes) studies owing to these features and to their ease of typing, including the small amount of template DNA required. Remember that heterozygosity can never exceed 50% for a diallelic marker, whereas STRPs can easily have heterozygosities >75%. Because of this high heterozygosity, STRPs are the markers on which the most detailed human linkage maps are based (Dib et al., 1996; Kong et al., 2002; Jorgenson et al., 2005). STRPs have also become the standard for forensics and paternity testing (Budowle et al., 2001). The larger minisatellite arrays (also referred to as VNTRs, variable number of tandem repeats) are also highly polymorphic and powerful markers in forensic and paternity studies (DNA fingerprinting) (Jeffreys et al., 1985; Armour et al., 1996). However, they are less common than STRPs and are not evenly distributed throughout the genome. Their larger sequence motifs make them less amenable to PCR technology and use in genomic screening analyses. STRPs and VNTRs, on the other hand, tend to be much more dynamic than the diallelic SNPs and Indels. The mutation rates are higher,

there are more alleles, and it is usually impossible to determine which is the ancestral allele.

AUTOSOMAL SHORT TANDEM REPEAT (STRs)

Microsatellites refer to DNA with varying numbers of short tandem repeats (Allor et al., 2005; Klintschar et al., 2006) between a unique sequence (Figures 1 and 2). DNA regions with repeat units that are 2 to 7 bp in length or most generally short tandem repeats (STRs) or simple sequence repeats (SSRs) are generally known as microsatellites (Ellegren, 2004). In the core repeated bases, long repeat units may contain several hundred to thousands (Butler and Hill, 2012). Within the DNA there are length and sequence polymorphisms (Silvia et al., 2009). DNA can be used to study human evolution using human genome analysis regions that are not subjected to selection pressure (Mats et al., 2007). Besides, information from DNA typing provides vital information in medico-legal with polymorphisms allowing for more biological studies (Walkinshaw et al., 1996). It has been found that microsatellites are evenly distributed in the genome on all chromosomes and all regions of the chromosome (Ensenberger et al., 2010). They can also be found inside gene coding regions, introns, and in the non-gene sequences. Most microsatellite loci are really small, ranging from a few to a few hundred repeats and this small size of microsatellite loci is important for PCR-facilitated genotyping. Basically microsatellites containing a higher number of repeats are more polymorphic.

STR LOCUS NOMENCLATURE

The nomenclature for DNA markers is fairly straightforward. If a marker is part of a gene or falls within a gene, the gene name is used in the designation. For example, the STR marker TH01 is from the human tyrosine hydroxylase gene located on chromosome 11. The '01' portion of TH01 comes from the fact that the repeat region in question is located within intron 1 of the tyrosine hydroxylase gene. Sometimes the prefix HUM- is included at the beginning of a locus name to indicate that it is from the human genome. Thus, the STR locus TH01 would be correctly listed as HUMTH01 (Butler, 2005). DNA markers that fall outside of gene regions may be designated by their chromosomal position. The STR loci D5S818 and DYS19 are examples of markers that are not found within gene regions. In these cases, the 'D' stands for DNA. The next character refers to the chromosome number, 5 for chromosome 5 and Y for the Y chromosome. The 'S' refers to the fact that the DNA marker is a single copy sequence. The final number indicates the order in which the marker was discovered and categorized for a particular chromosome. Sequential numbers are used to give uniqueness to each identified

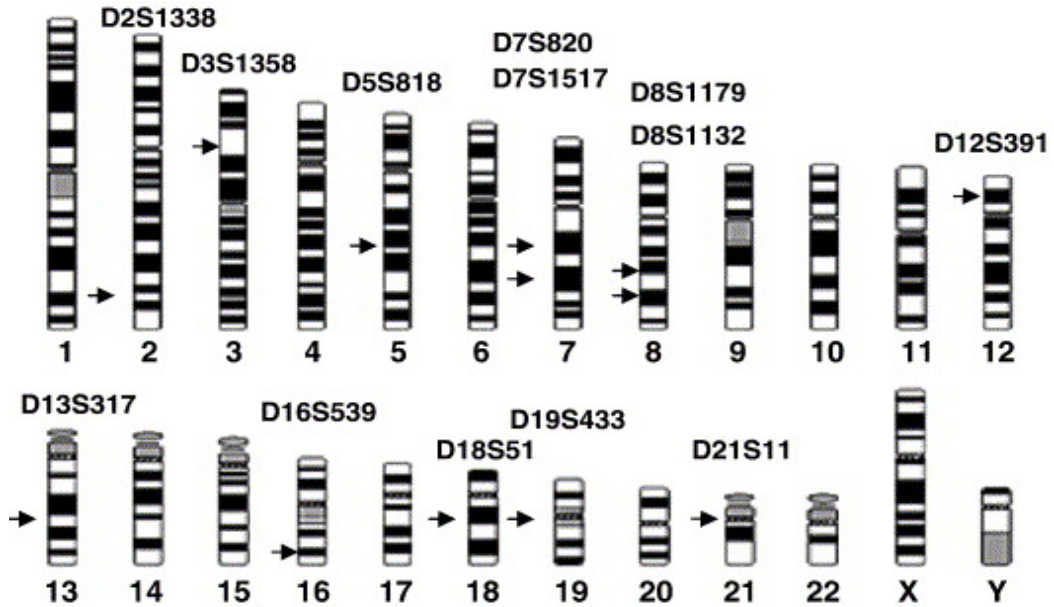


Figure 1. Exact physical location of 13 STR markers called CODIS : D2S1338 (q35), D3S1358 (3p21.31), D5S818 (5q23.2), D7S820 (7q21.11), D7S1517 (7q31.33), D8S1179 (8q24.13), D8S1132 (8q23.1), D12S391 (12p13.2), D13S317 (13q31.1), D16S539 (16q24.1), D18S51 (18q21.33), D19S433 (19q12), D21S11 (21q21.1). (Butler, 2006).

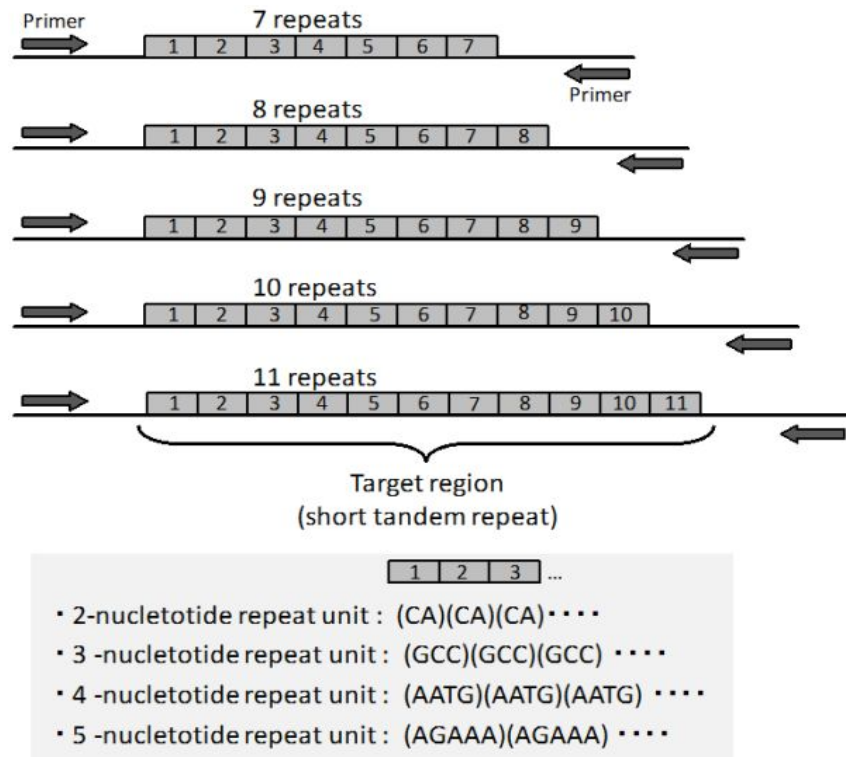


Figure 2. The structure of Short Tandem Repeat (STR) (Butler, 2006).

DNA marker (Butler, 2005). Thus, for the DNA marker D3S1358: D = DNA, 3 = Chromosome 3, S = single copy

sequence, 1358 = 1358th locus described on chromosome 3.

