HUMAN GENETICS • REVIEW

Current genetic methodologies in the identification of disaster victims and in forensic analysis

Ewa Ziętkiewicz · Magdalena Witt · Patrycja Daca ·
Jadwiga Żebracka-Gala · Mariusz Goniewicz ·
Barbara Jarząb · Michał Witt

Abstract This review presents the basic problems and currently available molecular techniques used for genetic profiling in disaster victim identification (DVI). The environmental conditions of a mass disaster often result in severe fragmentation, decomposition and intermixing of the remains of victims. In such cases, traditional identification based on the anthropological and physical characteristics of the victims is frequently inconclusive. This is the reason why DNA profiling became the gold standard for victim identification in mass-casualty incidents (MCIs) or any forensic cases where human remains are highly fragmented and/or degraded beyond recognition. The review provides general information about the sources of genetic material for DNA profiling, the genetic markers routinely used during genetic profiling (STR markers, mtDNA and single-nucleotide polymorphisms [SNP]) and the basic statistical approaches used in DNA-based disaster victim identification. Automated technological platforms that allow the simultaneous analysis of a multitude of genetic markers used in genetic identification (oligonucleotide microarray techniques and next-generation sequencing) are also presented. Forensic and population databases containing information on human variability, routinely used for statistical analyses, are discussed. The final part of this review is focused on recent developments, which offer particularly promising tools for forensic applications (mRNA analysis, transcriptome variation in individuals/populations and genetic profiling of specific cells separated from mixtures).

Keywords DNA profiling · Genetic markers in DVI · Statistical analysis in DVI · Automated technological platforms · Polymorphisms databases · Transcriptome analysis · DNA mixtures analysis

E. Ziętkiewicz (✉) · P. Daca · M. Witt
Institute of Human Genetics, Polish Academy of Sciences,
Poznań, Poland
e-mail: zietkiee@man.poznan.pl

M. Witt
Department of Disaster Medicine, University of Medical Sciences,
Poznań, Poland

J. Żebracka-Gala · B. Jarząb
Cancer Center and Institute of Oncology,
Gliwice, Poland

M. Goniewicz
Emergency Medicine Unit, Medical University,
Lublin, Poland

M. Witt
International Institute of Molecular and Cell Biology,
Warsaw, Poland

Introduction

Disasters have accompanied humankind since time immemorial. A disaster, as defined by the World Health Organization (WHO), is "a sudden ecological phenomenon of sufficient magnitude to require external assistance". It is important that, in this definition, the emphasis is on the inability to provide adequate assistance to all at need, rather than on the absolute number of victims.

The traditional classification divides disasters into natural and man-made ones. The first group includes floods, natural fires, avalanches, earthquakes, tornadoes, volcanic eruptions, droughts etc., while the events which fall into the other group include transportation disasters (e.g. road traffic accidents, rail, air and maritime transport disasters), construction disasters, fires and wars.
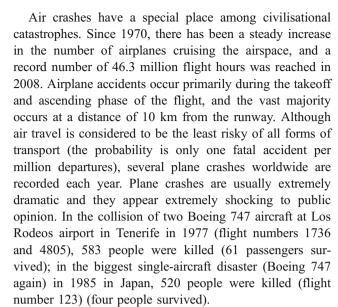
Changes taking place in the modern world make the use of weapons of mass destruction (e.g. biological, chemical)

an increasingly likely cause of disaster, both in an armed conflict and a terrorist act. These two events, war and terrorism, are examples of intentional human actions; the other examples are civilisational disasters which should be considered as the price which today's society has to pay for what can be broadly defined as technological progress.

As early as 1628 BC, the eruption of Santorini in the Aegean Sea was recorded—the biggest explosion since the birth of human civilisation. Another was the eruption of Vesuvius in 79 BC, which produced thousands of victims at Pompeii and Herculaneum. The recent 2,000 years of human civilisation also abounded in various types of disasters. Among them, one may mention the plague epidemic, which claimed 75 million people in the years 1347–1351, explosions of volcanoes in Indonesia in 1815 and 1833 (respectively, 90,000 and 40,000 deaths) or events in China, like flooding along the river Huang-Ho in 1887 (one million victims), land slides in the province of Kansu in 1920 (180,000 killed) or the blowing up of a dam on the Yang-tse-kiang in 1938 (890,000 deaths). In the more recent times, the worst natural disasters include: the cyclone in Bangladesh in 1970 (one million victims), the earthquake in Tangsham, China, in 1976 (750,000 deaths) and the earthquake in Haiti in January 2010 (230,000 deaths). The most recent disaster occurred in 2011 in northeastern Japan, where an earthquake with a force of about 9 degrees on the Richter scale was followed by a tsunami. As a result, 28,000 people were killed or unaccounted for. A nuclear power plant in Fukushima province was also severely damaged in the aftermath. Because of the radiation hazard in the area within a radius of 20 km from the power station, there was an urgent need to evacuate the residents. Four months later, repeated shocks complicated the situation even further, although they did not cause any new fatalities. Other recent man-made disasters include a chemical catastrophe in Bhopal, India, in 1984 (3,500 killed), the explosion of the Chernobyl nuclear reactor in Ukraine in 1986 (between 31,000 and 75,000 killed, depending on the source of information) and the terrorist attack on the Twin Towers of the World Trade Center in New York City on September 11, 2001, which killed nearly 3,000 people and left over 6,000 wounded.

The official record of major disasters in Poland, kept since the 1970s, shows several dramatic events, which challenged national forensic genetic laboratories. These were: a gas explosion in the Rotunda building in Warsaw in 1979 (49 killed, 100 wounded), the collision of two trains near Otłoczyn in 1980 (65 people killed, 64 injured), the sinking of the ferry "Jan Hevelius" near Ruegen in the Baltic Sea in 1993 (55 killed), a fire in the hall of the shipyard in Gdansk during a concert in 1994 (7 people killed, hundreds suffered burns), a gas explosion in an apartment building in Gdansk in 1995 (22 people killed) and the collapse of the roof of the International Fair Hall in Katowice during a pigeon exhibition in 2006 (64 people killed).

Air crashes have a special place among civilisational catastrophes. Since 1970, there has been a steady increase in the number of airplanes cruising the airspace, and a record number of 46.3 million flight hours was reached in 2008. Airplane accidents occur primarily during the takeoff and ascending phase of the flight, and the vast majority occurs at a distance of 10 km from the runway. Although air travel is considered to be the least risky of all forms of transport (the probability is only one fatal accident per million departures), several plane crashes worldwide are recorded each year. Plane crashes are usually extremely dramatic and they appear extremely shocking to public opinion. In the collision of two Boeing 747 aircraft at Los Rodeos airport in Tenerife in 1977 (flight numbers 1736 and 4805), 583 people were killed (61 passengers survived); in the biggest single-aircraft disaster (Boeing 747 again) in 1985 in Japan, 520 people were killed (flight number 123) (four people survived).

Over the last 40 years, there has been a number of fatal Polish airplane crashes: a crash near the Okecie Airport in 1980, with 87 people killed (flight number LO 007), another one in the Kabacki forest near Warsaw in 1987, with 183 people killed (flight number LO 5055), and, especially, the crash of a Polish government delegation flight near Smolensk in 2010 (PLF 101-I-M), in which 96 people lost their lives (among them, the President and his wife, parliamentarians, commanders of the Polish armed forces, ministers and many others).

One of the key issues during the disaster response is the need to identify victims. Disaster victim identification (DVI) can be performed in several traditional ways, like the physical identification of documents, jewellery and other belongings. More reliable methods are associated with the collection of fingerprints, hair and/or dental data.

Until recently, in the vast majority of cases, traditional methods were used for DVI, while DNA-typing played a supporting role or was the last resort when other methods failed. Currently, DVI is based mainly on molecular methodology, which is considered to be the most effective approach in victim identification, both in small-scale disasters and in mass-casualty events.

Early molecular biology procedures used for DVI were based on the analysis of protein or blood group polymorphisms. Most of these systems (with the exception of human leukocyte antigen [HLA]) suffered from a low degree of discrimination and, in addition, their use was limited because of an insufficient amount of genetic information in many tissues and low stability in biologically degraded forensic samples (Budowle and van Daal 2008). In 1985, a quarter of a century ago, the first paper describing the genetic identification of human individuals by DNA fingerprint analysis was published (Jeffreys 1985). This original approach, based on minisatellite hybridisation,

came to be known as DNA fingerprint analysis. Since then, DNA analysis has undergone rapid evolution (Decorte 2010). The developments were instigated by the need for more specific and more sensitive methods allowing identification based on less material of lower quality. The introduction of polymerase chain reaction (PCR)-based techniques allowed the targeting of a variety of genetic markers, from single-nucleotide polymorphisms (SNP) to variable number of tandem repeats (VNTRs); VNTRs include minisatellites (repeated motif of 6–100 bp) and microsatellites (repeated motif of 1–5 bp), often referred to as short tandem repeats (STRs) or simple sequence repeats (SSRs). The current methodology in genetic identification relies on the use of direct analysis of the DNA sequence and the availability of numerous genetic databases.

Human identity testing may be considered in a variety of contexts: forensic casework, analysis of relationship, paternity testing, missing person investigation and mass-casualty incident (MCI) victim identification. Although the ultimate goal is to obtain a match between two persons or between a biological material and a person, the specific context of each of these applications of human identity testing has its specific problems, ranging from technical approach, through statistical interpretation, to ethical issues. The main focus of this review is on genetic profiling in DVI.

Basic problems of genetic profiling in disaster victim identification (DVI)

The environmental conditions of an MCI often result in severe fragmentation, decomposition and intermixing of the remains of victims. In such situations, traditional identification methods, based on the anthropological and physical characteristics of the victims, fail. This is the reason why DNA profiling became the gold standard for victim identification in MCI or any forensic cases when, for various reasons, the human remains are highly fragmented and/or degraded beyond recognition.

The primary goals of DNA profiling are determined by the specificity of a catastrophe and may involve:

–   Identification of the victims (by comparing to DNA profiles from reference samples);
–   Association of body parts (sometimes from highly fragmented remains);
–   In certain cases, the identification of perpetrators (often from contact traces).

Depending on the primary goal, the procedures vary to a considerable extent. For example, the identification of a small part of the recovered remains is enough to identify the person, but it does not satisfy the need to "reassemble"

the victim's body for their family. On the other hand, the DNA-based association of body parts does not assure the personal identification of a victim if the reference DNA profile is not available.

The technical approach adopted in the process of DVI has to take into account the possible problems which can be associated with DNA profiling. Most of the difficulties are related to:

–   Scarcity and/or degradation of the analysable genetic material;
–   Non-homogeneous character of the material;
–   Lack of informative reference samples.

The scarcity of DNA available for the analysis may reflect the critically small amount of material used for DNA extraction. The source material may come in the form of contact traces, for example, on personal items of a victim (to be used as reference samples) or on forensic samples (meant to reveal the identity of a perpetrator). This problem influences the choice of extraction methods, of DNA profiling techniques and, to some extent, of markers to be analysed.

The scarcity of analysable DNA may also be related to its degradation, which depends on the quality of the remains, related to the circumstances of the catastrophe, to the time elapsed since the victim's death and to the environmental conditions. The degree of degradation also depends on the analysed tissue; for example, DNA from bones is usually less degraded than that from blood or soft tissues; DNA from blood is less degraded than that from saliva etc. (Dixon et al. 2006). If DNA degradation is very serious, it may present a serious challenge to DVI in an open type of a catastrophe. An inverse relation between the degradation and the length of DNA sequence that can be successfully and reproducibly analysed influences the choice of DNA profiling technology and of markers to be analysed (Dixon et al. 2006).

The non-homogeneous character of the genetic material in DVI depends on the circumstances of a catastrophe. In the case of severe fragmentation of victims' bodies, it may compromise the association of body parts, while still allowing the identification of the victims. The lack of material homogeneity poses the most serious problems when identification is to be based only on the analysis of contact traces (e.g. in the identification of perpetrators). The mixed DNA profiles of many individuals have a direct bearing on the choice of markers to be analysed and on the statistical methodology for the assessment of results.

Genetic methods used for the forensic identification of individuals rely on the comparison of the material under investigation with reference samples. Reference samples can take the form of biological material or the victim's personal use items previously acquired (used in direct matching analysis) or material obtained from the victim's known biological relatives (used in kinship analysis) (National
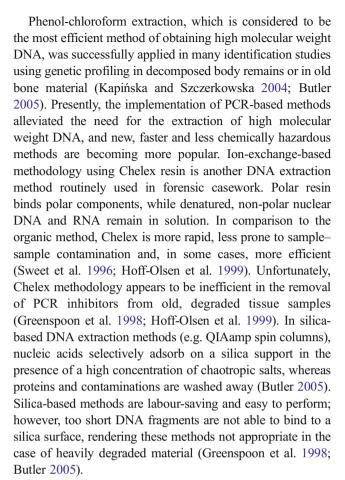
Institute of Justice 2006). In cases aiming to confirm the identity of a victim in a "closed" type incident, obtaining reference samples is relatively simple. The lack of informative reference samples presents the most serious challenge to DVI in "open" type catastrophes involving a large and precisely unknown number of victims of unknown identity. In some cases, solving this problem can be aided by using the information stored in a variety of specialised genetic databases (containing DNA information of blood, sperm or tissue donors, missing persons, criminals or victims etc.). An alternative way to shed light on the genetic origin of an unidentified biological sample is to compare it with population-specific data. While not direct, this approach allows the focus of further investigation to be narrowed. An efficient approach to the genetic identification of individuals based on population-specific data requires taking into account general information on human genetic diversity and strongly affects the choice of markers to be analysed.

