The handle http://hdl.handle.net/1887/20927 holds various files of this Leiden University dissertation.

Author: Westen, Antoinette-Andrea  
Title: Human identification & forensic analyses of degraded or low level DNA  
Issue Date: 2013-06-06
Chapter 8

General discussion
General discussion

Human identification is employed in diverse situations. When a clandestine grave is discovered in a forest, the circumstances are completely unlike those at an airplane crash site where bodies and body parts are scattered around. A tsunami area presents again an entirely different scenario as not only many casualties may occur, but also the complete infrastructure has flooded away. Nevertheless, in all such cases one of the main questions to answer is: “To whom belongs this body (part)?” Especially when bodies are beyond recognition, DNA-based methods can be very useful for identification. DNA, however, is subject to decay particularly when it has been exposed to harsh conditions, and this may hamper obtaining an interpretable DNA profiling result. The research described in Chapters 1 to 7 of this thesis (published as [1,2,3,4,5,6,7]) has been aimed at developing additional or alternative methods to extract information from a person’s DNA when standard DNA methodology is not sufficient for human identification. This chapter reviews various aspects of human identification and the technical and ethical choices at hand. Focus lies on (disaster) victim identification, although human identification also applies to relationship testing and forensic analysis.

Sample collection, DNA extraction and quantification

In case of a (mass) disaster, it is helpful when disaster victim identification (DVI) team members can rely on predefined guidelines, such as provided in “Mass fatality incidents: A guide for human forensic identification” by the National Institute of Justice [8] or the Interpol DVI guide [9]. These guides broadly outline the steps that need to be taken during DVI, but do not give specifics on how to perform certain procedures. In Chapter 1 (published as [1]), we have described a guideline for bone and tooth sample collection and contamination prevention for these samples to aid DNA-based victim identification. This guideline, which has been spread under the participants of “The use of DNA in disaster victim identification meeting” in Berlin in 2009, is based on common sense and practical experience. Additional experiences may provide further improvements. The two most important requirements for DNA-based identification are 1) collection of representative, high quality tissue samples from a victim and 2) the availability of reference samples, either from the alleged victim or from family members, with which the profiling results of the tissue samples can be compared [10,11,12,13]. Since the quality of the DNA is affected by time since death and environmental conditions (Box 1), samples need to be collected as soon as possible and stored under optimal conditions after sample collection to prevent further deterioration of the DNA [1,14,15]. To determine which samples are the most promising for DNA analysis, Milos et al. [16] analysed genotyping success of various skeletal elements (using skeletal samples from war victims in the former Yugoslavia). They verified empirically that the densest compact bones and teeth are the optimal samples for DNA recovery,
as is often assumed. Mundorff et al. [17] performed similar analyses (based on samples taken after the World Trade Center terrorist attacks) and suggest to sample patellae, metatarsals or foot phalanges since these provided similar genotyping results as femora and tibiae, but are much easier accessible; these skeletal elements can be excised with a disposable scalpel, thereby reducing contamination risk from reuse of instruments. Steinlechner et al. [18] advise to take (in addition to bone and tooth samples) two swabs from the internal organ or muscle surfaces that appear most intact (based on samples from the Sri Lanka site after the 2004 tsunami) and analyse these swabs first. The main advantage is that swabs are much easier to process than bone samples, but since DNA in soft tissues typically degrades faster than in bones and teeth, the chance of success is lower [18]. Evidently, any swab used for DNA-identification needs to be sterile and free of contaminating human DNA [19,20]. The actual sample collection procedure that is followed will depend on the condition of the remains, which to a large extent will rely on the type of disaster that has occurred as every disaster has its own specifics [17,21].