Table 1. Comparison of STR loci present in kits used in the United States.

Fix parameter here	Power Plex®18D	Power Plex®21	Power Plex®Fusion	Identifiler	Mini Filer	Extended ^a
TPOX	TPOX	TPOX	TPOX	TPOX	-	-
CSF1PO	CSF1PO	CSF1PO	CSF1PO	CSF1PO	CSF1PO	CSF1PO
D5S818	D5S818	D5S818	D5S818	D5S818	-	D5S818
D7S820	D7S820	D7S820	D7S820	D7S820	D7S820	D7S820
D13S317	D13S317	D13S317	D13S317	D13S317	D13S317	D13S317
FGA	FGA	FGA	FGA	FGA	FGA	FGA
vWA	vWA	vWA	vWA	vWA	-	vWA
D3S1358	D3S1358	D3S1358	D3S1358	D3S1358	-	D3S1358
D8S1179	D8S1179	D8S1179	D8S1179	D8S1179	-	D8S1179
D18S51	D18S51	D18S51	D18S51	D18S51	D18S51	D18S51
D21S11	D21S11	D21S11	D21S11	D21S11	D21S11	D21S11
TH01	TH01	TH01	TH01	TH01	-	TH01
D16S539	D16S539	D16S539	D16S539	D16S539	D16S539	D16S539
-	D2S1338	D2S1338	D2S1338	D2S1338	D2S1338	D2S1338
-	D19S433	D19S433	D19S433	D19S433	-	D19S433
-	-	D12S391	D12S391	-	-	D12S391
-	-	D1S1656	D1S1656	-	-	D1S1656
-	D2S441	-	-	-	-	D2S441
-	-	-	D10S1248	-	-	D10S1248
-	-	D6S1043	-	-	-	-
-	-	-	D22S1045	-	-	-
Penta D	Penta D	Penta D	Penta D	-	-	-
Penta E	Penta E	Penta E	Penta E	-	-	Penta E
-	-	-	DYS391	-	-	DYS391
Amelogenin	Amelogenin	Amelogenin	Amelogenin	Amelogenin	Amelogenin	Amelogenin

^a(see Ref. (Lareu et al., 1996) for more information on the proposed U.S. extended core).

COMMON COMMERCIAL KITS USED IN THE UNITED STATE

To further assist efforts in forensic DNA testing and human identification, Promega and Applied Biosystems have released a number of new STR typing kits Table 1. The European Standard Set of STR loci added to U.S. core loci used for the Combined DNA Index System are among the most recent STR kits were developed. Some of these kits are capable of simultaneous, multicolor fluorescence detection of 15 STRs and the sex-typing marker amelogenin in a single PCR reaction (Butler et al., 2006).

AUTOSOMAL STRs LOCUS INFORMATION

Autosomal STRs locus information were evaluated and selected at The Cooperative Human Linkage Center. <http://www.chlc.org> evaluates the genetic markers and the loci are selected from there. Additional STR loci, chromosomal location and repeat sequence for each core

STR locus are provided in Table 1 and Table 2 (Ruitberg et al., 2001; Klitschar et al., 2005; Klitschar et al., 2006; Imad et al., 2014a). According to the International Society of Forensic Genetics (ISFG) recommendation the repeat sequence motif be defined so that the first 5'-nucleotides on the Gene Bank forward strand define the repeat motif used. So the repeat motif for each STR marker is listed based on this. A significant fact is that STR allele sizes are measured relative to an internal size standard during electrophoresis. This depends on the DNA strand that is labeled using a dye that may have a different apparent measured size.

Y-CHROMOSOME SHORT TANDEM REPEAT (Y-STRs)

Chromosome Y microsatellites or STR's seem to be an ideal markers to delineate differences between human populations for several reasons: (i) they are transmitted in uniparental (paternal) fashion without recombination (De Knijff et al., 1997), (ii) they are very sensitive for genetic

Table 2. Information on 21 autosomal STR loci present in The PowerPlex® 21 System kits

STR Locus ^a n	Label	Physical position	Chromosomal location1	Repeat sequence 5'-3'
Amelogenin	Fluorescein	X and Y	Xp22.1-22.3 and Y	NA
D3S1358	Fluorescein	Chr 3 (45.582 Mb)	3p21.31	TCTA Complex
D1S1656	Fluorescein	Chr 1 (230.905 Mb)	1q42	TAGA Complex
D6S1043	Fluorescein	Chr 6 (92.450 Mb)	6q15	AGAT
D13S317	Fluorescein	Chr 13 (82.692 Mb)	13q31.1	TATC
Penta E	Fluorescein	Chr 15 (97.374 Mb)	15q26.2	AAAGA
D16S539	JOE	Chr.16(86.386Mb)	16q24.1	GATA
D18S51	JOE	Chr 18 (60.949 Mb)	18q21.33	AGAA
D2S1338	JOE	Chr 2 (218.879 Mb)	2q35	TGCC/TTCC
CSF1PO	JOE	Chr 5 (149.455 Mb)	5q33.1	AGAT
Penta D	JOE	Chr 21 (45.056 Mb)	21q22.3	AAAGA
TH01	TMR-ET	Chr 11 (2.192 Mb)	11p15.5	AATG (19)
vWA	TMR-ET	Chr12(6.093 Mb)	12p13.31	TCTA Complex (19)
D21S11	TMR-ET	Chr 21 (20.554 Mb)	21q21.1	TCTA Complex (19)
D7S820	TMR-ET	Chr 7 (83.789 Mb)	7q21.11	GATA
D5S818	TMR-ET	Chr 5 (123.111 Mb)	5q23.2	AGAT
TPOX	CXR-ET	Chr 2 (1.493 Mb)	2p25.3	AATG
D8S1179	CXR-ET	Chr 8 (125.907 Mb)	8q24.13 (125.976 Mb)	TCTA Complex (19)
D12S391	CXR-ET	Chr 12 (12.450 Mb)	12p12(12.341 Mb)	AGAT/AGAC Complex
D19S433	CXR-ET	Chr19(30.416 Mb)	(35.109 Mb)	AAGG Complex
FGA	CXR-ET	Chr 4 (155.509 Mb)	4q28 (155.866 Mb)	4q28 (155.866 Mb)

^adatabase of sequence-tagged sites (STSs) available on the NCBI website: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=unists>. ⁿThe 13 CODIS core loci are highlighted in bold font. Adapted from Cotton et al., 2000; Wiegand et al., 1993; physical positions are from Schneider et al., 1998.

drift, and (iii) they allow a simple highly informative haplotype construction (Kayser et al., 1997). Also, for forensic applications, this ability to differentiate distinct Y chromosomes makes Y-STR's an advantageous addition to the well characterized autosomal STR's. For a number of forensic applications Y-STR's could be superior to autosomal STR's. Especially in rape cases where (i) the differential extraction was unsuccessful, (ii) the number of sperm cells is very low, (iii) due to vasectomy epithelial cells instead of sperm cells from the ejaculate of the perpetrator have to be analyzed, and (iv) the perpetrator, due to a familial relationship shares many autosomal bands with the victim, Y-STR's could provide crucial evidence. Also, in the case of male-male rape or rape cases with multiple perpetrators Y-STR's could lead to essential qualitative evidence. In all such cases Y-STR's facilitates a simple and reliable exclusion of suspects (Redd et al., 1997).

COMMON COMMERCIAL KITS USED IN THE UNITED STATES

Two most popular commercial kits in the U.S., are namely PowerPlex®Y System (a,b) and AmpFISTR®Yfiler™ PCR Amplification Kit. There are

currently four online searchable Y-STR haplotype databases based in the United States and intended for forensic use. Three are maintained by commercial vendors: Reliagene, Inc., Promega Corporation and Applied Biosystems, Inc. The fourth is maintained by the University of Arizona. The National Center for Forensic Science (NCFS) also maintains a Y-STR database that will soon be available online. However, these databases are somewhat limited in the number of individuals and loci profiled, which sometimes limits their operational usefulness.

Y CHROMOSOME IS A USEFUL TOOL FOR TRACING HUMAN EVOLUTION

The Y chromosome is becoming a useful tool for tracing human evolution through male lineages (Jobling, 1994) as well as application in a variety of forensic situations (Kayser et al., 1997) including those involving evidence from sexual assault cases containing a mixture of male and female DNA (Prinz et al., 1997; Prinz et al., 2001). Y-chromosome specific methods can improve the chances of detecting low levels of male DNA in a high background of female DNA.