Sources of material and extraction methods for genetic profiling

Any type of biological remains (including soft tissue, bone, saliva and hair) can be used as a source of genetic material for DNA profiling (e.g. Mandrekar et al. 2002). However, the quality of the genetic material to be analysed is strongly influenced by the time since death and environmental conditions. When using degraded and limited samples, extraction methods have to be optimised for maximum yield and typing success. The International Society for Forensic Genetics (ISFG), building on experiences from recent mass disasters, formulated recommendations concerning practical procedures and guidance in the collection and storing of ante-mortem and post-mortem samples suitable for the DNA-based identification of victims (Prinz et al. 2007). Obviously, the need for optimisation is even more pronounced in cases when transcriptome profiling is considered (Ibberson et al. 2009).

Factors such as high temperature, contamination of remains with dirt or jet fuel, the putrefaction process or microbiological infestation, may significantly impair DNA quality. In comparison to soft tissues, hard tissues are greatly resistant to autolysis and decay caused by environmental factors, which is why bones, teeth and nails are the only source of DNA in some forensic cases (Graw et al. 2000). The choice of an isolation method depends on the type of biological remains and the degree of body decay. A number of methods for DNA extraction used in forensic applications have been tested for their performance in various tissue types (from bones to soft tissues) and under different environmental conditions (e.g. moisture, air drought) (Clayton et al. 1995; Greenspoon et al. 1998; Hoff-Olsen et al. 1999; Butler 2005).

Phenol-chloroform extraction, which is considered to be the most efficient method of obtaining high molecular weight DNA, was successfully applied in many identification studies using genetic profiling in decomposed body remains or in old bone material (Kapińska and Szczerkowska 2004; Butler 2005). Presently, the implementation of PCR-based methods alleviated the need for the extraction of high molecular weight DNA, and new, faster and less chemically hazardous methods are becoming more popular. Ion-exchange-based methodology using Chelex resin is another DNA extraction method routinely used in forensic casework. Polar resin binds polar components, while denatured, non-polar nuclear DNA and RNA remain in solution. In comparison to the organic method, Chelex is more rapid, less prone to sample–sample contamination and, in some cases, more efficient (Sweet et al. 1996; Hoff-Olsen et al. 1999). Unfortunately, Chelex methodology appears to be inefficient in the removal of PCR inhibitors from old, degraded tissue samples (Greenspoon et al. 1998; Hoff-Olsen et al. 1999). In silica-based DNA extraction methods (e.g. QIAamp spin columns), nucleic acids selectively adsorb on a silica support in the presence of a high concentration of chaotropic salts, whereas proteins and contaminations are washed away (Butler 2005). Silica-based methods are labour-saving and easy to perform; however, too short DNA fragments are not able to bind to a silica surface, rendering these methods not appropriate in the case of heavily degraded material (Greenspoon et al. 1998; Butler 2005).

DNA purification kits based on magnetic beads have become more dominant in the forensic laboratory in the last few years. This technique relies on DNA binding to the magnetic beads in an appropriate ionic charge of the environmental surroundings; all un-bound contaminations are washed away. Systems designed specifically for forensic casework, for example, the DNA IQ™ System (Promega) or the PrepFiler™Automated Forensic DNA Extraction Kit (Life Technologies), enables DNA extractions from very small quantities of material, like bloodstain, saliva or semen stain. The PrepFiler™Automated Forensic DNA Extraction Kit were also validated following the SWGDAM Guidelines (http://www.promega.com/products/pm/genetic-identity/dna-iq/dna-iq/; http://marketing.appliedbiosystems.com/images/Product_Microsites/PrepFiler/PDF/prepfiler_autovalidation_bulletin.pdf). The biggest advantages of the mentioned magnetic beads methodology is labour-saving procedures, as well as the possibility of automatisation, which is not only an attractive alternative to labour-intensive manual methods, but it is also is more resistant against contamination or human error. A variety of manufacturers offer automated machines for the extraction of DNA from forensic samples. Regardless of the manufacturer, all systems allow robust, fast and simultaneous isolation of several samples. Good DNA

recovery from low-yield samples using BioRobot EZ1 was lately reported by different research groups (Nagy et al. 2005; Kishore et al. 2006).

It is worth mentioning that the isolation of a relatively high amount of total DNA from biological material does not always guarantee the positive determination of a proper genetic profile. Genetic material extracted from human remains, in addition to human DNA, often contains bacterial or fungal DNA, which can act as inhibitors during the PCR assay (Golenberg et al. 1996; Greenspoon et al. 1998; Schmerer et al. 1999).

Genetic markers used in DVI

The choice of genetic markers for DVI consistent with the requirements discussed above reflects the fact that the variability of the human genome is not uniform. This trivial statement refers to a number of markers' characteristics, ranging from the distribution of markers in the genome, their informativity and population specificity, to the technical issues of robustness to degradation processes and amenability to multiplex and automated analysis. We further discuss the characteristics of different markers and the technical approaches presently used in personal genetic identification.

Short tandem repeats (STR) markers

The most frequently applied methodology of personal genetic identification relies on STR profiling. The number of alleles of the forensically relevant STR loci ranges usually from 5 to 20 common alleles (Budowle and van Daal 2008), and these STRs are characterised by high values of polymorphic informativity content (PIC) and have a high power of individual discrimination (Chakraborty and Kidd 1991). The multiallelic nature of STRs is a great advantage in analysing mixtures, since, in many cases, alleles from multiple contributors can be distinguished by means of their non-overlapping size. STRs can be analysed in a multiplex fashion, with as many as 16 loci amplified and typed simultaneously (Krenke et al. 2002; Collins et al. 2004), therefore, allowing the reduction of the amount of analysed material. Early work with STRs involved detection on silver-stained polyacrylamide gels, while the currently used technology is based on fluorescence detection methods involving gel or capillary electrophoresis and, more recently, ABI gel-based DNA sequencers (Butler 2006).

The Federal Bureau of Investigation (FBI) uses, and has done since 1997, a standard set of 13 specific STR loci, collectively referred to as CODIS markers (TPOX, D3S1358, D5S818, FGA, CSF1PO, D7S820, D8S1179, TH01, VWA, D13S317, D16S539, D18S51 and D21S11), and the sex-differentiating amelogenin locus AMEL-X/Y (Butler 2006).

The European Network of Forensic Science Institutes (ENFSI) and the Council of the European Union (2001) recommend a European Standard Set (ESS), which includes seven loci: D3S1358, FGA, D8S1179, TH01, VWA, D18S51 and D21S11 (http://www.promega.com/resources/articles/profiles-in-dna/2009/expansion-of-the-european-standard-set/). ESS forms the core STR system used in European national DNA databases; together with the amelogenin locus, it is equivalent to the Interpol standard set of loci, ISSOL (http://www.dna-conference.eu/ppt/Van%20der%20Beek.pdf). The UK and much of Europe utilise the so-called second-generation multiplex plus (SGMPLUS) profile, consisting of eight loci from the CODIS core set (FGA, TH01, VWA, D3S1358, D8S1179, D16S539, D18S51 and D21S11), along with two additional markers, D2S1338 and D19S433, and with the amelogenin locus (Butler 2006). A variety of commercial kits are available for robust multiplex amplification of the core STR loci (Butler 2006). For example, the PowerPlex 16 kit (Promega Corporation) amplifies the 13 core loci, two pentanucleotide loci, Penta D and Penta E, and amelogenin (Krenke et al. 2002). The 16plex AmpF⌀STR Identifiler kit (Life Technologies) amplifies the 13 core loci, two tetranucleotide loci, D2S1338 and D19S433, and amelogenin (Collins et al. 2004). The analysis of autosomal STR markers is often complemented by that of Y-chromosome STRs, which are very valuable in kinship/relationship testing; the commercially available Y-filer™ kit types 16 Y-chromosome loci and amelogenin.

To improve the discrimination power as well as the success in and effectiveness of degraded DNA analysis, the ENFSI and the European DNA Profiling Group (EDNAP) published in 2005 a list of three additional new mini-STR loci: D2S441, D10S1248 and D22S1045, as well as two additional poly-morphic loci, D1S1656 and D12S391 (Gill et al. 2006; Hill et al. 2011). According to the ENFSI/EDNAP group's recommendations, new and more robust kits have been developed and introduced, e.g. the PowerPlex® ESX 17 and ESI 17 Systems (Promega Company) that allow the amplification of 17 loci (D1S1656, D2S441, D2S1338, D3S1358, D8S1179, D10S1248, D12S391, D16S539, D18S51, D19S433, D21S11, D22S1045, FGA, TH01, vWA, SE33 and the sex-typing locus amelogenin) and the AmpF⌀STR® NGM™ PCR Amplification Kit (NGM) (Life Technologies), which is partly compatible with AmpF⌀STR® kits to maximise concordance with existing data sets (Hill et al. 2011).

Most of the STR markers used in commercial kits are located on separate chromosomes (or are located over 25 Mb apart). In most kits, full profiles of these independently segregating STR markers provide a power of discrimination ranging from 1 in 10^{10} to 1 in 10^{20}, depending on the number of markers used and the populations under investigation, e.g. estimated random match probabilities for the US Caucasian population, using PowerPlex16 and AmpF⌀STR Identifiler,

were 1 in 10^{18} (Butler 2006). Thanks to this high power of discrimination, STR analysis has proven to be a powerful tool in many mass fatality incidents since the 1990s. In fact, in many instances, STR analysis alone was sufficient to complete the DVI procedure.

In spite of being the most commonly and routinely applied markers in forensics, STRs have some shortcomings, limiting their use in DVI efforts. STRs are particularly informative on well-preserved soft tissue and bone samples. In original settings, the amplicon size was too long (150–450 bp) to allow efficient amplification from degraded DNA templates (Gill 2002). The development of mini-STRs (60–80 bp in length), obtained by designing primers located closer to the tandem repeat sequence, largely solved this problem (Wiegand and Kleiber 2001; Butler et al. 2003; Coble and Butler 2005), while maintaining consistency with the existing core loci. However, the template size needed for the analysis (even using mini-STRs) may still be too long to successfully analyse heavily degraded samples.

Another shortcoming of using STRs is that further extension of the set of 13–15 core loci, which, in some cases, is necessary in order to achieve reliable discrimination, is not straightforward and the analysis is not fully automated. Other potential technical problems include the presence of triallelic patterns, null alleles, alleles with size not matching the standardised allelic ladders etc. (Butler 2006). The high mutation rate of STR loci (10^{-3}–10^{-5}), while rendering them very informative, also makes them less stable (Huang et al. 2002; Dupuy et al. 2004). The resulting problems with distinguishing the alleles identical by state from those identical by descent compromise the use of STRs in lineage or ethnicity analysis (Butler 2006). Finally, STRs do not provide phenotypic hints regarding the analysed samples.

Single-nucleotide polymorphism (SNP) markers

SNP markers offer a useful and increasingly important extension to a routine STR-based DNA profiling. SNPs provide an almost unlimited source of human genome diversity to be analysed (Cooper et al. 1985; Wang et al. 1998). Obviously, the number of over 30 mln refSNP clusters reported in 2010 (build 132 of the NCBI dbSNP, http://www.ncbi.nlm.nih.gov/projects/SNP/) includes variants which, because of the low frequency of their rare alleles, are not useful for forensic identification. Based on the average nucleotide diversity, estimated for the haploid human genome at 0.1% (Li and Sadler 1991; Reich et al. 2002), one can expect a total of about 3 million bases differentiating haploid genomes of two randomly chosen individuals.

The application of SNP profiling for DNA identification offers some advantages as compared to the use of STR markers, but it also suffers from limitations. These positive and negative aspects of using SNPs overlap, because the required characteristic of a set of SNPs depends on their specific applications (Budowle and van Daal 2008).