After (DVI) samples are taken, the DNA analyses are preferably performed in specialised and/or accredited (forensic) laboratories. When DNA is to be extracted from bone fragments or teeth, pulverisation of the samples is needed beforehand, which is usually done by cryogenic grinding (Box 2.1). According to Loreille et al., the DNA

### Box 1. DNA degradation

After death, cells and organelles disintegrate, thereby releasing lysosomal endonucleases that mediate DNA degradation. Further degradation can be caused by, for instance, exogenous nucleases from micro-organisms (that may grow on released nutrient-rich fluids after cell death) and environmental invertebrates. These enzymatic reactions can be followed by non-enzymatic or spontaneous DNA degradation processes that are generally much slower. These include hydrolytic reactions (at the glycosidic base-sugar bond, which can result in abasic sites, DNA cleavage or deamination for DNA bases with secondary amino groups), oxidative reactions (induced by $O_2$ or $H_2O_2$ and resulting in modifications such as formation of hydantoins that may block PCR), DNA cross-linking (within or between DNA strands or between DNA and proteins, which also blocks PCR) and radiation (that may induce a variety of lesions) [146,147]. The extent and manner of DNA degradation is influenced by environmental factors and exposure time to these factors. These environmental factors include (amongst others) river- or seawater; warm humid air; fire, nuclear radiation, UV-light, highly acidic soil and chemical agents that are used in deliberate attempts to degrade the DNA of war victims [1,148,149,150,151,101].
yield can be increased by total demineralisation of the bone powder after grinding [22] although some bone samples cannot be demineralised completely (personal communication, T. Parsons), which results in a smaller increase in yield. Next, the DNA has to be extracted from the sample and purified from possible PCR-inhibiting substances, such as humic acid (that can be present in soil) or soot (after fire). Many DNA extraction methods exist that mostly rely on phenol/chloroform extraction or silica (column or bead)-based purification (e.g. [23,24,25,26]). Comparative studies on human bones showed either that these methodologies perform equally well [27,28], or that the silica-based strategy (e.g. QIAgen kits, Venlo, the Netherlands) outperforms phenol/chloroform extraction [29,30] (L. Schoenmakers & A.A. Westen, unpublished results). An interesting finding by Guo et al. [28] is that the DNA extraction method has an influence on the ratio between the extracted nuclear DNA and mitochondrial DNA (mtDNA). They conclude that when a high yield of mtDNA is essential for your research, organic extraction is preferred over column-based extraction methods, while for nuclear DNA phenol/chloroform and QIAgen extraction methods performed equally well [28]. In general, silica-based methods are very good (and better than the organic extraction procedures) in removing non-DNA compounds that may negatively affect the PCR [29,31,32]. In specific cases, a pre-PCR DNA clean-up kit can be used such as NucleoSpin gDNA Clean-up (Macherey-Nagel) [33,34] or Powerclean DNA Clean-Up [35], to obtain DNA as pure as possible.

A completely different approach to obtain DNA from (bone) cells is based on pressure cycling technology (PCT). PCT makes use of hydrostatic pressure that alternates between ultrahigh and ambient levels to physically disrupt the cellular material. This method is reported to extract proteins, lipids and nucleic acids in parallel from the same sample [36]. The DNA recovery is dependent on the cell type and the efficiency (compared to DNeasy (Qiagen) methodology) ranges between 100 % and 30 % [36]. An additional interesting aspect of PCT is the ability to change the conformation of compounds under high pressure, which may positively affect the PCR when inhibiting substances are present [37].

After isolation, the DNA is quantified, preferably by a human-specific methodology, to determine the PCR input in order to obtain balanced results in the generally multi-target DNA profiling assays. Different quantification methods amplify DNA targets of different fragment lengths. Since the DNA in DVI samples is often degraded, it is important to realise that quantification methods relying on short DNA amplicons (such as Quantifier™ Human (Applied Biosystems (AB)), having a 62 bp target [38]) may provide higher quantifications than methods with longer targets (such as Plexor HY™ (Promega) with a 99 bp [39], Alu-assay with a 127 bp [40], Quantifier™ Duo (AB) with a 140 bp [38] or Quantiplex HYres™ (Qiagen) with a 146 bp target). Given that most short tandem repeat (STR) multiplex systems currently comprise amplicons between 70 and 400 bp, the quantification methods with the longer DNA targets will give more accurate predictions of genotyping success than those with shorter targets.
STR profiling of compromised samples