The European Y chromosome typing community has

established a "minimal haplotype" and an "extended haplotype" for inclusion of common loci into a central DNA database. These haplotypes consist of results from the following Y STR markers: DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS385, and YCAII (Roewer et al., 2001). The combination of these Y STRs can distinguish approximately 74 to 90% of male individuals in various local European populations (Kayser et al., 1997). The minimal haplotype, the results of which have been accepted for court use in Europe (Pascali et al., 1999), is most commonly used and is typically obtained via two or three separate multiplex amplifications. In this study, a Y STR 20 plex capable of the simultaneous amplification of 20 polymorphic Y chromosome-specific PCR products is introduced. This 20 plex includes all of the markers in the European extended haplotype (Roewer et al., 2001) and also contains the trinucleotide loci DYS388 (M. Kayser et al., 1997) and DYS426 (Jobling et al., 1996), the tetranucleotide loci DYS437 (Ayub et al., 2000; Grignani et al., 2000), DYS439 (Ayub et al., 2000; Grignani et al., 2000), GATA A7.1 (DYS460) and H4 (White et al., 1999; Gonzalez-Neira et al., 2001), the pentanucleotide loci DYS438 (Ayub et al., 2000; Grignani et al., 2000) and DYS447 (Redd et al., 2002), and the hexanucleotide marker DYS448.

Analysis of Y-chromosome microsatellite haplotypes in globally human population

Variability at microsatellite or short tandem repeat (STR) loci is being used, in various species, for linkage analysis, individual identification (Hammond et al., 1994; Redd et al., 1997), and population-genetic analyses. Autosomal STR loci have also been successfully applied to reconstruct human evolutionary history (Goldstein et al., 1995). The resulting phylogenetic trees reveal evolutionary relationships similar to those based on mtDNA sequence variation, and it has been a long lasting wish to add to these trees the one strictly based on Y chromosome-specific markers. Approximately 5 years ago, the number of published STR loci on the human Y chromosome was <15, and only one Y chromosome-specific minisatellite was known (Roewer et al., 1992). The number of known Y chromosomal single-nucleotide polymorphisms (SNPs) was <10 (Nakahori et al., 1989; Jobling, 1994; Mathias et al., 1994; Seielstad et al., 1994; Hammer, 1995; Whitfield et al., 1995), and only one Alu-insertion polymorphisms had been discovered. This picture, however, has totally changed, owing to the recent introduction of many Y SNPs (Underhill et al., 2000) and STRs (White et al., 1999), and additional new markers are to be expected.

Recently, Y-STR variability has been used both for the dating of SNP mutations, in order to draw conclusions about the origins and history of human populations

(Underhill et al., 1996; Zerjal et al., 1997), and for human identification in forensic casework (Kayser et al., 1997; Prinz et al., 1997). Nevertheless, global studies of Y-STR variability are still rare (Seielstad et al., 1999), and most of these analyses have been based on combined single-locus information. The major advantage of analyzing the nonrecombining part of the Y chromosome is that single-locus information can be used to construct compound haplotypes that allow male lineages to be characterized in a much more detailed fashion. It has been questioned whether the tracing of human migration history can be achieved solely on the basis of Y STRs (De Knijff et al., 1997).

Y-STR database

When the DNA profile of a known suspect or victim matches the DNA profile from crime scene evidence, the individual is "included" as a potential source of that evidence. In the U.S., the strength of the match is most often expressed as a statistic that describes the estimated frequency of occurrence of the DNA profile in unrelated individuals within various population groups. Due to the lack of recombination along most of the length of the Y chromosome, Y-STR loci are not statistically independent of one another (unlike standard autosomal STR markers) and are co-inherited as extended haplotypes of linked markers. Therefore, multiplication of single-locus allele frequencies to obtain estimated Y-STR haplotype frequencies is not appropriate. An estimation of the frequency of occurrence of a particular Y-STR haplotype necessitates the use of a counting method, which, with the limited sizes of databases available, produces an estimate that depends entirely upon the size of the database used. Thus, large databases of multi-locus Y-STR haplotypes need to be generated to maximize the probity of Y-STR evidence. A large comprehensive European-based Y-STR database is maintained by the Institute of Legal Medicine, Charité-University Medicine Berlin (www.yhrd.org).

However, although a subset of this database comprises the SWGDAM core loci, it is less useful for frequency estimates from haplotypes that have been generated using the two most popular commercial kits in the U.S., namely the PowerPlex®Y System(a,b) and AmpFISTR®Yfiler™ PCR Amplification Kit. There are presently four online searchable Y-STR haplotype databases based in the United States and intended for forensic use. Three are maintained by commercial vendors: Reliagene, Inc., Promega Corporation and Applied Biosystems, Inc. The fourth is maintained by the University of Arizona. The National Center for Forensic Science (NCFS) also maintains a Y-STR database that will soon be available online. These databases differ in the number of Y-STR markers and individuals, although all possess the SWGDAM core loci. However, these

Table 3. Current U.S. - Based Y-STR databases.

Agency	URL	Number of markers	Number of samples
National Center for Forensic Science	To be determined	76	1.396
University of Arizona	http:// Amadeus.biosci.arizona.edu/kcaldero/str	83	2.518
Applied Biosystems	www. Appliedbiosystems.com/yfilerdatabase	17	3.561
Promega Corporation	www.promega.com/techserv/tools/pplexy/	12	4.004
Reliagen	www.reliagen.com/index.asp?menu_id=rdcontent	11	4.623
Potential Size of National Y-STR Database			16.102

databases are somewhat limited in the number of individuals and loci profiled, which sometimes limits their operational usefulness. For example, the biggest U.S.-based database comprises haplotypes from 4,623 individuals. By combining data from these U.S. databases, a much larger Y-STR database of approximately 16,000 individuals can be created, resulting in a significant increase in the probative value of Y-STR evidence. Also, merging the NCFS and University of Arizona databases will increase the number of samples with extended Y-STR loci haplotypes, which may be of assistance to those interested in developing the next generation of Y-STR multiplex systems. Establishing a national database that incorporates data from a multitude of sources requires the implementation of a number of quality indicator metrics. Quality assurance procedures must be developed to govern the suitability and quality of data from diverse sources. For example, it may be necessary for donors of data to establish analytical prowess by testing externally provided proficiency samples. Since each commercial kit or academic multiplex system uses different primer sets, it will also be essential to ensure that allele calls are equivalent regardless of the multiplex system employed.

Importantly, merged data must be purged of duplicate samples that have been submitted by the same donor to multiple databases. To effectively manage the data, a Y-STR Database Consortium comprised of database stakeholders from commercial companies, academia, the FBI and U.S. crime laboratories was formed at the February 2006 AAFS meeting in Seattle (Table 3). It was agreed that NCFS, a program of the National Institute of Justice (NIJ) hosted by the University of Central Florida, would maintain and manage the consolidated Y-STR database on behalf of stakeholders. The National Institute of Justice is funding this effort. As a group, we are working to collate existing Y-STR data from various commercial and academic sources and have enlisted the aid of geographically diverse crime laboratories to furnish additional samples.

MITOCHONDERIA

Structure of mitochondria

The typical human cell has several hundred mitochondria,

cytoplasmic organelles that convert energy to forms that can be used to drive cellular reactions. Without them cells would be dependent on anaerobic glycolysis for all their adenosine triphosphate (ATP). The mitochondria have a characteristic double membrane structure, in which the outer membrane contains large channel-forming proteins (called porin) and is permeable to all molecules of 5000 daltons or less, while the inner membrane is impermeable to most small ions and is intricately folded, forming structures called cristae. The large surface area of the inner mitochondrial membrane accommodates respiratory chain and ATP synthase enzymes involved in the process of oxidative phosphorylation (OXPHOS). The mitochondrial matrix contains hundreds of enzymes, including those required for the oxidation of pyruvate and fatty acids and those active in the tricarboxylic acid (TCA) cycle. The matrix also contains several identical copies of the mitochondrial DNA, mitochondrial ribosomes, tRNAs and various enzymes required for the transcription and translation of mitochondrial genes (Alberts et al., 1994).