An important, general advantage of using SNPs for DVI purposes stems from the possibility of analysing heavily degraded fragments: SNPs can be amplified using DNA templates of less than 60 bp in length (approximately the length of two flanking primers). The biallelic nature of SNPs facilitates their typing, but renders them not very informative for identity testing on a per-locus basis. The informativity may be increased by analysing many unlinked SNPs, but, in order to achieve the level of discrimination typical for the 13 core STR loci (i.e. 10^{-15}–10^{-16}), large panels of SNPs (50–100) need to be genotyped (Chakraborty et al. 1999; Gill 2001; Li et al. 2006). Fortunately, the recent progress in high-throughput technologies enables large multiplex assays and complete automation of the process, so that multiple SNP profiling is cost-effective and standardisable (see *Automated technological platforms in genetic identification*). On the negative side, biallelic SNPs are more difficult for interpretation in situations which involve mixtures of contributing samples, since the deconvolution of a mixed sample into individual profiles is difficult and not reliable (Budowle and van Daal 2008).

An important feature of autosomal SNPs is that they exhibit considerable variability in their heterozygosity levels across the genome. This is likely due to differences in the recombination rate, as reflected by the presence of haplotype blocks characterised by high levels of linkage disequilibrium (Daly et al. 2001; Gabriel et al. 2002; Reich et al. 2002). The extent of LD blocks in diverse human populations varies, providing clues about the evolutionary history of human lineages. On the other hand, genotyping a set of tightly linked SNPs allows converting these biallelic markers into a multiallelic haplotype. Autosomal SNPs within a single LD block form lineage-specific markers, best known from the analysis of maternally inherited mitochondrial DNA (mtDNA) or paternally inherited non-recombining portion of the Y-chromosome (NRY).

SNPs are characterised by a low mutation rate, on the order of 10^{-8}/site/generation (Nachman and Crowell 2000; Reich et al. 2002), which makes them stable genetic markers, with alleles' identity by state reflecting their identity by descent. This is why differences in SNPs' frequency and distribution across ethnically diverse populations may be assumed to represent their distinct evolutionary and demographic histories (Ziętkiewicz et al. 2001, 2003). Most studies of autosomal, X- and Y-chromosome, and mtDNA markers show higher levels of genetic variation in African compared to non-African populations, reflecting the ancient history of the origin of modern humans (Tishkoff and Williams 2002; Zitkiewicz et al. 2001, 2003). Other differences in SNP markers distribution reflect more recent demographic histo-

ries, such as migrations, population bottlenecks, isolation, admixture etc. (Tishkoff and Verrelli 2003), and may be used for inferring the geographic or ethnic origin of an individual.

The possibility of using SNPs for the inference of ethnic origin of the analysed sample, which can help in predicting the geographic origin or certain physical traits of a donor, can be an important asset in DVI efforts, when no match is found between the victim's remains and any living person or available database records. Mitochondrial SNP haplo-types belong to continental-specific haplogroups and can be used to indicate the matrilineal origin of the analysed sample (the complementary patrilineal information most often comes from the analysis of Y-chromosome STRs), but their informativity is limited. The use of ancestry-informative markers (AIMs), autosomal SNPs which are characterised by large inter-population differences of their allele frequencies, offers much better potential for assessing the biogeographical origin of analysed samples (Frudakis et al. 2003; Shriver and Kittles 2004; Phillips et al. 2009; Nassir et al. 2009). Efficient discrimination of ancestry requires genotyping a large number of AIMs, usually using microarray-based technology. Currently available panels of AIM SNPs (73 AIMs DNAWitness Bio-Geographical Ancestry Kit, 320 AIMs Eurasian panel, 1,476 AIMs European panel, references in Budowle and van Daal [2008]) allow addressing continental differences and rela-tive levels of admixture among some populations; addi-tional panels will be required to reliably discriminate between individuals from other specific populations within continents (Budowle and van Daal 2008; Nassir et al. 2009; Decorte 2010).

The complicated genetic background of DVI-relevant physical traits (height, pigmentation, facial features) is the reason why direct prediction of these traits from DNA variability data has very limited success (Sturm 2009; Decorte 2010). The basic approach relies on the identification of SNPs within genomic regions (or genes) associated with a particular trait. Some success has been achieved in revealing the association of SNPs in a number of pigmentation genes with various human hair, skin and eye colour phenotypes: some sequence polymorphisms in the melano-cortin 1 gene are indicators of the red hair phenotype, while genome-wide association studies identified HERC2 as a human iris colour gene (Grimes et al. 2001; Rees 2003; Kayser et al. 2008).

In summary, the informed choice of SNPs for the purpose of genetic identification has to combine informa-tion on the technical and statistical aspects of SNP analysis with that on SNPs' population genetics and, at the same time, taking into account the purpose of the study (Budowle and van Daal 2008; Decorte 2010). A set of unlinked SNPs characterised by high heterozygosity values and low inter-population variance is suitable for personal identification,

where the assessment of match probability is based on the allele frequencies in a general population (Dixon et al. 2005; Kidd et al. 2006; Sanchez et al. 2006). In contrast, in cases where no match is found to relatives, reference samples or personalised databases, using AIM-SNPs, may help to establish the ethnicity of analysed samples, providing important clues for further investigation (Phillips et al. 2009). Finally, when SNP-based identification is to be based on kinship analyses, the markers of choice are SNPs residing on mtDNA or Y-chromosome loci, or lineage-specific autosomal SNPs within genomic haploblocks (Budowle and van Daal 2008).

Mitochondrial DNA (mtDNA) in forensic science

The analysis of nuclear markers may fail to give reliable results, due to the degradation of nuclear DNA in the available material or because of the insufficient amount of DNA in samples like bones, hair shafts and teeth (Parson et al. 2004; Parson and Bandelt 2007; Nilsson et al. 2008). In such cases, mtDNA is a much more convenient marker, due to its higher resistance to degradation and a high copy number per cell. mtDNA analysis was first applied as court evidence in 1996, in the case of Paul William Ware, who was accused of a child rape and murder. The only exhibit found at the crime scene was a single hair in the victim's throat, which, as mtDNA analysis indicated, belonged to the suspect. Since then, mtDNA analysis has become an increasingly frequent tool in forensic criminal casework (Holland and Parson 1999).

Because of the lack of recombination, mtDNA can be analysed as a single, highly informative multiallelic locus (haplotype). This, together with the maternal mode of inheritance, allows using the mtDNA of any maternally related individual as a reference for personal identification (Carracedo et al. 2000; Parson and Bandelt 2007). By convention, mtDNA sequences (haplotypes) are recorded by listing only the differences from the universal reference sequence, also called the Cambridge or Anderson's sequence (Isenberg and Moore 1999).

A big advantage of using mtDNA in forensic applica-tions is its high sequence variation, which results from the high mutation rate in the mitochondrial genome. mtDNA is almost ten times more prone to mutation compared to nuclear DNA. Two hypervariable segments of the control region, HVSI (positions 16,024–16,365) and HVSII (73–340), are the most variable between individuals, and are, therefore, of special interest for forensic studies. Routine investigation is based on a determination of sequence variants of the mtDNA control region (PCR and sequenc-ing). When an mtDNA profile derived from an evidence sample matches that obtained from reference samples, the statistical interpretation of a match significance depends on

the population frequency of the haplotype in question (Parson et al. 2004). If it is rare, the probability that the two samples which share this haplotype are derived from the same maternal line is high, and the evidence obtained in this way has a strong discrimination power (Holland and Parson 1999; Salas et al. 2007). In the European population, where about 7% of the population share the most common HVSI/HVSII type, standard sequencing analysis of the control region is insufficient for forensic discrimination. To improve the discrimination power, SNPs located in the mtDNA coding region, which is more evolutionarily stable, are analysed (Coble et al. 2004, 2006; Nilsson et al. 2008; Daca et al. 2008). As a result of the growing number of completely sequenced mtDNA genomes, a lot of SNPs have been published in the last few years. For instance, population-specific SNP panels can discriminate the population of Eastern Eurasia from that of Western Eurasia. Some of these panels have been optimised and successfully applied in forensic science laboratories (Brandstätter et al. 2003, 2006; Quintáns et al. 2004; Daca et al. 2008).

Interpreting the weight of mtDNA evidence in forensic testing is complicated by the non-homogeneous mutation rate in the mitochondrial DNA genome. In some positions (e.g. 16,093, 16,126, 16,172, 146, 150 and 152), mutations are more frequent than in other sites. Heteroplasmy, the presence of more than one mtDNA type in an individual, is another uncommon feature of the mitochondrial genome, which may complicate the interpretation of forensic evidence. It may be observed at many levels; for example, an individual can be heteroplasmic in one tissue type and homoplasmic in another. Heteroplasmy usually manifests as substitutions, but it can also affect the number of bases residing within a homopolymeric stretch (for example, C stretches) (Carracedo et al. 2000).

To standardise forensic mtDNA analysis, the DNA Commission of the International Society for Forensic Genetics created guidelines for mtDNA typing in a forensic science laboratory. The relevant document regulates standard amplification and sequencing methodology, criteria for inclusion or exclusion in case of mismatches between the analysed samples, as well as general principles of heteroplasmy or indel interpretation (Carracedo et al. 2000; Tully et al. 2001).

Statistical analysis in DVI

Statistics plays an important role in evaluating the evidential weight of DNA profiling, especially in more complicated cases like those involving DNA mixtures and subdivided populations (Fung et al. 2006). The underlying theory and methodology are well developed (e.g. Balding

2005; Buckleton et al. 2005). The choice of an approach used in DNA-based DVI reflects the compromise between the quality of material available for analysis, the availability of the reference material, the informativity of markers analysed and the statistical threshold required to make the identification.

The statistical threshold, which allows the conclusion that a match between DNA profiles from a victim and from a reference sample is not coincidental, may be calculated in many ways (Butler 2006). In DVI work, the odds of a given DNA profile favouring identification are represented as likelihood ratios that permit DNA results to be combined among multiple genetic systems or with other non-DNA evidence. The likelihood ratio estimate is inversely related to the random match probability, which, in turn, depends on the genotype frequencies of the analysed markers, and should be based on population-specific genotype frequencies. The probability that a profile composed of multiple independently segregating markers would accidentally match any individual in the population can be estimated as the product of genotype frequencies for all of these markers. In a homogeneous population, the values of random match probability for a profile composed of 13 STRs range from 1 in 10^{14} to 1 in 10^{21} (Butler 2005). The National Research Council recommends taking population structure into account when estimating random match probabilities; this is especially important when the ethnic origin of the person to be identified is unknown. If the population structure and size are not known, the use of a correction factor is advised: 0.01 for a large population and 0.03 for a small and isolated population where inbreeding takes place. Such a correction causes a two orders of magnitude reduction of the likelihood ratio, but when the initial ratio is high (e.g. 10^{22}), this does not significantly affect the results (Butler 2005).

One of the commonly used statistical approaches is based on estimating Bayesian posterior probability, which combines the odds of a given DNA profile favouring identification with a prior probability. The prior probability describes the chance that any sample of the available remains belongs to a particular individual. Typically, it is inversely related to the estimated number of victims **N** (in other words, the number of victims strongly affects this a priori component, which, only in case of typical paternity cases, is set as 50%). The DNA laboratory has to analyse a set of markers yielding a random match probability that, combined with the prior probability, would meet the required posterior probability. The posterior probability **p** of correctly identifying all N victims is based on the formula $(1-1/L)^N < p$. For example, if the required posterior probability in a catastrophe involving $N=1{,}000$ victims is $p=99.9\%$, the threshold for a random match probability would have to be 1 in 10^9 (Brenner and Weir 2003).

When analysing probabilities, it is essential to avoid using incorrect assumptions. The probability of a match cannot be automatically translated into the probability of identification ("prosecutor's fallacy"); the theoretical probability of a match cannot be directly translated into practical probability; the rates of laboratory errors (especially false-positive probability) have to be taken into consideration; population data have to be considered to compute a match probability etc. Finally, a body part can be identified if its DNA profile has a sufficient similarity to a direct personal reference (DNA profile from a known biological relic of the victim) to an indirect reference (DNA profile of the victim's family member) or to a profile in a database ("cold hit"); all of these situations require different statistical approaches (Brenner and Weir 2003).

Automated technological platforms in genetic identification

A high level of multiplexing in genetic profiling is of special importance in DVI, where the aim is to analyse a large number of informative markers in a limited amount of genetic material. The ultimate level of multiplexing is represented by the whole-genome approaches, allowing the simultaneous analysis of a multitude of genetic markers without the need for their annotation. The recent progress in high-throughput technologies enabled achieving unprecedented levels of affordable automation and multiplexing of the genomic sequence analysis. On the other hand, technologies are being developed which make sequence analysis using single molecule templates possible. All of these developments opened new venues for forensic applications, including DVI. Two main groups of recent technologies include hybridisation to oligonucleotide microarrays and massively parallel sequencing.

Hybridisation to oligonucleotide microarrays

Oligonucleotide microarray techniques (Sobrino et al. 2005) are based on the hybridisation of target nucleic acids to small, high-density arrays, containing tens of thousands of short (~25 bp in length) synthetic oligonucleotides immobilised on a solid support (chips or beads). A prior PCR amplification of the region around the variable site is usually required. The basic SNP genotyping techniques involve allele-specific oligonucleotide hybridisation, primer extension and allele-specific oligonucleotide ligation.