Usually, the next step towards human identification is an STR PCR to amplify and fluorescently label DNA markers before detection by capillary electrophoresis (CE). The PCR may be inhibited by several substances that were co-isolated with the DNA, resulting in no profiles or profiles with lower peak heights than expected from the amount of template DNA. PCR inhibitors can be classified into three groups by their source: 1) intrinsic inhibitors, such as heme in blood or melanin in hair; 2) inhibitors from the substrate, such as humic acid in soil or indigo dye in denim and 3) other inhibitors, such as substances used during the extraction process (e.g. SDS or chelex resin) or EDTA, which is used to buffer the DNA extracts [41]. Using less of the DNA extract may improve the results, but may also lead to partial DNA profiles when the quantity of the DNA is low, as is often the case in human identification samples. The PCR process may benefit from modified DNA polymerases that are more tolerant to PCR inhibition. Some of these modified enzymes are even able to directly amplify whole blood such as Hemo KlenTaq (a TaqI polymerase that lacks the first 280 amino acids and has several additional mutations) [42] and Phusion™ Flash Polymerase (a Pyrococcus-like polymerase to which a DNA-binding domain is fused) [43]. As Phusion™ Flash Polymerase is in addition accurate and fast, this enzyme was incorporated in an in-house developed “DNA-6-hours” service that assists police investigations by rapidly deriving DNA information from trace evidence [44,45]. Besides, different DNA polymerases vary for the capacity to deal with difficult templates such as GC-rich or looped sequences.

STRs are complex templates in the sense that they comprise repeated sequences, which invokes slipped strand displacement during amplification due to which back and forward stutter products are formed [5]. As stutter peaks resemble real alleles, they can hamper the interpretation of unequal mixtures. DeepVent™ DNA polymerase [46] is a polymerase adapted for use with difficult templates, but it did not produce less stutter products than the standard AmpliTaq™ Gold polymerase (Box 2.2). Other options that have been explored to improve PCR performance relate to adjustment of the PCR parameters, such as adding twice the recommended amount of Taq polymerase [47,48], elongation of the annealing time to reduce allele drop-out [48], reducing PCR volume, or (and this strategy is the most widely used) increasing the number of PCR cycles for the analysis of low template DNA [49,47,50,51]. Some of these aspects have been accounted for during the development of the current generation STR kits (e.g. AmpFISTR® NGM™ (AB), the PowerPlex® ESX™ and ESI™ systems (Promega) and Investigator® ESSplex™ (Qiagen)); these kits show an improved resistance to PCR inhibitors due to optimised buffer systems, have elongated annealing times to reduce allele drop-out rates and make use of one or two additional PCR cycles (i.e. 29 or 30 cycles) to increase sensitivity [52]. The newest kits aim specifically for efficient analysis of unprocessed reference samples. These kits accommodate direct amplification of both
buccal and blood specimens and achieve amplification in impressively short times (e.g., below 50 minutes for the Investigator® IDplex Direct™ kit (beta testing for Qiagen)). Since commercial companies developed these kits, the compounds effectuating these improvements are unknown.