Organization of the human mitochondrial genome

The human mitochondrial genome is 16,569 base pairs (bp) in length (Anderson et al., 1981), a closed, circular molecule located within the mitochondrial matrix and present in thousands of copies per cell. Mitochondrial DNA has two strands, a guanine rich heavy (H) strand and a cytosine-rich light (L) strand. The heavy strand contains 12 of the 13 polypeptide-encoding genes, 14 of the 22 tRNA-encoding genes and both rRNA encoding genes. Introns are absent in mtDNA, and all of the coding sequences are contiguous (Anderson et al., 1981; Wallace et al., 1992; Zeviani et al., 1996).

The only non-coding segment of mtDNA is the displacement loop (D-loop), a region of 1121 bp that contains the origin of replication of the H-strand (OH) and the promoters for L and H-strand transcription. The mtDNA is replicated from two origins. DNA replication is initiated at OH using an RNA primer generated from the L-strand transcript. H-strand synthesis proceeds two-thirds of the way around the mtDNA, displacing the parental H strand until it reaches the L-strand origin (OL), situated in a cluster of five tRNA genes. Once exposed on the displaced H- strand, OL folds a stem-loop structure

and L-strand synthesis is initiated and proceeds back along the H-strand template. Consequently, mtDNA replication is bidirectional but asynchronous (Clayton, 1982). MtDNA transcription is initiated from two promoters in the D-loop, PL and PH. Transcription from both promoters proceeds around the mtDNA circle, creating a polycistronic RNA. The tRNA genes which punctuate the larger rRNA and mRNA sequences then fold within the transcript and are cleaved out. The mRNAs and rRNAs liberated are posttranscriptionally polyadenylated and the tRNAs are modified and the 3' terminal CCA added (Attardi et al., 1982; Attardi and Montoya, 1983; Clayton, 1984; Wallace, 1992; Taanman, 1999).

Mitochondrial DNA replication

MtDNA is replicated by the DNA polymerase gamma complex which is composed of a 140 kDa catalytic DNA polymerase encoded by the POLG gene and a 55 kDa accessory subunit encoded by the POLG2 gene (Van der Tovar, 2005). Replication of mitochondrial DNA begins on only one strand within the non-coding "control" region. As this strand is replicated, the opposite strand of the original DNA duplex is displaced and forms a single-stranded loop hence the name "D-loop" for Displacement loop. The displaced strand is then broken down and the replication process begins again displacing a single strand in the process (Schon et al., 2012).

Reasons for using mitochondrial dna rather than nuclear dna

First, multiple copies: Each mitochondrion contains its own DNA, with many copies of the circular mitochondrial DNA in every cell. It is thought that each mitochondrion contains between 1 and 15, with an average of 4 to 5, copies of the DNA (Reynolds et al., 2000) and there are hundreds, sometimes thousands, of mitochondria per cell. The result is that there are many thousands of copies of the mitochondrial DNA in every cell. This compares with only two copies of nuclear DNA. Second, Better protection: The mitochondrion also has a strong protein coat that protects the mitochondrial DNA from degradation by bacterial enzymes. This compares to the nuclear envelope that is relatively weak and liable to degradation.

Third, higher rate of evolution: DNA alterations (mutations) occur in a number of ways. One of the most common ways by which mutations occur is during DNA replication. An incorrect DNA base may be added; for example, a C is added instead of a G. This creates a single base change, or polymorphism, resulting in a new form. These single base mutations are rare, but occur once in every 1,200 bases in the human genome. The

result is that the rate of change, or evolutionary rate, of mitochondrial DNA is about five times greater than nuclear DNA (Bar et al., 2000). This is important in species testing, as even species thought to be closely related may in time accumulate differences in the mitochondrial DNA but show little difference in the nuclear DNA. Finally, maternal inheritance: A further reason for the use of mitochondrial DNA in species testing, and in forensic science, is its mode of inheritance. Mitochondria exist within the cytoplasm of cells, including the egg cells. Spermatozoa do not normally pass on mitochondria and only pass on their nuclear DNA. The resulting embryo inherits all its mitochondria from its mother (Tully et al., 2004; Brown et al., 2002; Brown et al., 2000).

The result is that mothers pass on their mitochondrial DNA type to all their offspring, but only the daughters will pass on the mitochondrial DNA to the next generation. Mitochondrial DNA is therefore passed from generation to generation down the maternal line. Mechanisms for this include simple dilution (an egg contains 100,000 to 1,000,000 mtDNA molecules, whereas a sperm contains only 100 to 1000), degradation of sperm mtDNA in the fertilized egg, and, at least in a few organisms, failure of sperm mtDNA to enter the egg. Whatever the mechanism, this single parent (uniparental) pattern of mtDNA inheritance is found in most animals, most plants and in fungi as well. Also, most mitochondria are present at the base of the sperm's tail, which is used for propelling the sperm cells. Sometimes the tail is lost during fertilization. Also, unlike nuclear DNA, where there is a shuffling of the chromosomes at every generation, the mitochondrial DNA does not recombine with any other DNA type and remains intact from generation to generation (Brown et al., 2002; Brown et al., 2000; Pastores et al., 1994; Guntheroth et al., 1990). The role of DNA is to encode protein and RNA molecules, and the mitochondrial DNA is no different. All mammalian mitochondrial DNA is very similar, with the order and position of the genes being the same. The general structure of the mitochondrial DNA is shown in Figure 3. As with nuclear DNA, to indicate the significance of the match, analysts usually estimate the frequency of the sequence in some population. The estimation procedure is actually much simpler with mtDNA. It is not necessary to combine any allele frequencies because the entire mtDNA sequence, whatever its internal structure may be, is inherited as a single unit (a "haplotype"). In other words, the sequence itself is like a single allele, and one can simply see how often it occurs in a sample of unrelated people (Lutz et al., 2003; Alvarez et al., 2007).

This polymorphism allows scientists to compare mtDNA from crime scenes to mtDNA from given individuals to ascertain whether the tested individuals are within the maternal line (or another coincidentally matching maternal line) of people who could have been the source of the trace evidence (Coble et al., 2006).