Microarray technology has been first introduced for the transcriptome analysis. Several microarray platforms for RNA analysis have been developed by different companies (e.g. Affymetrix and Illumina). They differ in the chemistry and design of microarray preparation, and in detailed protocols of RNA processing, labelling and hybridisation. Protocols for template preparation typically include the synthesis of cDNA from the total (poly-A) RNA isolated from biological samples and in vitro transcription to produce cRNA, which is then fragmented and used for hybridisation with the microarray. The recent, most updated version of the expression array produced by Affymetrix, GeneChip Human Gene 1.0 ST Array, interrogates 28,869 well-annotated human genes with 764,885 distinct probes. Each gene is represented on the array by approximately 26 probes, spread across the full length of the gene. The most popular HumanHT-12 v4 Expression BeadChips system produced by Illumina contains 12 arrays on a single BeadChip, each with >47,000 probes derived from human genes annotated in the NCBI Reference Sequence and UniGene databases.

The alternative possibility of gene expression analysis, accessible on the Illumina platform, DASL procedure (cDNA mediated annealing, selection, extension and ligation), is designed mainly for the analysis of partially degraded RNA from formalin-fixed, paraffin-embedded (FFPE) samples. Total RNA is reverse-transcribed to cDNA; unique pairs of query oligos (one of which with the unique gene address) are annealed to the cDNA and subjected to primer extension and ligation. The products are then PCR amplified in high multiplex, using common primers, and hybridised (by means of the address sequence in the query oligo) to the HumanRef-8 Expression BeadChip designed to interrogate specific genes. This procedure could be of special interest for forensic applications, as RNA fragments as small as 60 bp may be addressed (Haller et al. 2006). HumanRef-8 Expression BeadChip allows the interrogation of a specific set of 512 genes.

Microarray technology was later adapted for the genome analysis, in particular, for SNPs genotyping and for resequencing. There are many methods and commercial platforms for SNP analysis. Here, we comment only on those achievable on the Illumina and Affymetrix platforms. Illumina offers BeadChips targeting from 300,000 to 2.5 million SNPs. The most advanced Affymetrix array, the Genome-Wide Human SNP Array 6.0, contains probes for about 906,600 SNPs. The sample preparation protocols are similar for both platforms, although the amount of DNA is about 2.5 times lower for Illumina (about 200 ng sample). An important difference concerns the number of samples which can be analysed during a single run: Illumina enables the analysis of 2 to 12 samples, while Affymetrix allows the examination of only one sample.

Targeted resequencing using microarrays consists of microarray hybridisation, used for the template preparation, followed by new-generation sequencing techniques (see the next section). Illumina offers two kits for targeted resequencing: TruSeq DNASamplePrep and TruSeq ExomeEn-

richment. The procedure includes preparing a SureSelect capture library and processing of the DNA sample. After a few cycles of DNA amplification, the enriched library is loaded on an Illumina flow cell and undergoes standard clonal amplification on the Illumina Cluster Station. After that, clusters are generated and a flow cell is transferred to the Genome Analyzer for massively parallel sequencing by synthesis. Some kits offered by Affymetrix, like the GeneChip Human Mitochondrial Resequencing Array 2.0, are based on a similar sample preparation protocol.

Methylome analysis is offered exclusively by Illumina. The Infinium Methylation Assay provides quantitative methylation measurement at the single CpG-site level, offering the highest resolution for understanding epigenetic changes.

The DNA analysis using microarray technology has found its application in forensic testing for sequencing or rese-quencing purposes, for SNP genotyping, paternity testing and individual identification. SNP typing by oligonucleotide probe microarrays designed to target 34 different SNPs was evaluated in a sample of Chinese individuals; the study has shown that the microarray assay was useful for paternity testing and individual identification (Li et al. 2006). Combined with the 28 SNPs loci distributed on HLA-DRB1 and ABO genes (Li et al. 2004), the analysis of 59 SNP loci provided the discrimination power of 9×10^{-15}, yielding, thus, the same power in forensic identification as the CODIS STRs currently used.

On the other hand, the application of oligonucleotide microarrays for DVI is not free of problems. For example, evaluation of the Affymetrix GeneChip Mitochondrial Resequencing Array for forensic applications in compari-son to dideoxy fluorescent sequencing indicated that the array platform had difficulty calling instances of point heteroplasmy, insertions, deletions, poly C regions and some closely spaced polymorphisms (Vallone et al. 2007). The use of microarrays for identification on the basis of RNA has been only rarely addressed and requires even more research than DNA-based analysis. The potential application of microarray-based analysis of the methylation level for individual identification purposes also awaits evaluation.

The main shortcoming of all of the described platforms is that they contain only previously known, presynthesised probes and, therefore, do not detect all of the nucleic acid sequences present in the analysed biological samples.

Next-generation genome sequencing

Next-generation sequencing (NGS) allows the simultaneous analysis of hundreds of loci or even of the whole genome, without being dependent on genome annotation. NGS produces massively parallel sequence data, in some cases in excess of one billion short reads per instrument run (Tucker et al. 2009; Li et al. 2009a; Harismendy et al. 2009; Pareek et al. 2011). Cloning or template amplification of the DNA fragments that are being sequenced is fully automated and intimately coupled with the sequence reading process. NGS provides an increased depth of sequence coverage, which leads to a higher accuracy of the reads and allows inferences on the copy number of analysed segments (Tucker et al. 2009).

The three best-known NGS platforms are: Illumina Genome Analyzer, ABI SOLID and Roche 454 Genome Sequencer. All of them involve the template amplification step, but they use different sequencing chemistries and amplification approaches (Tucker et al. 2009). Illumina uses bridge-PCR on a solid surface, while SOLID and 454 rely on PCR in emulsion; the sequencing chemistry is polymerase-based (with fluorescently labelled nucleotides), ligation-based and pyrosequencing, respectively. Illumina and SOLID use fluorescent-based signal detection and 454 is based on the detection of light emitted upon pyrophos-phate disintegration. The length of the reads is 75 bp for Illumina GA, up to 50 bp for ABI SOLID and up to 400 bp for Roche 454, while the raw base accuracy is 99.5% for Illumina and Roche, and 99.94% for ABI Solid (Bosch et al. 2008; Tucker et al. 2009).

The newer Helicos technology is a polymerase-based method, which is unique in that it does not include the amplification step and uses a single DNA molecule as a template. The accuracy of the reads (30–35 bp) using this platform is 99% (Braslavsky et al. 2003).

Most of the NGS platforms (except Roche 454) provide very short reads (50–100 bp). While the possibility to read very short sequences is an important advantage for application to degraded DNA material, this characteristics makes sequence assembly of repetitive and rearranged DNA segments more difficult. Short length of the reads also complicates interpretation when it is necessary to determine the phase of variants, as in the case of mixed DNA samples. Currently, NGS is more expensive in compar-ison to the array-based technology. As this may change in the future, a search for potential applications of NGS to DVI is fully justified.

NGS can be used for both genome and transcriptome analysis. In genome analysis, it permits high-quality variant calling for SNPs, insertions and deletions, and allows the analysis of copy number variants (CNVs) and other structural rearrangements. In transcriptome analysis, NGS can be used to annotate coding SNPs, discover transcript isoforms, identify regulatory RNAs, characterise splice junctions and determine the relative abundance of tran-scripts. It allows to generate full sequence information from any poly(A)-tailed RNA, analyse gene expression and identify novel or rare transcripts in one experiment.

Massively parallel NGS of forensically relevant loci could play a significant role in solving criminal cases and the identification of human remains. Targeted sequencing of candidate genes and genomic intervals in four individuals, using Illumina GA, SOLID and 454, demonstrated the suitability of NGS platforms for population-based targeted sequencing studies, although it also indicated the need for further optimisation (Harismendy et al. 2009). The biggest disadvantage of NGS is a necessity of performing PCR amplification that may introduce base sequence errors or favour certain sequences over the other. Third-generation HT-Technology, also known as sequencing from a single DNA molecule, overcomes this problem because it is based on sequencing by synthesis without any amplification step. The available on the market HT-NGS platforms, offered by a number of companies, represent a different technical approach to sequencing by synthesis.

The newer Helicos technology is a polymerase-based method, which is unique in that it does not include the amplification step and uses a single DNA molecule as a template. The accuracy of the reads (30–35 bp) using this platform is 99% (Braslavsky et al. 2003).

SMRT sequencing technology is based on a single-molecule, real-time sequencing. The process of enzymatic incorporation of labelled nucleotides is performed on zero-mode waveguides (ZMWs) located on the SMRT chip. The whole process is captured by a CCD camera (http://www.pacificbiosciences.com/smrt-biology/overview).

A different approach is represented by the Nanopore DNA sequencer and Ion Torrent sequencing technology. Both platforms are based on physicochemical processing that occurs during the DNA sequencing process rather than nucleotide labelling detection. In the Nanopore sequencer, the DNA sequence is determined by the disruption of ionic current in a protein nanopore, while Ion Torrent technology is based on pH change caused by nucleotide incorporation into a DNA strand. In contrast to all of the mentioned platforms, PGM (Ion Personal Genome Machine) sequencer is laser-capture-independent because it uses ion sensors densely spaced on the bottom of a high-density array of micro-machined wells. This new approach allows producing more sequence reads on time and faster gathering of information than light-based technology platforms. It also enables to perform a wider range of sequencing applications, such as multiplexing amplicons, small RNA sequencing, paired-end sequencing and methylation analysis (http://www.iontorrent.com/; Pareek et al. 2011)

Examples of databases used for identification purposes

A multitude of databases are available, gathering information of the human genome variability. While all can be, to some extent, useful in forensics, only some of them, which provide direct links between genetic variation and individuals, can be directly applied in DVI procedures.

STR databases

The Combined DNA Index System, CODIS, stores STR-based DNA profiles of US offenders and is organised in two databases (indexes). The Convicted Offender Index contains DNA profiles of individuals convicted of felony sex offenses and other violent crimes; the Forensic Index contains DNA profiles developed from crime scene evidence. CODIS utilises dedicated computer software to automatically search its two indexes for matching DNA profiles. This technology is authorised by US law (DNA Identification Act of 1994). As of August 2007, CODIS contained over 5 million DNA profiles in its Convicted Offender Index and about 188,000 DNA profiles collected from crime scenes but not connected to a particular offender (http://www2.fbi.gov/hq/lab/codis/clickmap.htm). Interpol and the European laboratories use a common number of STRs (European Standard Set of loci, ESS) in order to allow comparisons across borders. Their records are stored in the National DNA databases (e.g. NDNAD in the UK and FNAEG in France).

The drawback of using forensic databases as the source of reference in DNA profiling for the purpose of DVI is that most of the data concern offenders and/or crime victims. Most of the population never being involved in any forensic investigation would have no matching records.

Population databases of genomic variation

In contrast to the personal information contained in the forensic STR databases, the resources regarding SNP markers are mostly population-oriented. dbSNP stores data on SNP allele frequencies determined in four major populations (http://www.ncbi.nlm.nih.gov/projects/SNP/). The HapMap effort resulted in a large amount of data on LD blocks, characterising SNP-based haplotypes, but this information is again related to populations and not individuals. Some STR databases also have a population-rather than individual-based character, like the database of Y-chromosome STR haplotypes (http://www.hapmap.org/) (Thorisson et al. 2005). All of these resources, although not directly usable for personal identification, may be used in DVI efforts for the DNA-based determination of the individual's ancestry/ethnicity/geographic origin.

To evaluate the population frequency of Y-chromosome haplotypes, the open-source YHRD database is commonly recommended (Kayser et al. 2002; Jobling and Gill 2004). This database contains haplotypes from the global population from many continents, genotyped for a core set of

Y-STR markers (YS19, DYS385, DYS891I/II, DYS390, DYS391, DYS392 and DYS393). The YHRD is divided into sets of populations sharing a common linguistic, genetic or geographic background. Such a structure is very convenient during the statistical evaluation of haplotype matches (Willuweit and Roewer 2007).

Helpful in estimating the frequency of mtDNA profiles are some freely available public databases, which provide information on mitochondrial haplotypes in a number of populations. These databases contain data on total mtDNA sequences (http://www.mtdb.igp.uu.se/), hypervariable regions (HvrBase++) or mtDNA haplotypes (http://www.nationalgeographic.com/genographic; Behar et al. 2007). Another database storing mtDNA haplotypes, EMPOP (http://empop.org/), is referred to as "forensic", since the data were collected in accordance with the requirements of forensic data collection (Parson and Dür 2007). However, unlike CODIS, this database does not allow personal identification.

Some of the public mtDNA profile resources have been heavily criticised for the unacceptably high rate of errors (Bandelt et al. 2002, 2004; Yao et al. 2004). Mistaken analysis or clerical errors introduced in the process of data rewriting (Bandelt et al. 2001; Parson et al. 2004) result in "phantom" polymorphisms, which disturb the assessment of the population frequency of a particular mtDNA haplotype. Although sources of these errors are well known, problems still persist (Parson and Bandelt 2007). To solve the problem of artifactual recombinations, phylogenetic analysis may be applied, based on the knowledge of evolutionary patterns in human mitochondrial genome (Parson and Bandelt 2007; Salas et al. 2007).