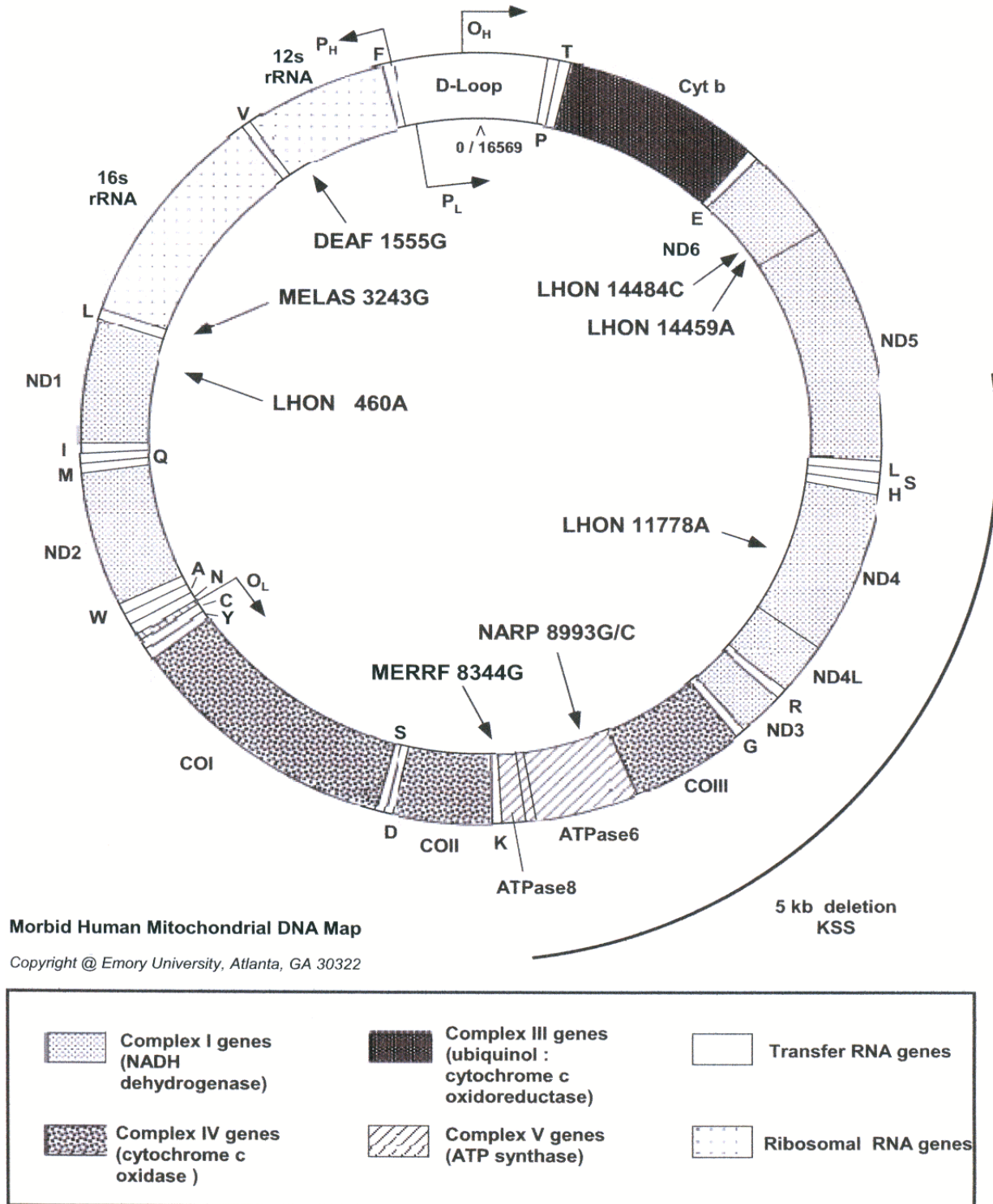


Figure 3. The Human Mitochondrial DNA Genome. The genes encoded by the mitochondrial DNA (mtDNA) genome are noted. Point mutations associated with mitochondrial diseases are noted in the center of the genome (Brown et al., 2002).

Mitochondrial coding region as a source for variability

Forensic mtDNA analysis commonly investigates the HVI

and HVII regions, but analysis of the coding region is gaining interest as this region harbours highly polymorphic mutations that may increase the discrimination of most common mtDNA types (Parsons and Coble et al.,

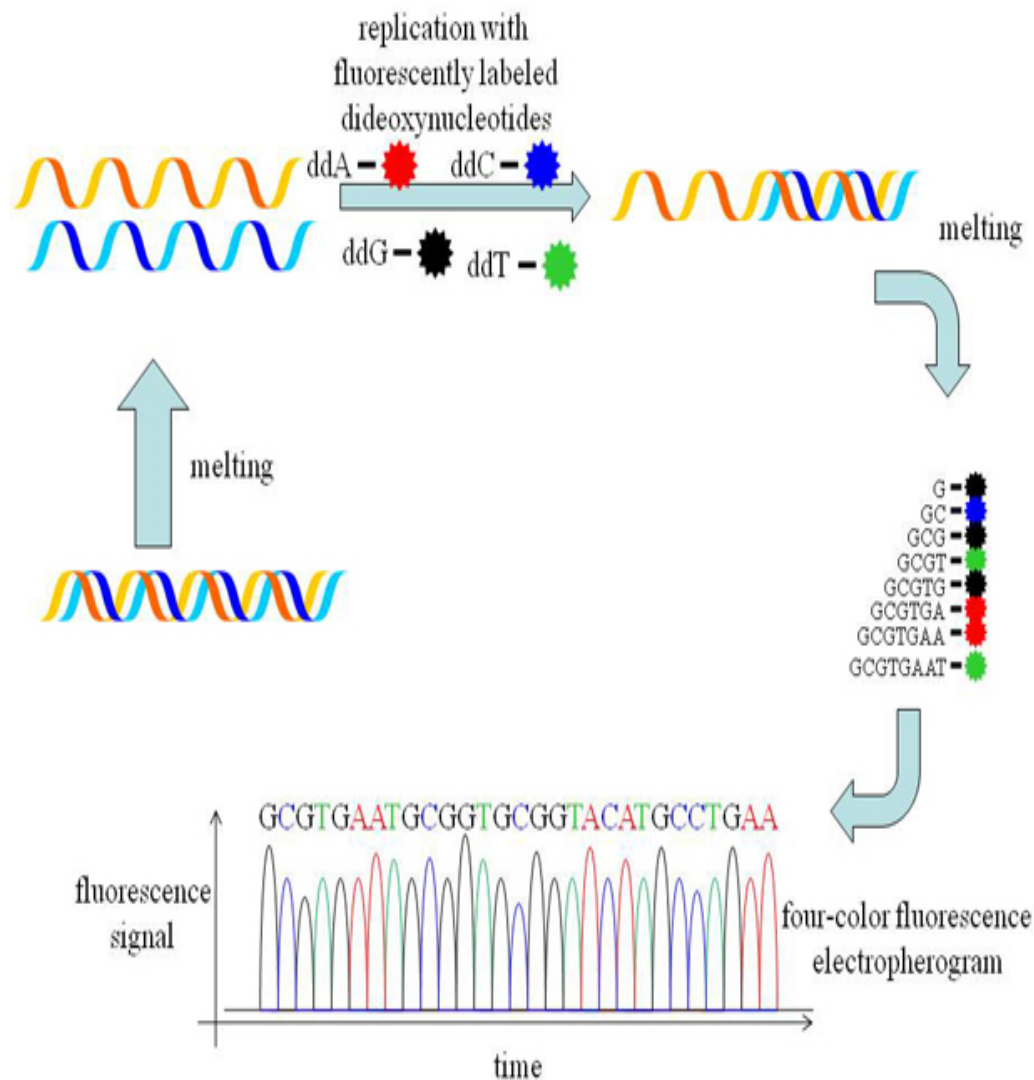


Figure 4. The sketch of Sanger capillary sequencing.

2001; Andreasson et al., 2002; Imad et al., 2014b). The development of high throughput capillary sequencing has made whole mitochondrial genome analysis feasible and large collections of coding region data are now progressing (Herrnstadt et al., 2002). Thus, selection of new markers suitable for forensic coding mtDNA analysis is feasible. In a study by Coble et al. (2004) a panel of coding region SNPs were selected to resolve the most common HVI and HVII sequences including the H haplogroup. An SNP assay was chosen as whole mitochondrial genome sequencing is not a practical approach for the analysis of forensic samples. Moreover, coding SNPs were also chosen to resolve four additional commonly observed HVI and HVII haplotypes as the H-defining SNPs may not be strained by only choosing silent and non-private polymorphism (variant observed) in more than one individual for ethical and practical reasons, nearly a four-fold increase in discrimination of the five

most common haplogroups (H, J, T, V and U) was obtained (Coble et al., 2004).

INTERPRETATION OF SEQUENCE DATA

Current Sanger capillary sequencing systems are like the widely used Applied Biosystems (Figure 4), the dNTP/ddNTP mixture causes random, non-reversible termination of the extension reaction, creating from the different copies molecules extended to different lengths. Following denaturation and cleaning up of free nucleotides, primers, and the enzyme, the resulting molecules are sorted by their molecular weight (corresponding to the point of termination) and the label attached to the terminating ddNTPs is read out sequentially in the order created by the sorting step. Additionally, the advent of more sensitive detection systems and several rounds

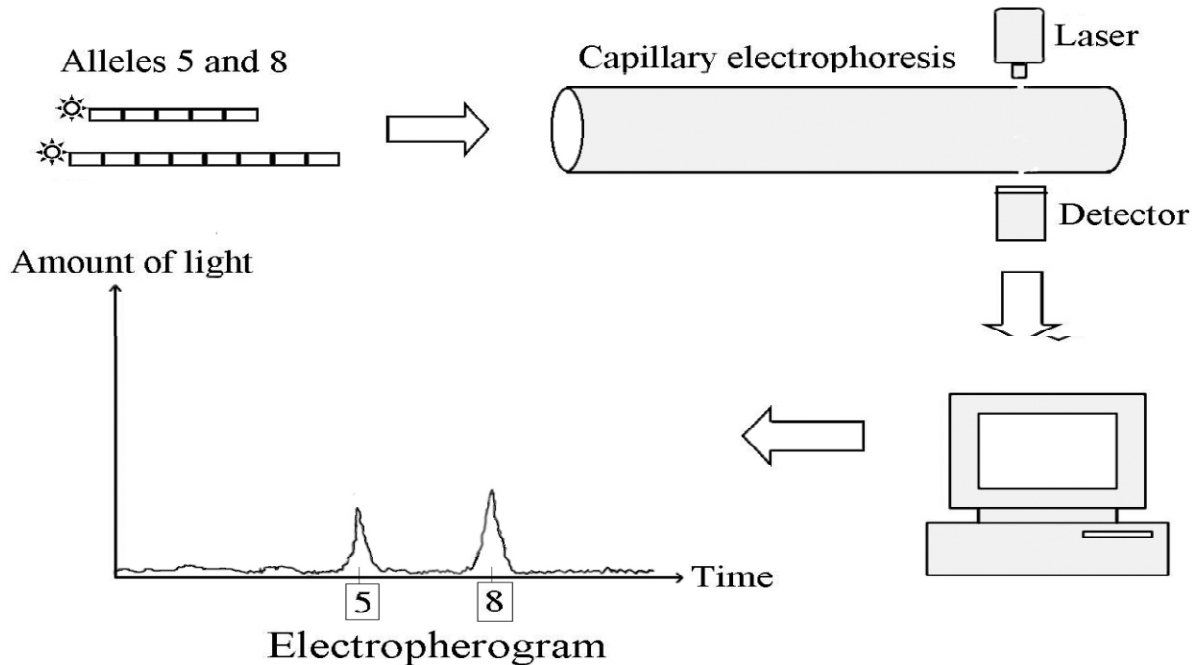


Figure 5. Shows the sketch of an electropherogram for two D16S539 alleles. One allele has eight repeats of the sequence GATA and the other has five. Small rectangle represents a GATA repeat. For illustration here only one copy of each allele (with a fluorescent molecule, or "tag" attached) is shown. However, PCR generates many more copies from the DNA sample with these alleles at the D16S539 locus. Whilst these copies are drawn through the capillary tube, the tags glow as the STR fragments move pass the laser beam. The colored light from the tags is measured using an electronic camera. A computer is finally used to produce the electropherogram based on the signals received.

of primer extensions (equivalent to a linear amplification) permit smaller amounts of starting DNA to be used for modern sequencing reactions. Using current Sanger sequencing technology, it is technically possible for up to 384 sequences (Emrich et al., 2002) of between 600 and 1,000 nt in length to be sequenced in parallel.

Base designations:

"A" designation—green peaks

"G" designation—black peaks

"T" designation—red peaks

"C" designation—blue peaks

"N" designation—peaks that, for whatever reason, are not clear enough to designate as A, G, T, or C. These bases are generally inconclusive. Often, a position that has an inconclusive (N) base in one direction may appear very clear in the other direction.

TYPES OF POLYMORPHISMS

Purines to purines or pyrimidines to pyrimidines -
Transversions: (purines to pyrimidines or pyrimidines to purines).

Insertions: an extra base is present when compared to the Anderson reference sequence.

Deletions: a base is missing when compared to the Anderson reference sequence.

SEQUENCHER™ 4.7

The validate mitotypes command allows ours to compare the results of separate analysts. Select two contigs which have been assembled to the same reference sequence (SEQUENCHER™ 4.7 User Manual for Windows © 1991 - 2007 Gene Codes Corporation, Inc. All rights reserved).

After the raw sequencing data has been collected by the 310 genetic analyzer and analyzed by the Sequencing Analysis program, it must be reviewed manually and interpreted with the aid of the Sequencer™ software (Figure 5). Sometimes the data will present characteristics that make interpretation more complex, though not impossible. Interpretive skills improve with experience, but listed below are suggestions for those less familiar with mtDNA analysis on dealing with some of the more common features that affect interpretation. However, all data (especially indications of mixtures, heteroplasmy, etc.) should be reviewed on a case-by case basis, and interpretations made

by the analyst based on his or her own expertise.

The forensic community generally recognizes the HV1 region as consisting of base positions 16024-16365 and HV2 as base positions 73-340, as numbered according to the Cambridge Reference Sequence (also referred to as the Anderson Sequence) (Bainbridge et al., 2011). This laboratory will attempt to sequence all the bases within these defined regions, as well as an additional 25 base pairs in either direction (that is, 15999-16390 and 48-365). While it is preferable to have confirmation of each sequence by comparing the forward and reverse strands, it is sometimes necessary to use two forward strands or two reverse strands as confirmation of the sequence (Al-Hagggar et al., 2013; Pareek et al., 2011). In these cases it is suggested that the sample be cycle sequenced a second time to obtain the confirmatory strand in the same direction. In addition to case samples, a positive and negative control must be sequenced. The purpose of controls is to show that each stage of the analysis is working properly. Therefore, only one positive and negative control is necessary for each step of the procedure (for example, if samples from three amplification runs are combined into one cycle sequencing run, only one positive and one negative must be carried through to show the sequencing reagents and thermal cyclers have performed as expected). If the results obtained for the controls are not as anticipated, the evaluation of the results will be determined on a case-by-case basis (Goldberg et al., 2006).

Conflict of Interests

The author(s) have not declared any conflict of interests.

REFERENCES

- Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD (1994). Energy conversion: Mitochondria and chloroplasts. In: Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD (eds) *Molecular biology of the cell*. Garland Publishing, Inc., New York, pp.653-720.
- Al-Hagggar MMS, Khair-Allaha BA, Islam MM, Mohamed AS (2013). Bioinformatics in High Throughput Sequencing: Application in Evolving Genetic Diseases. *J Data Mining Genomics Proteomics*, 4:131.
- Allor C, Einum DD, Marco S (2005). Identification and Characterization of Variant Alleles at CODIS STR Loci. *J. For. Sci.* 50(5):1128-1133.
- Alvarez-Iglesias V, Jaime JC, Carracedo A, Salas A (2007) "Coding Region Mitochondrial DNA SNPs: Targeting East Asian and Native American Haplogroups." *Forensic Science International. Genetics* 1(1):44-55.
- Anderson S (1981). Sequence and organization of the human mitochondrial genome. *Nature* 290:457.
- Andrea V, Nicoletta C, Fausta G, Anna P, Elena M, Francesco DF (2008). Population Data for 15 Autosomal STRs Loci and 12 Y Chromosome STRs Loci in a Population Sample from the Sardinia Island (Italy). *Leg Med.*, 11(1):37-40.
- Armour J, Anttinen T, May CA, Vega EE, Sajantila A, Kidd JR, Kidd KK, Jeffreys AJ (1996). Minisatellite diversity supports a recent African origin of modern humans. *Nature Genetics*, 13:154-160.
- Attardi G, Montoya J (1983). Analysis of human mitochondrial RNA. *Methods Enzymol.* 97:435-469.
- Attardi G, Chomyn A, Montoya J, Ojala D (1982). Identification and mapping of human mitochondrial genes. *Cytogenet. Cell Genet.* 32:85-98.
- Ayub AM, Qamar R, Mazhar K, Zerjal T, Mehdi SQ, Tyler-Smith C (2000). Identification and characterisation of novel human Y-chromosomal microsatellites from sequence database information. *Nucleic Acids Res.* 28:e8.
- Bainbridge MN, Wiszniewski W, Murdock DR, Friedman J, Gonzaga-Jauregui C (2011). Whole-genome sequencing for optimized patient management. *Sci. Transl. Med.* 3:87re3.
- Ballantyne KN (2010). Mutability of Y-chromosomal microsatellites: rates, characteristics, molecular bases, and forensic implications. *Am. J. Hum. Genet.* 87:341-353.
- Bar W (2000). DNA commisionof the international society for forensic genetic : guidelines for mitochondrial DNAtyping. *Int . J. Legal Med.* 113(4):193-196.
- Beleza C, Alves A, Gonzalez-Neira M, Lareu A, Amorim A (2003). Extending STR markers in Y chromosome haplotypes. *Int. J. Legal Med.* 117:27-33.
- Brown JR, Bechenbach AT, Smith MJ (1993). Intraspecific DNA sequence variation of the mitochondrial control region of white sturgeon (*Acipenser transmontanus*). *Mol. Biol. Evol.* 10:326-341.
- Brown MD (2000). The role of MtDNA background in disease expression: Anew primary LHON mutation associated with Western Eurasian haplogroup. *J. Hum. Genet.* 110(2):130-138.
- Brown MD (2002). Functional analysis of lymphoplast and cybrid mitochondria containing the 3460. 11778. or 14484 Lebersheredity optic neuropathy mitochondrial DNA mutation. *J. Biol. Chem.* 275(51):39831-39836.
- Budowle B, Masibay A, Anderson SJ, Barna C, Biega L, Brenneke S (2001). STR primer concordance study. *Forensic Science International*, 124:47-54.
- Butler J (2006). Genetics and genomics of core STR loci used in human identity testing. *J. For. Sci.* 51(2):253-265.
- Butler JM (2005). *Forensic DNA Typing. Biology, Technology and Genetic of STR markers*. Second Edition. Elsevier Academic Press. San Diego, CA.,
- Butler JM (2007). Short tandem repeat typing technologies used in human identity testing. *Biotechniques* 43(4):2-5.
- Butler JM, Buel E, Crivellente F, McCord B (2009). Forensic DNA typing by capillary electrophoresis using the ABI Prism 310 and 3100 genetic analyzers for STR analysis. *Electrophoresis* 25:1397-412.
- Butler JM, Hill CR (2012). Biology and genetics of new autosomal STR loci useful for forensic DNA analysis. *Forensic Sci. Rev.* 24(1):15-26.
- Butler JM, Ruitberg CM, Vallone PM (2001). Capillary electrophoresis as a tool for optimization of multiplex PCR reactions. *Fresenius J. Anal. Chem.* 369:200.
- Butler JM, Ruitberg CM, Vallone PM (2001). Capillary electrophoresis as a tool for optimization of multiplex PCR reactions. *Fresenius J. Anal. Chem.* 369(3-4):200-205.
- Butler JM, Schoske R, Vallone PM, Kline MC, Redd AJ, Hammer MF (2002). A Novel Multiplex for Simultaneous Amplification of 20 Y-Chromosome STR Markers. *Forensic Sci. Int.* 129(1):10-24.
- Butler JM, Schoske R, Vallone PM, Kline MC, Redd AJ, Hammer MF (2002). A Novel Multiplex for Simultaneous Amplification of 20 Y-Chromosome STR Markers. *Forensic Sci. Int.* 129(1):10-24.
- 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.
- Carolina N, Miriam B, Cecilia S, Yolanda C, Jianye G, Bruce B (2010). Reconstructing the Population History of Nicaragua by Means of mtDNA, Y-Chromosome STRs, and Autosomal STR Markers. *Am. J. phys. Anthropol.* 143(4):591-600.
- Carter AB, Salem AH, Hedges DJ, Deegan CN, Kimball B, Walker JA (2004). Genome-wide analysis of the human Alu Yb-lineage. *Human Genomics*, 1:167-178.
- Clayton DA (1982). Transcription of the maintain mitochondrial genome. *Annu. Rev. Biochem.* 53:573-594.
- Coble MD (2004). Single nucleotide polymorphism over the entire MtDNA genome that increase the power of forensic testing in Caucasian . *Int. J. Legal Med.* 5(2):116-127.
- Conrad DF (2011). Variation in genome-wide mutation rates within and