The EMPOP database, recommended by the European DNA Profiling (EDNAP) group, appears to be the most inspected among the mtDNA databases. It contains about 12,785 mtDNA profiles from all of the continents, and consists of forensic datasets obtained from forensic laboratories and literature data; all of the sequences have been precisely verified in order to prevent the release of erroneous data (Parson and Dür 2007; Zimmermann et al. 2011). Unlike many other databases, the EMPOP database contains an application of quasi-median (QM) networks, which is a powerful tool in the quality testing of mtDNA data by graphically representing the genetic structure of the lineages in a data set (Parson and Dür 2007; Zimmermann et al. 2011).

A database's size and coverage is another issue which is very important in forensic identification based on mtDNA analysis. Due to the very high level of variability in mtDNA sequences, large databases are required to achieve good representation of mtDNA diversity. Too small haplotype coverage in a database, representing only a small fraction of population diversity, results in wrong statistical

evaluation. This is especially important in a situation when the suspect's mtDNA haplotype is common in the relevant geographical region, but has no appropriate representation in the database (Salas et al. 2007). Analyses of Portuguese, German and Polish populations indicated a lack of saturation with respect to mtDNA haplotypes, using sample sizes of 549 from the Portuguese population, 1,200 from the German population and 1,253 from the Polish population (Pfeiffer et al. 2001; Pereira et al. 2004; Daca et al. 2010). Even a large mtDNA database of the Genographic Population Project, where 78,590 samples from distinct regions of the world are stored, indicated a lack of full haplotype coverage (Behar et al. 2007). This shows that a constant database build-up is required for achieving an appropriate representation of mtDNA haplotype diversity in populations and that sample sizes used for forensic purposes need to be very large.

Recent developments

New biotechnological methodologies and technologies are constantly being developed. While their direct application for DVI is often not possible at the present moment, these alternative directions are worthy to be considered in the context of developing new approaches for future efforts aiming to identify individuals or to assess their population affiliation.

mRNA analysis in present forensic applications

Successful recovery and cDNA amplification of mRNA extracted from post-mortem tissues (Phang et al. 1994; Cummings et al. 2001) or from tissue samples stored for many years (Matsubara et al. 1992; Zhang and McCabe 1992; Karlsson et al. 2003) demonstrated the potential of RNA for forensic studies. Real-time PCR results using RNA extracted from various tissues indicated that RNA degradation did not correlate with the post-mortem interval, but, rather, with other parameters, like environmental impact and circumstances of death (Li et al. 2004; Heinrich et al. 2007). Today, body fluid identification (blood, saliva, semen, vaginal secretion and menstrual blood) relying on the quantification of specific mRNAs is an established technique in many forensic laboratories (Juusola and Ballantyne 2005; Nussbaumer et al. 2006; Zubakov et al. 2008); other possible applications (reviewed in Bauer 2007) include wound age estimation (Takamiya et al. 2003) and determination of the age of biological stains by means of RNA decay analysis (Bauer et al. 2003; Anderson et al. 2005). Stable RNA markers were ascertained for the identification of tissue specificity in old, degraded samples (Zubakov et al. 2008, 2009).

Recently, the advances in multiplex real-time PCR technology and the introduction of high-throughput sequence analysis platforms enabled the quantitative profiling of the whole transcriptome. Importantly, quantitative analysis of the transcriptome (using Illumina or Affymetrix arrays addressing thousands of expressed genes) opened the gates for the genome-wide analysis of differential levels of gene expression in individuals and populations.

Transcriptome variation in individuals and populations

Many studies during the last decade have shown that gene expression levels in humans differ not only among cell types, but also among individuals. Allelic variation in human gene expression has already been reported in 2002 (Yan et al. 2002). The following years brought a number of microarray-based gene expression studies, addressing several hundred to thousands of transcripts and performed in tens to hundreds of human lymphoblastoid cell lines (LCLs) derived from individuals of European ancestry. These studies have shown that variation in human gene expression is extensive, has a strong heritable component and maps to many chromosomes across the human genome (Cheung et al. 2003; Monks et al. 2004; Morley et al. 2004; Cheung et al. 2005). The genome-wide linkage analysis in LCLs from European CEPH families revealed the existence of genetic determinants that acted in *cis* or *trans* to the differentially expressed genes (Morley et al. 2004). Genotype data generated by the International HapMap Project (The International HapMap Consortium 2003), used to test for local and genome-wide association between expression phenotypes and SNPs in unrelated CEPH individuals, confirmed that both *cis*- and *trans*-acting loci regulate variation in the expression levels of genes (Cheung et al. 2005; Stranger et al. 2005); significant associations of genomic SNP regulators were more often detected at the locations *cis*-proximal than distal and *trans* to the variably expressed genes (Stranger et al. 2005).

These first studies, limited to transcriptome analysis in a single European population, were soon followed by others, designed to assess quantitative variation in gene expression across populations (Spielman et al. 2007; Stranger et al. 2007; Storey et al. 2007; Price et al. 2008; Zhang et al. 2008). Analysed LCLs were derived from ethnically homogeneous populations used in the International HapMap Project (CEU – Utah individuals of European ancestry, CHB – Han Chinese, JPT – Japanese, YRI – Nigerians) as well as from an admixed population of African Americans (AA). Expression phenotypes were analysed using Illumina or Affymetrix arrays, which addressed several thousand genes expressed in human LCLs. Spielman et al. 2007 compared expression in 142 individuals from CEU, CHB and JPT populations. The

mean expression level of almost 25% of the genes differed significantly between European and two combined Asian populations; at the same time, significant differences in the expression levels between CHB and JPT populations were observed in 1–3% of the genes. Although the majority of the expression differences were modest, taken collectively, they allowed to significantly differentiate the analysed populations. Storey et al. (2007), in a study performed in 16 CEU and YRI individuals, demonstrated that less than ~17% of the genes were differentially expressed among populations, while ~83% of the genes were differentially expressed among individuals. This observation, resembling the distribution of human variance observed at the DNA level, underlined the importance of estimating how gene-expression variation is apportioned within and among human populations. Zhang et al. 2008 evaluated gene expression in a much larger sample of 60 CEU and YRI parents–offspring trios. They estimated that 34–67% (depending on the thresholds used) genes were differentially expressed between the two populations (using more stringent cutoffs, 4.2−5.1% of the gene clusters were differentially expressed). Stranger et al. (2007) performed gene expression profiling in all of the 270 individuals genotyped in the HapMap; differentiation between populations was in agreement with earlier small-scale studies. To minimise the possible influence of non-genetic factors, which could contribute to differential expression across continental populations, Price et al. (2008) compared the transcription levels in 89 African-American individuals. The observed expression differences corresponded well to those predicted by the variable proportions of two continental ancestries in African-American individuals (assessed from the HapMap data), supporting the view that continental differences in transcriptome phenotypes were heritable and directly linked to genomic variation.

Genome-wide association analyses using HapMap SNP profiles, performed to elucidate genomic determinants of transcriptomic differences across populations, yielded results similar to those obtained for individuals from a single population: population differences in the expression levels were found to be associated with the differences in SNP allelic frequencies (Spielman et al. 2007; Storey et al. 2007; Zhang et al. 2008). The most comprehensive, fine-scale association analysis performed by Stranger et al. (2007) addressed over 2.2 million common SNPs per population in LCLs from all 270 individuals genotyped in the phase II HapMap (The International HapMap Consortium 2008). Their results (~1,350 differentially expressed genes with SNP association signals detected in *cis* and ~180 in *trans*), strongly supported the prevalence of genes with the expression regulated by *cis*-elements, but did not exclude a significant contribution of *trans*-effects that are often more indirect and, therefore, weaker and
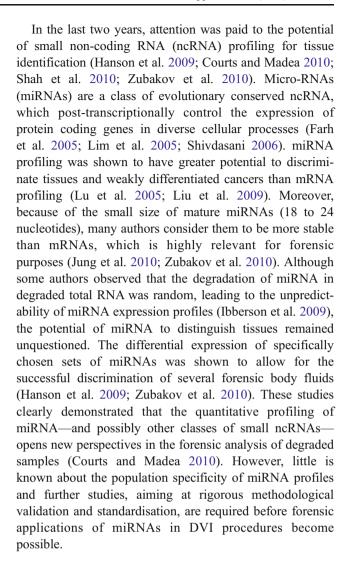
more difficult to detect. In contrast, Price et al. (2008), in their study of African-Americans, estimated that only 12–13% of the expression phenotypes were due to *cis* variants, indicating that distant or *trans*-acting elements can contribute substantially to the population differences in gene expression.

Although the exact contribution of *cis* or *trans*-acting regulatory genomic elements to heritable variation in expression remains to be established, the fact that gene expression differences among populations and individuals are associated with heritable DNA variation remains unquestioned, justifying the potential use of transcriptome analysis for determining the population specificity of the analysed samples. Furthermore, a number of biological processes were found to be enriched in the differentially expressed and differentially spliced genes between continental samples (Stranger et al. 2007; Zhang et al. 2008). By further analysing these processes, their possible roles in contributing to the phenotypic population differences can be revealed.

Determining whether associated SNPs are actual causative factors of expression phenotypes remains a challenge, and findings from genome-wide association studies need to be confirmed and verified by functional analyses. Apart from the possibility of just being linked to real regulators, SNP alleles can be indicative of CNVs, can themselves influence transcription or transcript stability or can be involved in alternative splicing.

CNVs are a potential source of expression-level differences among individuals and/or populations. It was shown that CNVs are sufficient to discriminate continental populations and that CNVs involving genes modify, to a certain extent, their expression levels (Armengol et al. 2009). However, transcriptome analysis in CEU and YRI populations has shown that the majority of differential expression phenotypes were not within genomic regions of known CNVs, indicating that the possibility of CNVs being major contributors to the expression differences is unlikely (Zhang et al. 2008).

Alternative splicing is likely to play an important role in regulating gene expression given the estimates that between one-third and two-thirds of all human genes may undergo alternative splicing (Sorek et al. 2004). Analysis of the HapMap Project data in European LCLs has recently shown that common genetic variation (SNPs) correlates with heritable changes in alternative splicing (Hull et al. 2007; Kwan et al. 2007). The genome-wide association study using 176 LCLs from individuals of European and African ancestry has shown that SNP variation does contribute to a substantial fraction of transcript isoform variation across these populations, although some non-genetic factors could also influence the observed differences (Zhang et al. 2008).

In the last two years, attention was paid to the potential of small non-coding RNA (ncRNA) profiling for tissue identification (Hanson et al. 2009; Courts and Madea 2010; Shah et al. 2010; Zubakov et al. 2010). Micro-RNAs (miRNAs) are a class of evolutionary conserved ncRNA, which post-transcriptionally control the expression of protein coding genes in diverse cellular processes (Farh et al. 2005; Lim et al. 2005; Shivdasani 2006). miRNA profiling was shown to have greater potential to discriminate tissues and weakly differentiated cancers than mRNA profiling (Lu et al. 2005; Liu et al. 2009). Moreover, because of the small size of mature miRNAs (18 to 24 nucleotides), many authors consider them to be more stable than mRNAs, which is highly relevant for forensic purposes (Jung et al. 2010; Zubakov et al. 2010). Although some authors observed that the degradation of miRNA in degraded total RNA was random, leading to the unpredictability of miRNA expression profiles (Ibberson et al. 2009), the potential of miRNA to distinguish tissues remained unquestioned. The differential expression of specifically chosen sets of miRNAs was shown to allow for the successful discrimination of several forensic body fluids (Hanson et al. 2009; Zubakov et al. 2010). These studies clearly demonstrated that the quantitative profiling of miRNA—and possibly other classes of small ncRNAs— opens new perspectives in the forensic analysis of degraded samples (Courts and Madea 2010). However, little is known about the population specificity of miRNA profiles and further studies, aiming at rigorous methodological validation and standardisation, are required before forensic applications of miRNAs in DVI procedures become possible.

Potential problems with using transcriptome analysis in forensics

Using transcriptome analysis for tissue identification is possible thanks to the fact that gene expression differs across tissues and that, within a particular tissue or cell type, it changes with time (e.g. depending on the RNA life time and developmental stage of the organism) and in response to external environmental conditions (diet, stress, infection etc.). These very characteristics make the application of expression profiling for the identification of individuals and/or populations very challenging. Such applications require an extensive effort to sieve away the effects of tissue-specific and environmentally triggered differences from the individual and population-specific variation in gene expression.

The majority of the presently available data on differential expression levels in individuals or populations are limited to a model of lymphoblastoid cell lines (EBV-immortalised B-cells). Expression phenotypes remain to be characterised

in a variety of tissues and environmental conditions. Furthermore, the accurate assessment of the expression differences among populations other than continental groups represented by HapMap LCLs requires further studies of the transcriptome in various human groups.

Ascertainment of genetic variants—SNPs, CNVs, splice variants—associated with or underlying transcriptome phenotypes that would stably and specifically differentiate individuals, populations and tissues would require even more rigorous standardisation, accounting for the complexity of the genomic regulation of expression phenotypes. This complexity was illustrated in the study by Dimas et al. (2009), who performed gene expression profiling and association with SNPs in three types of cells: LCLs, T-cells and primary fibroblasts. They demonstrated that the majority (80–90%) of genomic variants affecting gene regulation act in a cell type-specific manner, and even closely related cell types share only a minority of genomic *cis* regulatory elements; in addition, multiple regulatory SNPs existed per gene, unique or shared among cell types, and positively correlated with the number of transcripts per gene.

The application of transcriptome analysis for DVI poses additional challenges, related to the degradation of the biological material. While the degradation of the material cannot be avoided, an accurate assessment of the RNA template integrity remains a prerequisite for the reliable interpretation of quantitative gene expression analysis (Guhaniyogi and Brewer 2001). The commonly used ratio of ribosomal RNAs (28S to 18S) was found to provide only a weak correlation with RNA integrity (Auer et al. 2003; Miller et al. 2004; Schroeder et al. 2006; Copois et al. 2007). Recently, an effective standardised algorithm, called the RNA Integrity Number (RIN), was developed (Schroeder et al. 2006). The RIN approach uses a combination of microfluidics, microcapillary electrophoresis and fluorescence detection—the degradation of rRNA is reflected by a continuous shift towards shorter fragment sizes, providing values ranging from 10 (intact) to 1 (totally degraded).

The non-homogeneous character of most forensic samples presents another hindrance in the potential application of transcriptome analysis for DVI purposes. Analysis in mixed samples would require the deconvolution of signals specific for populations, individuals, different tissues and environmental factors.

Genetic profiling of specific cells separated
from heterogeneous mixtures

Forensic identification based on DNA or RNA profiles in mixtures of samples is a challenging task (Torres et al. 2003). Although procedures to be applied in such situations

exist and have been used in forensic practice, they require experience and complicated statistical evaluation (Clayton et al. 1998; Butler 2005). As a rule, mixed genotypes have a lower probative value. Interpretation is especially complicated in cases of an unknown number of contributors and their unbalanced proportion or different degree of degradation in the mixture (Torres et al. 2003). In such cases, physical separation of cells at the very beginning of investigation, allowing the isolation of pure cell populations from heterogeneous samples, would provide the best solution for forensic applications (Vandewoestyne and Deforce 2010).

There are various methods for the physical separation of cells. Differential lysis, filtration, sedimentation and adsorption methods exploit the differences in physical properties, like the shape and size of cells; fluorescence-activated or magnetic-activated cell sorting rely on the immunolabelling of cellular proteins (Vandewoestyne and Deforce 2010). All of these methods have already been applied in forensic investigations, mostly in rape cases, to distinguish cells of the male offender from those of the female victim (Witt et al. 1989, 1993). However, the effectiveness of these methods (efficiency of separation and yield of the specific cells) is limited, especially if the amount of initial biological evidence is very small. Laser capture microdissection (LCM) a technology combining microscopy with laser beam technology, offers a more suitable tool for the separation and collection of specific cell pools from mixed samples (Curran et al. 2000; Elliott et al. 2003; Li et al. 2009b; Vandewoestyne and Deforce 2010).

In LCM, specific cells are visualised via microscopy, separated from the others using laser beam energy and isolated into separate tubes, where the collected pools can be used for direct DNA or RNA extraction and analysis (Curran et al. 2000; Vandewoestyne and Deforce 2010). The discrimination of different types of cells is one of the biggest challenges in mixture separation by LCM. Fluorescence in situ hybridisation (FISH) probes, specific for X and Y chromosomes, are used to distinguish male from female cells. This application becomes widespread in forensic investigations of sexual assault cases with low numbers of sperm cells or male/female mixtures of blood or epithelial cells (Anslinger et al. 2005; Vandewoestyne and Deforce 2010). In medical sciences, LCM is used in cancer research to determine the cell-specific gene expression profile from heterogeneous human tissues (Curran et al. 2000; Cho-Vega et al. 2005). LCM can also be used for the isolation of blood cells from different cell mixtures or to obtain DNA profiles from biological samples contaminated with dirt or other components which may inhibit genetic analysis.

The problem in DVI is that, when LCM is performed in cells stained with histological stains, a stain providing good

discrimination of the cell type may negatively influence the DNA quality (Vandewoestyne and Deforce 2010). DNA profiling in mixed samples of the same sex would require "barcoding" specific cells, for example, with FISH probes specifically discriminating individuals (or populations). Another technical difficulty of LCM-based DNA profiling in mixtures is that the number of cells isolated by the method is too small for routinely applied tools to extract DNA. The most widely used, single-tube, DNA extraction method after LCM is the method based on proteinase K. Another recently developed technology (AmpliGrid slide, Advalytix AG, Germany) uses chemically structured microscope slides designed for ultra-low-volume applications. After the laser microdissection of cells (even of a single cell) into one of the hydrophilic reaction sites on the slide, DNA extraction and PCR can be performed directly on these reaction sites (Schmidt et al. 2006; Vandewoestyne and Deforce 2010). The combination of the AmpliGrid technology with micromanipulation has been already successfully applied by Li et al. (2009a) for a forensic analysis of different types of material evidence (Li et al. 2009b). Since the first application of LCM to separate spermatozoa from the victim's cells (Elliott et al. 2003), this technology has been used to separate cellular mixtures of blood and saliva, to collect single hair follicles or to distinguish foetal tissues from maternal tissues (Elliott et al. 2003; Budimlija et al. 2005; Vandewoestyne and Deforce 2010).

Conclusions

Mass-casualty incidents (MCIs) were present throughout the whole history of mankind. Rapidly developing biotechnologies offer an almost perfect tool for victims identification. The advantage of molecular genetic typing over any other method of personal identification is obvious. However, this highly complicated methodology is still undergoing development and should not be considered uncritically. The major, constantly debated technical topics are: how to deal with sample degradation, how to decrease background signals, how to increase the sensitivity and specificity of the typing. In fact, all of these factors arise from each other, circulating in a closed system of cause and effect.

Most countries allow classic DNA profiling, where markers are located only in non-coding genome regions. However, some of these non-coding markers are associated with visible traits. Furthermore, many of the markers provide information on an individual's biogeographic ancestry. Hence, the practical forensic application of scientist-developed methodology may violate the spirit of particular laws restricting the application of DNA analysis in certain countries. Furthermore, DNA profiling laws

usually only cover samples taken from convicts or suspects and from consenting volunteers, but these laws do not include all traces to which molecular typing could be applied. Many problems arise from a naïve adaptation of existing molecular methodologies by users who have no clue as to how deep the genetic insight into a personal heritage could be. Civil-law systems in many countries explicitly prohibit the determination of gender and race for forensic purposes as an act conducive to racial and/or social discrimination; this is especially important if the methodology is frequently applied for tracing suspects from minority groups. Thus, a combination of common sense, strict law enforcement and, above all, simple honesty is required for the use of genome-based technologies for particular purposes.

References

Anderson S, Howard B, Hobbs GR, Bishop CP (2005) A method for determining the age of a bloodstain. Forensic Sci Int 148:37–45

Anslinger K, Mack B, Bayer B, Rolf B, Eisenmenger W (2005) Digoxigenin labelling and laser capture microdissection of male cells. Int J Legal Med 119:374–377

Armengol L, Villatoro S, González JR, Pantano L, García-Aragonés M, Rabionet R, Cáceres M, Estivill X (2009) Identification of copy number variants defining genomic differences among major human groups. PLoS One 4(9):e7230

Auer H, Lyianarachchi S, Newsom D, Klisovic MI, Marcucci G, Kornacker K (2003) Chipping away at the chip bias: RNA degradation in microarray analysis. Nat Genet 35:292–293

Balding DJ (2005) Weight-of-evidence for forensic DNA profiles. John Wiley, New York

Bandelt HJ, Lahermo P, Richards M, Macaulay V (2001) Detecting errors in mtDNA data by phylogenetic analysis. Int J Legal Med 115:64–69

Bandelt HJ, Quintana-Murci L, Salas A, Macaulay V (2002) The fingerprint of phantom mutations in mitochondrial DNA data. Am J Hum Genet 71:1150–1160

Bandelt HJ, Salas A, Lutz-Bonengel S (2004) Artificial recombination in forensic mtDNA population databases. Int J Legal Med 118:267–273

Bauer M (2007) RNA in forensic science. Forensic Sci Int Genet 1:69–74

Bauer M, Polzin S, Patzelt D (2003) Quantification of RNA degradation by semi-quantitative duplex and competitive RT-PCR: a possible indicator of the age of bloodstains? Forensic Sci Int 138:94–103

Behar DM, Rosset S, Blue-Smith J, Balanovsky O, Tzur S, Comas D, Mitchell RJ, Quintana-Murci L, Tyler-Smith C, Wells RS; Genographic Consortium (2007) The Genographic Project public participation mitochondrial DNA database. PLoS Genet 3:e104

Bosch JR, Grody WW (2008) Keeping up with the next generation: massively parallel sequencing in clinical diagnostics. J Mol Diagn 10:484–492

Brandstätter A, Parsons TJ, Parson W (2003) Rapid screening of mtDNA coding region SNPs for the identification of west European Caucasian haplogroups. Int J Legal Med 117:291–298

Brandstätter A, Salas A, Niederstätter H, Gassner C, Carracedo A, Parson W (2006) Dissection of mitochondrial superhaplogroup H using coding region SNPs. Electrophoresis 27:2541–2550

Braslavsky I, Hebert B, Kartalov E, Quake SR (2003) Sequence information can be obtained from single DNA molecules. Proc Natl Acad Sci USA 100:3960–3964

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

Buckleton JS, Triggs CM, Walsh SJ (2005) Forensic DNA evidence interpretation. CRC Press, London

Budimlija ZM, Lechpammer M, Popiolek D, Fogt F, Prinz M, Bieber FR (2005) Forensic applications of laser capture microdissection: use in DNA-based parentage testing and platform validation. Croat Med J 46:549–555

Budowle B, van Daal A (2008) Forensically relevant SNP classes. Biotechniques 44:603–608, 610

Butler JM (2005) Forensic DNA typing, second edition: biology, technology, and genetics of STR markers. Elsevier Academic Press, New York

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

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

Carracedo A, Bär W, Lincoln P, Mayr W, Morling N, Olaisen B, Schneider P, Budowle B, Brinkmann B, Gill P, Holland M, Tully G, Wilson M (2000) DNA commission of the international society for forensic genetics: guidelines for mitochondrial DNA typing. Forensic Sci Int 110:79–85

Chakraborty R, Kidd KK (1991) The utility of DNA typing in forensic work. Science 254:1735–1739

Chakraborty R, Stivers DN, Su B, Zhong Y, Budowle B (1999) The utility of short tandem repeat loci beyond human identification: implications for development of new DNA typing systems. Electrophoresis 20:1682–1696

Cheung VG, Conlin LK, Weber TM, Arcaro M, Jen KY, Morley M, Spielman RS (2003) Natural variation in human gene expression assessed in lymphoblastoid cells. Nat Genet 33:422–425

Cheung VG, Spielman RS, Ewens KG, Weber TM, Morley M, Burdick JT (2005) Mapping determinants of human gene expression by regional and genome-wide association. Nature 437:1365–1369

Cho-Vega JH, Troncoso P, Do KA, Rago C, Wang X, Tsavachidis S, Medeiros LJ, Spurgers K, Logothetis C, McDonnell TJ (2005) Combined laser capture microdissection and serial analysis of gene expression from human tissue samples. Mod Pathol 18:577–584

Clayton TM, Whitaker JP, Sparkes R, Gill P (1998) Analysis and interpretation of mixed forensic stains using DNA STR profiling. Forensic Sci Int 91:55–70

Clayton TM, Whitaker JP, Maguire CN (1995) Identification of bodies from the scene of a mass disaster using DNA amplification of short tandem repeat (STR) loci. Forensic Sci Int 76:7–15

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

Coble MD, Just RS, O'Callaghan JE, Letmanyi IH, Peterson CT, Irwin JA, Parsons TJ (2004) Single nucleotide polymorphisms over the entire mtDNA genome that increase the power of forensic testing in Caucasians. Int J Legal Med 118:137–146

Coble MD, Vallone PM, Just RS, Diegoli TM, Smith BC, Parsons TJ (2006) Effective strategies for forensic analysis in the mitochondrial DNA coding region. Int J Legal Med 120:27–32

Collins PJ, Hennessy LK, Leibelt CS, Roby RK, Reeder DJ, Foxall PA (2004) Developmental validation of a single-tube amplification of the 13 CODIS STR loci, D2S1338, D19S433, and amelogenin: the AmpF\STR® Identifiler® PCR amplification kit. J Forensic Sci 49:1265–1277