- between human families. *Nature Genetics*. 43(7):712-714.
- De Knijff P, Kayser M, Caglia A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F (1997). Chromosome Y microsatellites: population genetic and evolutionary aspects. *Int. J. Legal Med*. 110:134-140.
- Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Missasseau P, Marc S (1996). A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature*, 380(Suppl A1-A138):152-154.
- Dupuy MS, Egeland M, Olaisen B (2004). Y chromosomal microsatellite mutation rates: differences in mutation rate between and within loci. *Hum. Mut.* 23:117-124.
- Ellegren H (2004). Microsatellites: simple sequences with complex evolution. *Nat. Rev. Genet.* 5:435-445.
- Emrich CA, Tian H, Medintz IL (2002). Microfabricated 384-lane capillary array electrophoresis bioanalyzer for ultra high-throughput genetic analysis. *Anal. Chem.* 74:5076-5083.
- Ensenberger MG, Thompson J, Hill B, Homick K, Kearney V, Mayntz-Press KA, Mazur P (2010). Developmental validation of the PowerPlex 16 HS System: an improved 16-locus fluorescent STR multiplex. *Forensic. Sci. Int. Genet.* 4(4):257-264.
- Giulietta DB, Ivane SN, Michele S, Lorendana N (2000). Mitochondrial DNA sequences in prehistoric human remains from the Alps. *Eur. J. Hum. Genet.* 8:669-677.
- Goldberg SM, Johnson J, Busam D, Feldblyum T, Ferreira S (2006). A Sanger/pyrosequencing hybrid approach for the generation of high-quality draft assemblies of marine microbial genomes. *Proc. Natl. Acad. Sci.*, 103:11240-11245.
- Goldstein DB, Ruiz LA, Cavalli-Sforza LL, Feldman MW (1995). An evaluation of genetic distances for use with microsatellite loci. *Genet.* 139(1):463-471.
- Grignani GP, Fattorini P, Previdere C (2000). Highly informative Y-chromosomal haplotypes by the addition of three new STRs DYS437, DYS438 and DYS439. *Int. J. Legal Med.* 114:125-129.
- Guntheroth WG, Lohmann R, Spiers PS (2006). Risk of sudden infant death syndrome in subsequent sibling. *J. Pediatr.* 116(4):520-524.
- Hacia JG, Fan JB, Ryder O, Jin L, Edgemon K, Ghandour G, Mayer R, Sun B (1999). Determination of ancestral alleles for human single-nucleotide polymorphisms using high-density oligonucleotide arrays. *Nat. Genet.* 22:164-167.
- Hammer MF (1995). A recent common ancestry for human Y chromosome. *Nature*, 378:376-378.
- Hammond HA, Jin L, Zhong Y, Caskey CT, Chakraborty R (1994). Evaluation of 13 short tandem repeat loci for use in personal identification applications. *Am. J. Hum. Genet.* 55:175-189.
- Hanson EK, Ballantyne J (2007). An Ultra-High Discrimination Y Chromosome Short Tandem Repeat Multiplex DNA Typing System. *PLoS ONE*; 2(8):e688.
- Helgason A, Hrafnkelsson B, Gulcher JR (2003). A populationwide coalescent analysis of Icelandic matrilineal and patrilineal genealogies: evidence for a faster evolutionary rate of mtDNA lineages than Y chromosomes. *Am. J. Hum. Genet.* 72:1370-1388.
- Holland MM, Parsons TJ (1999). Mitochondrial DNA sequence analysis -validation and use for forensic casework, *Forensic Science Reviews*, 11:21-48.
- Iafate AJ, Feuk L, Rivera M, Listewnik ML, Donahoe PK, Qi Y (2004). Detection of large-scale variation in the human genome. *Nature Genetics*, 9:949-951.
- Imad H, Abeer F, Cheah Y, Mohammed J, Aamera O (2013b). Discovery of Three Newly Described Single Nucleotide Polymorphisms in Mitochondrial DNA Hypervariable Region I (HVI) and Estimation of Variants and Haplotypes Encompassing Nucleotide Positions 16024-16365. *J. Forensic Res.* 5(1):1-6.
- Imad H, Cheah Q, Mohammad J, Aamera O (2013a). Genetic variation of 17 Y-chromosomal short tandem repeats (STRs) loci from unrelated individuals in Iraq. *Int. J. Biotechnol. Mol. Biol. Res.* 4(8):119-129.
- Imad HH, Ameer IA, Mohammed AJ, Cheah YK, Aamera JO (2014a). Haplotypes and variable position detection in the mitochondrial DNA coding region encompassing nucleotide positions 10,716-11,184. *Mitochondrial DNA*. 1-6.
- Imad HH, Ameer IA, Mohammed AJ, Cheah YK, Aamera JO (2014b). Genetic variation of twenty autosomal STR loci and evaluate the importance of these loci for forensic genetic purposes. *Afr. J. Biotechnol.* 13:1-9.
- Ingman M, Gyllensten U (2003). Mitochondrial Genome Variation and Evolutionary History of Australian and New Guinean Aborigines. *Genome Res.* 13(7):1600-1606.
- Jobling MA, Pandya A, Tyler-Smith C (1997). The Y chromosome in forensic analysis and paternity testing. *Inter. J. Legal Med.* 110:118-124.
- Jorgenson E, Tang H, Gadde M, Province M, Leppert M, Kardia S, Schork N, Cooper R (2005). Ethnicity and human genetic linkage maps. *Am. J. Human Gen.* 76:276-290.
- Kayser M, Caglia A, Corach D, Fretwell N, Gehrig C, Graziosi G (1997). Evaluation of Y-chromosomal STRs: a multicenter study. *Int. J. Legal Med.* 110:125-133.
- Klitschkar M, Immela UD, Kleiber M, Wiegand P (2006). Physical location and linked genes of common forensic STR markers. *Inter. Congress Series*, 1288:801-803.
- Klitschkar M, Immel UD, Stiller D, Kleiber M (2005). TH01, a tetrameric short tandem repeat locus in the tyrosine hydroxylase gene: association with myocardial hypertrophy and death from myocardial infarction? *Dis Markers* 21:9-13.
- Kraytsberg Y, Schwartz M, Brown TA, Ebralidse K, Kunz WS, Clayton DA, Vissing J, Khrapko K (2004). Recombination of Human Mitochondrial DNA. *Science* 304:981.
- Kuppareddi B, Suhasini G, Vijaya M, Kanthimathi S, Nicole M, Martin T (2010). Y Chromosome STR Allelic and Haplotype Diversity in Five Ethnic Tamil Populations from Tamil Nadu, India. *Leg. Med.*, 12:265-269.
- Kwak KD, Jin HJ; Shin DJ, Kim JM, Roewer L, Krawczak M (2005). Y-Chromosomal STR Haplotypes and Their Applications in Forensic and Population Studies in East Asia. *Int. J. of Leg. Med.* 119(4):195-201.
- Lander ES, Linton LM (2001). Initial sequencing and analysis of the human genome. *Nat.*, 409(6822):860-921.
- Litt M, Luty JA (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am. J. Human Gen.* 44:397-401.
- Lutz-Bonengel S, Ulrike S, Tanja S, Stefan P (2003). "Sequence Polymorphisms Within the Human Mitochondrial Genes MTATP6, MTATP8 and MTND4." *Inter. J. Legal Med.* 117(3):133-142.
- Mats O, Mo H, Erik W, Tonia S, Natasha L, Tobias U (2007). Mating system variation and morph fluctuations in a polymorphic lizard. *Mol. Ecol.*, 16:5307-5315.
- Mohammad A, Imad H (2013). Autosomal STR: From locus information to next generation sequencing technology. *Res. J. Biotechnol.* 8(10):92-105.
- Nakamura Y, Leppert M, O'Connell P, Wolff R, Holm T (1987). Variable number of repeat (VNTR) markers for human gene mapping. *Science*, 235:1616-1622.
- Pareek CS, Smoczynski R, Tretyn A (2011). Sequencing technologies and genome sequencing. *J. Appl. Genet.* 52: 413-435.
- Park MJ, Lee HY, Chung U, Kang SC, Shin KJ (2007). Y-STR analysis of degraded DNA using reduced-size amplicons. *J. Leg. Med.* 121(2):152-157.
- Parson W, Niederstätter H, Brandstätter A, Berger B (2003). Improved Specificity of Y-STR Typing in DNA Mixture Samples. *Int. J. Leg Med.* 117(2):109-114.
- Pascali VL, Dobosz M, Brinkmann B (1999). Coordinating Y-chromosomal STR research for the courts. *Int. J. Legal Med.* 112: 1.
- Pastore GM (1994). Leigh syndrome and hypertrophic cardiomyopathy in an infant with a mitochondrial DNA point mutation (T8993). *Am. J. Med. Genet.* 50:265-271.
- Pierce KE, Wangh LJ (2007). "Linear-after-the-exponential polymerase chain reaction and allied technologies Real-time detection strategies for rapid, reliable diagnosis from single cells". *Methods Mol Med. Methods in Molecular Medicine™* 132:65-85.
- Redd ABA, Kearney VA, Contreras VA, Karafet T, Park H, de Knijff P, Butler JM, Hammer MF (2002). Forensic value of 14 novel STRs on the human Y chromosome. *Forensic Sci. Int.* 130:97-111.
- Reynolds R (2000). Detection of sequence variation in the HV2 region of human mitochondrial genome in 689 individuals using immobilized