Cooper DN, Smith BA, Cooke HJ, Niemann S, Schmidtke J (1985) An estimate of unique DNA sequence heterozygosity in the human genome. Hum Genet 69:201–205

Copois V, Bibeau F, Bascoul-Mollevi C, Salvetat N, Chalbos P, Bareil C, Candeil L, Fraslon C, Conseiller E, Granci V, Mazière P, Kramar A, Ychou M, Pau B, Martineau P, Molina F, Del Rio M (2007) Impact of RNA degradation on gene expression profiles: assessment of different methods to reliably determine RNA quality. J Biotechnol 127:549–559

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

Cummings TJ, Strum JC, Yoon LW, Szymanski MH, Hulette CM (2001) Recovery and expression of messenger RNA from postmortem human brain tissue. Mod Pathol 14:1157–1161

Curran S, McKay JA, McLeod HL, Murray GI (2000) Laser capture microscopy. Mol Pathol 53:64–68

Daca P, Mielnik M, Rogalla U, Skonieczna K, Linkowska K, Grzybowski T (2008) The application of minisequencing reactions for haplogroup assignment of mitochondrial DNA. Arch Med Sadowej Kryminol 58:212–217

Daca P, Mielnik-Sikorska M, Bednarek J, Grzybowski T (2010) Saturation of the Polish mitochondrial DNA database. Arch Med Sadowej Kryminol 60:263–269

Daly MJ, Rioux JD, Schaffner SF, Hudson TJ, Lander ES (2001) High-resolution haplotype structure in the human genome. Nat Genet 29:229–232

Decorte R (2010) Genetic identification in the 21st century—current status and future developments. Forensic Sci Int 201:160–164

Dimas AS, Deutsch S, Stranger BE, Montgomery SB, Borel C, Attar-Cohen H, Ingle C, Beazley C, Gutierrez Arcelus M, Sekowska M, Gagnebin M, Nisbett J, Deloukas P, Dermitzakis ET, Antonarakis SE (2009) Common regulatory variation impacts gene expression in a cell type-dependent manner. Science 325:1246–1250

Dixon LA, Murray CM, Archer EJ, Dobbins AE, Koumi P, Gill P (2005) Validation of a 21-locus autosomal SNP multiplex for forensic identification purposes. Forensic Sci Int 154:62–77

Dixon LA, Dobbins AE, Pulker HK, Butler JM, Vallone PM, Coble MD, Parson W, Berger B, Grubwieser P, Mogensen HS, Morling N, Nielsen K, Sanchez JJ, Petkovski E, Carracedo A, Sanchez-Diz P, Ramos-Luis E, Briōn M, Irwin JA, Just RS, Loreille O, Parsons TJ, Syndercombe-Court D, Schmitter H, Stradmann-Bellinghausen B, Bender K, Gill P (2006) Analysis of artificially degraded DNA using STRs and SNPs—results of a collaborative European (EDNAP) exercise. Forensic Sci Int 164:33–44

Dupuy BM, Stenersen M, Egeland T, Olaisen B (2004) Y-chromosomal microsatellite mutation rates: differences in mutation rate between and within loci. Hum Mutat 23:117–124

Elliott K, Hill DS, Lambert C, Burroughes TR, Gill P (2003) Use of laser microdissection greatly improves the recovery of DNA from sperm on microscope slides. Forensic Sci Int 137:28–36

Farh KK, Grimson A, Jan C, Lewis BP, Johnston WK, Lim LP, Burge CB, Bartel DP (2005) The widespread impact of mammalian microRNAs on mRNA repression and evolution. Science 310:1817–1821

Frudakis T, Venkateswarlu K, Thomas MJ, Gaskin Z, Ginjupalli S, Gunturi S, Ponnuswamy V, Natarajan S, Nachimuthu PK (2003) A classifier for the SNP-based inference of ancestry. J Forensic Sci 48:771–782

Fung WK, Hu YQ, Chung YK (2006) On statistical analysis of forensic DNA: theory, methods and computer programs. Forensic Sci Int 162:17–23

Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D (2002) The structure of haplotype blocks in the human genome. Science 296:2225–2229

Gill P (2001) An assessment of the utility of single nucleotide polymorphisms (SNPs) for forensic purposes. Int J Legal Med 114:204–210

Gill P (2002) Role of short tandem repeat DNA in forensic casework in the UK—past, present, and future perspectives. Biotechniques 32:366–368, 370, 372, passim

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

Golenberg EM, Bickel A, Weihs P (1996) Effect of highly fragmented DNA on PCR. Nucleic Acids Res 24:5026–5033

Graw M, Weisser HJ, Lutz S (2000) DNA typing of human remains found in damp environments. Forensic Sci Int. 113:91–95

Greenspoon SA, Scarpetta MA, Drayton ML, Turek SA (1998) QIAamp spin columns as a method of DNA isolation for forensic casework. J Forensic Sci 43:1024–1030

Grimes EA, Noake PJ, Dixon L, Urquhart A (2001) Sequence polymorphism in the human melanocortin 1 receptor gene as an indicator of the red hair phenotype. Forensic Sci Int 122:124–129

Guhaniyogi J, Brewer G (2001) Regulation of mRNA stability in mammalian cells. Gene 265:11–23

Haller AC, Kanakapalli D, Walter R, Alhasan S, Eliason JF, Everson RB (2006) Transcriptional profiling of degraded RNA in cryopreserved and fixed tissue samples obtained at autopsy. BMC Clin Pathol 6:9

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

Harismendy O, Ng PC, Strausberg RL, Wang X, Stockwell TB, Beeson KY, Schork NJ, Murray SS, Topol EJ, Levy S, Frazer KA (2009) Evaluation of next generation sequencing platforms for population targeted sequencing studies. Genome Biol 10:R32

Heinrich M, Matt K, Lutz-Bonengel S, Schmidt U (2007) Successful RNA extraction from various human postmortem tissues. Int J Legal Med 121:136–142

Hill CR, Duewer DL, Kline MC, Sprecher CJ, McLaren RS, Rabbach DR, Krenke BE, Ensenberger MG, Fulmer PM, Storts DR, Butler JM (2011) Concordance and population studies along with stutter and peak height ratio analysis for the PowerPlex ® ESX 17 and ESI 17 Systems. Forensic Sci Int Genet 5:269–275

Hoff-Olsen P, Mevåg B, Staalstrøm E, Hovde B, Egeland T, Olaisen B (1999) Extraction of DNA from decomposed human tissue. An evaluation of five extraction methods for short tandem repeat typing. Forensic Sci Int 105:171–183

Holland MM, Parson TJ (1999) Mitochondrial DNA sequence analysis—validation and use for forensic casework. Forensic Sci Rev 11:21

Huang QY, Xu FH, Shen H, Deng HY, Liu YJ, Liu YZ, Li JL, Recker RR, Deng HW (2002) Mutation patterns at dinucleotide microsatellite loci in humans. Am J Hum Genet 70:625–634

Hull J, Campino S, Rowlands K, Chan MS, Copley RR, Taylor MS, Rockett K, Elvidge G, Keating B, Knight J, Kwiatkowski D (2007) Identification of common genetic variation that modulates alternative splicing. PLoS Genet 3:e99

Ibberson D, Benes V, Muckenthaler MU, Castoldi M (2009) RNA degradation compromises the reliability of microRNA expression profiling. BMC Biotechnol 9:102–108

Isenberg AR, Moore JM (1999) Mitochondrial DNA analysis at the FBI Laboratory. Forensic Sci Commun 1:2

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

Jobling MA, Gill P (2004) Encoded evidence: DNA in forensic analysis. Nat Rev Genet 5:739–751

Jung M, Schaefer A, Steiner I, Kempkensteffen C, Stephan C, Erbersdobler A, Jung K (2010) Robust microRNA stability in degraded RNA preparations from human tissue and cell samples. Clin Chem 56:998–1006

Juusola J, Ballantyne J (2005) Multiplex mRNA profiling for the identification of body fluids. Forensic Sci Int 152:1–12

Kapińska E, Szczerkowska Z (2004) Personal identification based on nuclear DNA extracted from bones of deceased individuals. Probl Forensic Sci 60:104–116

Karlsson H, Guthenberg C, von Döbeln U, Kristenssson K (2003) Extraction of RNA from dried blood on filter papers after long-term storage. Clin Chem 49:979–981

Kayser M, Brauer S, Willuweit S, Schädlich H, Batzer MA, Zawacki J, Prinz M, Roewer L, Stoneking M (2002) Online Y-chromosomal short tandem repeat haplotype reference database (YHRD) for U.S. populations. J Forensic Sci 47:513–519

Kayser M, Liu F, Janssens AC, Rivadeneira F, Lao O, van Duijn K, Vermeulen M, Arp P, Jhamai MM, van Ijcken WF, den Dunnen JT, Heath S, Zelenika D, Despriet DD, Klaver CC, Vingerling JR, de Jong PT, Hofman A, Aulchenko YS, Uitterlinden AG, Oostra BA, van Duijn CM (2008) Three genome-wide association studies and a linkage analysis identify HERC2 as a human iris color gene. Am J Hum Genet 82:411–423

Kidd KK, Pakstis AJ, Speed WC, Grigorenko EL, Kajuna SL, Karoma NJ, Kungulilo S, Kim JJ, Lu RB, Odunsi A, Okonofua F, Parnas J, Schulz LO, Zhukova OV, Kidd JR (2006) Developing a SNP panel for forensic identification of individuals. Forensic Sci Int 164:20–32

Kishore R, Reef Hardy W, Anderson VJ, Sanchez NA, Buoncristiani MR (2006) Optimization of DNA extraction from low-yield and degraded samples using the BioRobot EZ1 and BioRobot M48. J Forensic Sci 51:1055–1061

Krenke BE, Tereba A, Anderson SJ, Buel E, Culhane S, Finis CJ, Tomsey CS, Zachetti JM, Masibay A, Rabbach DR, Amiott EA, Sprecher CJ (2002) Validation of a 16-locus fluorescent multiplex system. J Forensic Sci 47:773–785

Kwan T, Benovoy D, Dias C, Gurd S, Serre D, Zuzan H, Clark TA, Schweitzer A, Staples MK, Wang H, Blume JE, Hudson TJ, Sladek R, Majewski J (2007) Heritability of alternative splicing in the human genome. Genome Res 17:1210–1218

Li WH, Sadler LA (1991) Low nucleotide diversity in man. Genetics 129:513–523

Li L, Li CT, Li RY, Sun M, Liu Y, Li Y, Lin Y, Que T, Cheng D, Yan P, Fang J, Zhao Z, Shen M, Du Z (2004) ABO genotyping by duplex amplification and oligonucleotide arrays assay. Fa Yi Xue Za Zhi 20:193–196

Li L, Li CT, Li RY, Liu Y, Lin Y, Que TZ, Sun MQ, Li Y (2006) SNP genotyping by multiplex amplification and microarrays assay for forensic application. Forensic Sci Int 162:74–79

Li C, Qi B, Ji A, Xu X, Hu L (2009a) The combination of single cell micromanipulation with LV-PCR system and its application in forensic science. Forensic Sci Int Genet Suppl Ser 2: 516–517

Li R, Li Y, Fang X, Yang H, Wang J, Kristiansen K, Wang J (2009b) SNP detection for massively parallel whole-genome resequencing. Genome Res 19:1124–1132

Lim LP, Lau NC, Garrett-Engele P, Grimson A, Schelter JM, Castle J, Bartel DP, Linsley PS, Johnson JM (2005) Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature 433:769–773

Liu A, Tetzlaff MT, Vanbelle P, Elder D, Feldman M, Tobias JW, Sepulveda AR, Xu X (2009) MicroRNA expression profiling outperforms mRNA expression profiling in formalin-fixed paraffin-embedded tissues. Int J Clin Exp Pathol 2:519–527

Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, Sweet-Cordero A, Ebert BL, Mak RH, Ferrando AA, Downing JR, Jacks T, Horvitz HR, Golub TR (2005) MicroRNA expression profiles classify human cancers. Nature 435:834–838

Mandrekar PV, Flanagan L, Tereba A (2002) Forensic extraction and isolation of DNA from hair, tissue and bone. Profiles DNA 5:11–13

Matsubara Y, Ikeda H, Endo H, Narisawa K (1992) Dried blood spot on filter paper as a source of mRNA. Nucleic Acids Res 20:1998

Miller CL, Diglisic S, Leister F, Webster M, Yolken RH (2004) Evaluating RNA status for RT-PCR in extracts of postmortem human brain tissue. Biotechniques 36:628–633

Monks SA, Leonardson A, Zhu H, Cundiff P, Pietrusiak P, Edwards S, Phillips JW, Sachs A, Schadt EE (2004) Genetic inheritance of gene expression in human cell lines. Am J Hum Genet 75:1094–1105