- sequence specific oligonucleotide probe. *J. Forensic Sci.* 45(6):1210-1231.
- Roach JC (2010). Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science* 328:636-639.
- Rocheleau M, Grey R, Chen D (1992). Formamide modified polyacrylamide gels for DNA sequencing by capillary gel electrophoresis. *Electrophoresis* 13:484.
- Roewer MK, Willuweit S, Nagy M, Alves C, Amorim A, Anslinger K (2001). Online reference database of European Y chromosomal short tandem repeat (STR) haplotype. *Forensic Sci. Int.* 118:106-113.
- Ruitberg CM (2001). STRBase: a short tandem repeat DNA database for the human identity testing community. *Nucleic Acids Res.* 29(1):320-322.
- Schon EA, DiMauro S, Hirano M (2012). Human mitochondrial DNA: roles of inherited and somatic mutations. *Nat. Rev. Genet.* 13:878-890.
- Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P (2004). Large-scale copy number polymorphism in the human genome. *Science*, 305(5683):525-528.
- Seielstad M, Bekele E, Ibrahim M, Toure A, Traore M (1999). A view of modern human origins from Y chromosome microsatellite variation. *Genome Res.* 9:558-567.
- Taanman JW (1999). The mitochondrial genome: structure, transcription, translation and replication. *Biochim. Biophys. Acta.* 1410:103-123.
- Tseng W, Chang H (2001). A new strategy for optimizing sensitivity, speed and resolution in capillary electrophoretic separation of DNA. *Electrophoresis* 22:763.
- Tully G (2004). Results of collaborative study of the EDNAP group regarding mitochondrial DNA heteroplasmy and segregation in hair shafts. *Forensic Science Int.* 140(1):1-11.
- Tzen C, Wu T, Liu H (2001). Sequence polymorphism in the coding region of the mitochondrial genome encompassing position 8389-8865. *Forensic Science International.* 120:204-209.
- Ukhee C, Hwan YL, Ji-Eun Y (2005). Mitochondrial DNA CA dinucleotide repeats in Koreans: the presence of length heteroplasmy. *Int. J. Legal Med.* 119:50-53.
- Underhill PA, Shen P, Lin AA, Jin L, Passarino G, Yang WH, Kauffman E, Bonne-Tamir B, Bertranpetit J, Francalacci P, Ibrahim M, Jenkins T, (2000). Y chromosome sequence variation and the history of human populations. *Nat. Genet.* 26:358-361.
- Van der GM, Tovar J (2005). Degenerate mitochondria. *EMBO Rep.* 6:525-530.
- Walkinshaw M, Strickland L, Hamilton H, Denning K, Gayley T (1996). DNA Profiling in two Alaskan Native Populations Using HLA-DQA1, PM, and D1S80 Loci. *J. Forensic Sci.* 41:478-484.
- Wallace DC, Shoffner JM, Watts RL, Juncos JL, Torroni A (1992). Mitochondrial oxidative phosphorylation defects in Parkinson's disease. *Ann. Neurol.* 32:113-114.
- Weber JL, David D, Heil J, Fan Y, Xhao C, Marth G (2002). Human diallelic insertion/deletion polymorphisms. *Am. J. Human Gen.* 71:854-862.
- White PS, Tatum OL, Deaven LL, Longmire JL (1999). New, male-specific microsatellite markers from the human Y chromosome. *Genomics* 57:433-437.
- Whitfield LS, Sulston JE, Goodfellow PN (1995). Sequence variation of the human Y chromosome. *Nature* 378:379-380.
- Wren JD (2000). Repeat polymorphisms within gene regions: phenotypic and evolutionary implications. *Am. J. Hum. Genet.* 67(2):345-356.
- Yamamoto T, Uchihi R, Nozawa H, Huang XL, Leong YK, Tanaka M (1999). Allele distribution at nine STR loci-D3S1358, vWA, FGA, TH01, TPOX, CSF1PO, D5S818, D13S317 and D7S820 in the Japanese population by multiplex PCR and capillary electrophoresis. *J. Forensic Sci.* 44(1):167-170.
- Zerjal T, Dashnyam B, Pandya A, Kayser M, Roewer L, Santos FR, Schiefenho" vel W, Fretwell N (1997). Genetic relationships of Asian and northern Europeans, revealed by Y-chromosomal DNA analysis. *Am. J. Hum. Genet.* 60:1174-1183.
- Zeviani M, Bertagnolio B, Uziel G (1996). Neurological presentations of mitochondrial diseases. *J. Int. Metab. Dis.* 19:504-520.