Morley M, Molony CM, Weber TM, Devlin JL, Ewens KG, Spielman RS, Cheung VG (2004) Genetic analysis of genome-wide variation in human gene expression. Nature 430:743–747

Nachman MW, Crowell SL (2000) Estimate of the mutation rate per nucleotide in humans. Genetics 156:297–304

Nagy M, Otremba P, Krüger C, Bergner-Greiner S, Anders P, Henske B, Prinz M, Roewer L (2005) Optimization and validation of a fully automated silica-coated magnetic beads purification technology in forensics. Forensic Sci Int 152:13–22

Nassir R, Kosoy R, Tian C, White PA, Butler LM, Silva G, Kittles R, Alarcon-Riquelme ME, Gregersen PK, Belmont JW, De La Vega FM, Seldin MF (2009) An ancestry informative marker set for determining continental origin: validation and extension using human genome diversity panels. BMC Genet 10:39

Nilsson M, Andréasson-Jansson H, Ingman M, Allen M (2008) Evaluation of mitochondrial DNA coding region assays for increased discrimination in forensic analysis. Forensic Sci Int Genet 2:1–8

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

Pareek CS, Smoczynski R, Tretyn A (2011) Sequencing technologies and genome sequencing. J Appl Genet (in press). doi:10.1007/s13353-011-0057-x

Parson W, Bandelt HJ (2007) Extended guidelines for mtDNA typing of population data in forensic science. Forensic Sci Int Genet 1:13–19

Parson W, Dür A (2007) EMPOP—a forensic mtDNA database. Forensic Sci Int Genet 1:88–92

Parson W, Brandstätter A, Alonso A, Brandt N, Brinkmann B, Carracedo A, Corach D, Froment O, Furac I, Grzybowski T, Hedberg K, Keyser-Tracqui C, Kupiec T, Lutz-Bonengel S, Mevag B, Ploski R, Schmitter H, Schneider P, Syndercombe-Court D, Sørensen E, Thew H, Tully G, Scheithauer R (2004) The EDNAP mitochondrial DNA population database (EMPOP) collaborative exercises: organisation, results and perspectives. Forensic Sci Int 139:215–226

Pereira L, Cunha C, Amorim A (2004) Predicting sampling saturation of mtDNA haplotypes: an application to an enlarged Portuguese database. Int J Legal Med 118:132–136

Pfeiffer H, Forster P, Ortmann C, Brinkmann B (2001) The results of an mtDNA study of 1,200 inhabitants of a German village in comparison to other Caucasian databases and its relevance for forensic casework. Int J Legal Med 114:169–172

Phang TW, Shi CY, Chia JN, Ong CN (1994) Amplification of cDNA via RT-PCR using RNA extracted from postmortem tissues. J Forensic Sci 39:1275–1279

Phillips C, Prieto L, Fondevila M, Salas A, Gómez-Tato A, Alvarez-Dios J, Alonso A, Blanco-Verea A, Brión M, Montesino M, Carracedo A, Lareu MV (2009) Ancestry analysis in the 11-M Madrid bomb attack investigation. PLoS One 4:e6583

Price AL, Patterson N, Hancks DC, Myers S, Reich D, Cheung VG, Spielman RS (2008) Effects of cis and trans genetic ancestry on gene expression in African Americans. PLoS Genet 4(12):e1000294

Prinz M, Carracedo A, Mayr WR, Morling N, Parsons TJ, Sajantila A, Scheithauer R, Schmitter H, Schneider PM; International Society for Forensic Genetics (2007) DNA Commission of the International Society for Forensic Genetics (ISFG): recommendations regarding the role of forensic genetics for disaster victim identification (DVI). Forensic Sci Int Genet 1:3–12

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

Rees JL (2003) Genetics of hair and skin color. Annu Rev Genet 37:67–90

Reich DE, Schaffner SF, Daly MJ, McVean G, Mullikin JC, Higgins JM, Richter DJ, Lander ES, Altshuler D (2002) Human genome sequence variation and the influence of gene history, mutation and recombination. Nat Genet 32:135–142

Salas A, Bandelt HJ, Macaulay V, Richards MB (2007) Phylogeographic investigations: the role of trees in forensic genetics. Forensic Sci Int 168:1–13

Sanchez JJ, Phillips C, Børsting C, Balogh K, Bogus M, Fondevila M, Harrison CD, Musgrave-Brown E, Salas A, Syndercombe-Court D, Schneider PM, Carracedo A, Morling N (2006) A multiplex assay with 52 single nucleotide polymorphisms for human identification. Electrophoresis 27:1713–1724

Schmerer WM, Hummel S, Herrmann B (1999) Optimized DNA extraction to improve reproducibility of short tandem repeat genotyping with highly degraded DNA as target. Electrophoresis 20:1712–1716

Schmidt U, Lutz-Bonengel S, Weisser HJ, Sänger T, Pollak S, Schön U, Zacher T, Mann W (2006) Low-volume amplification on chemically structured chips using the PowerPlex16 DNA amplification kit. Int J Legal Med 120:42–48

Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, Lightfoot S, Menzel W, Granzow M, Ragg T (2006) The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC Mol Biol 7:3

Shah AA, Meese E, Blin N (2010) Profiling of regulatory microRNA transcriptomes in various biological processes: a review. J Appl Genet 51:501–507

Shivdasani RA (2006) MicroRNAs: regulators of gene expression and cell differentiation. Blood 108:3646–3653

Shriver MD, Kittles RA (2004) Genetic ancestry and the search for personalized genetic histories. Nat Rev Genet 5:611–618

Sobrino B, Brión M, Carracedo A (2005) SNPs in forensic genetics: a review on SNP typing methodologies. Forensic Sci Int 154:181–194

Sorek R, Shamir R, Ast G (2004) How prevalent is functional alternative splicing in the human genome? Trends Genet 20:68–71

Spielman RS, Bastone LA, Burdick JT, Morley M, Ewens WJ, Cheung VG (2007) Common genetic variants account for differences in gene expression among ethnic groups. Nat Genet 39:226–231

Storey JD, Madeoy J, Strout JL, Wurfel M, Ronald J, Akey JM (2007) Gene-expression variation within and among human populations. Am J Hum Genet 80:502–509

Stranger BE, Forrest MS, Clark AG, Minichiello MJ, Deutsch S, Lyle R, Hunt S, Kahl B, Antonarakis SE, Tavaré S, Deloukas P, Dermitzakis ET (2005) Genome-wide associations of gene expression variation in humans. PLoS Genet 1:e78

Stranger BE, Nica AC, Forrest MS, Dimas A, Bird CP, Beazley C, Ingle CE, Dunning M, Flicek P, Koller D, Montgomery S, Tavaré S, Deloukas P, Dermitzakis ET (2007) Population genomics of human gene expression. Nat Genet 10:1217–1224

Sturm RA (2009) Molecular genetics of human pigmentation diversity. Hum Mol Genet 18:R9–R17

Sweet D, Lorente M, Valenzuela A, Lorente JA, Alvarez JC (1996) Increasing DNA extraction yield from saliva stains with a modified Chelex method. Forensic Sci Int 83:167–177

Takamiya M, Saigusa K, Nakayashiki N, Aoki Y (2003) Studies on mRNA expression of basic fibroblast growth factor in wound healing for wound age determination. Int J Legal Med 117:46–50

The Council of the European Union (2001) Council Resolution of 25 June 2001 on the exchange of DNA analysis results

The International HapMap Consortium (2003) The International HapMap Project. Nature 426:789–796

The International HapMap Consortium (2008) A second generation human haplotype map of over 3.1 million SNPs. Nature 449:851–861

Thorisson GA, Smith AV, Krishnan L, Stein LD (2005) The International HapMap Project Web site. Genome Res 15:1592–1593

Tishkoff SA, Williams SM (2002) Genetic analysis of African populations: human evolution and complex disease. Nat Rev Genet 3:611–621

Tishkoff SA, Verrelli BC (2003) Patterns of human genetic diversity: Implications for human evolutionary history and disease. Annu Rev Genomics Hum Genet 4:293–340

Torres Y, Flores I, Prieto V, López-Soto M, Farfán MJ, Carracedo A, Sanz P (2003) DNA mixtures in forensic casework: a 4-year retrospective study. Forensic Sci Int 134:180–186

Tucker T, Marra M, Friedman JM (2009) Massively parallel sequencing: the next big thing in genetic medicine. Am J Hum Genet 85:142–154

Tully G, Bär W, Brinkmann B, Carracedo A, Gill P, Morling N, Parson W, Schneider P (2001) Considerations by the European DNA profiling (EDNAP) group on the working practices, nomenclature and interpretation of mitochondrial DNA profiles. Forensic Sci Int 124:83–91

Vallone PM, Jakupciak JP, Coble MD (2007) Forensic application of the Affymetrix human mitochondrial resequencing array. Forensic Sci Int Genet 1:196–198

Vandewoestyne M, Deforce D (2010) Laser capture microdissection in forensic research: a review. Int J Legal Med 124:513–521

Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, Ghandour G, Perkins N, Winchester E, Spencer J, Kruglyak L, Stein L, Hsie L, Topaloglou T, Hubbell E, Robinson E, Mittmann M, Morris MS, Shen N, Kilburn D, Rioux J, Nusbaum C, Rozen S, Hudson TJ, Lipshutz R, Chee M, Lander ES (1998) Large scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science 280:1077–1082

Wiegand P, Kleiber M (2001) Less is more—length reduction of STR amplicons using redesigned primers. Int J Legal Med 114:285–287

Willuweit S, Roewer L, International Forensic Y Chromosome User Group (2007) Y chromosome haplotype reference database (YHRD): update. Forensic Sci Int Genet 1:83–87

Witt M, Erickson RP (1989) A rapid method for detection of Y-chromosomal DNA from dried blood specimens by the polymerase chain reaction. Hum Genet 82:271–274

Witt M, Michalczak K, Latos-Bieleńska A, Jaruzelska J, Kuczora I, Lopez M (1993) An improved, non-isotopic method of screening cells from patients with abnormalities of sexual differentiation for Y chromosomal DNA content. J Med Genet 30:304–307

Yan H, Yuan W, Velculescu VE, Vogelstein B, Kinzler KW (2002) Allelic variation in human gene expression. Science 297:1143

Yao YG, Bravi CM, Bandelt HJ (2004) A call for mtDNA data quality control in forensic science. Forensic Sci Int 20(141):1–6

Zhang YH, McCabe ER (1992) RNA analysis from newborn screening dried blood specimens. Hum Genet 89:311–314

Zhang W, Duan S, Kistner EO, Bleibel WK, Huang RS, Clark TA, Chen TX, Schweitzer AC, Blume JE, Cox NJ, Dolan ME (2008) Evaluation of genetic variation contributing to differences in gene expression between populations. Am J Hum Genet 82:631–640

Ziętkiewicz E, Labuda D (2001) Modern human origins in light of the nuclear DNA diversity in world populations. In: Donnelly P, Foley RA (eds) Genes, fossils and behaviour: an integrated approach to human evolution. IOS Press, Amsterdam, The Netherlands, pp 79–97

Ziętkiewicz E, Yotova V, Gehl D, Wambach T, Arrieta I, Batzer M, Cole DE, Hechtman P, Kaplan F, Modiano D, Moisan JP, Michalski R, Labuda D (2003) Haplotypes in the dystrophin DNA segment point to a mosaic origin of modern human diversity. Am J Hum Genet 73:994–1015

Zimmermann B, Röck A, Huber G, Krämer T, Schneider PM, Parson W (2011) Application of a west Eurasian-specific filter for quasi-median network analysis: sharpening the blade for mtDNA error detection. Forensic Sci Int Genet 5:133–137

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

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

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

The worst Natural Disasters ever; http://www.scaruffi.com/politics/disaster.html

National Institute of Justice (2005) Mass fatality incidents: a guide for human forensic identification; http://www.ncjrs.gov/pdffiles1/nij/199758.pdf

National Institute of Justice (2006) Lessons learned from 9/11: DNA identification in mass fatality incidents; http://www.massfatality.dna.gov

http://www.ornl.gov/sci/techresources/Human_Genome/elsi/forensics.shtml

U.S. Department of Justice's Advancing Justice Through DNA Technology

http://www2.fbi.gov/hq/lab/codis/clickmap.htm

http://www.hvrbase.org/

http://www.ncbi.nlm.nih.gov/projects/SNP/

http://www.hapmap.org

http://www.mtdb.igp.uu.se/

http://www.nationalgeographic.com/genographic

http://empop.org

http://www.dna-conference.eu/ppt/Van%20der%20Beek.pdf

http://www.promega.com/resources/articles/profiles-in-dna/2009/expansion-of-the-european-standard-set/

http://www.iontorrent.com/

http://www.pacificbiosciences.com/smrt-biology/overview

http://www.promega.com/products/pm/genetic-identity/dna-iq/dna-iq/

http://marketing.appliedbiosystems.com/images/Product_Microsites/PrepFiler/PDF/prepfiler_autovalidation_bulletin.pdf