Low template DNA amplification (as is frequently encountered in human identification and forensic research) is typically affected by stochastic effects such as allele drop-out, near-threshold peaks, heterozygous peak imbalances, increased stutters and allele drop-in. To aid analyses, stochastic thresholds can be derived, and when allele peak heights are below this threshold it is anticipated that stochastic effects may have occurred. In Chapter 5 (published as [5]) we determined stochastic thresholds for the analysis of NGM™ DNA profiles. There are several options to improve low template DNA profiling results, such as the above-described methods to increase PCR performance (of which performing additional PCR cycles [50,51,47] is most widely and frequently used). In addition, adaptations at the post-PCR level can be made such as purification of amplified products (as described in Chapter 7 (published as [7])) [53,54,55,56,57]. Performing additional PCR cycles (e.g., 34 cycles as used at the NFI [51,5]) is only recommended for samples with very low DNA inputs (< 31 pg, using NGM™ [5]), as with higher inputs the profiles will become overloaded [51]. For low template samples with higher inputs, enhanced injection settings are advised preceded by post-PCR purification to reduce dye-blobs in the DNA profile (e.g., 9 kV for 15 seconds combined with DTR gel filtration, as suggested in Chapter 2 (published as [2])). This method retrieves almost as many alleles as 34 cycle PCR, but shows less drop-ins, and can be performed on the remaining PCR product after standard analysis (thus no additional PCR or use of extra DNA extract is needed) [51,5]; this method is embedded in the ISO17025 accreditation of the NFI and in use since November 2007.

When some of the peaks in the standard DNA profile are relatively high, for instance on the shorter loci in profiles of degraded samples or for the major component in an unequal mixture, specific post-PCR purification strategies (such as a size-selective method based on AMPure® XP beads that especially enriches for longer amplicons) may be beneficial, as described in Chapter 7 (published as [7]). The above-described methods are applied when standard STR profiling resulted in peaks below the stochastic threshold and/or allele drop-outs. Consequently, these methods are to be regarded as low template techniques, and low template interpretation strategies are appropriate. The most applied interpretation strategy involves replicate PCR analyses from which a consensus profile is derived [51]. Benschop et al. have compared several methods to generate a consensus profile, and recommend the n/2 method in which alleles are included in the consensus when they are designated in at least half of the replicates (with n = 3 or n = 4 as optimal replicate number) [51]. An upcoming alternative is the use of statistical interpretation tools that harbour the occurrence.
of drop-outs and drop-ins [58, 59, 60]. When peaks in the DNA profiles are low and sufficient DNA extract remains, a final possibility is to concentrate the DNA extract to a volume befitting the PCR set up (for instance by ethanol precipitation). Since this will consume most (if not all) DNA extract, this method should only be performed when it is anticipated that the resulting allele peak heights will be above the stochastic threshold. Otherwise, it might be sensible to store the DNA extract in expectation of future development of more sensitive methods.

Whole genome amplification (mostly used for single cells in pre-implantation genetic diagnostic research) has been proposed as an amplification method for forensic samples with extremely low amounts of template DNA [61, 62]. Several techniques exist and have been compared [62], and increased amplification success was obtained especially with multiple displacement amplification [63, 62]. However, WGA suffers from preferential amplification and this difficulty (that translates into locus and allele drop-out) has not yet been overcome (Box 2.3) [64]. Another concept that has been tested for relatively low template DNA samples, is employing the unused genomic DNA that still resides in PCR mixtures (of which only a small amount has been used for CE) after amplification as template for a second PCR based on a distinct marker system. Although products for the second marker system were generated, the primers from the first amplification strongly interfered in the second amplification and this approach was regarded unsuccessful (Box 2.4). Compromised samples may suffer not only from reduced quantity, but also from low quality, which means that the DNA is degraded. Therefore, attempts have been made to repair the degraded DNA before amplification by using DNA repair enzymes that \textit{in vivo} correct genomic DNA damage. Several groups applied the commercially available PreCR™ Repair Mix (New England Biolabs, MA, USA), with variable success rates [4, 65, 66], which mainly relate to the amount of template DNA used. Full STR profiles are obtained when an input of 25 ng DNA (from an old bloodstain) is repaired [66], but thus far no success is reported for (mock) casework samples (Box 2.5) [4, 65].

### Box 2. Supplemental results generated during this thesis work

#### 2.1 Optimisation of cryogenic grinding

In order to optimise the DNA yield from bone samples, these samples need to be grinded before DNA isolation. In a comparison of three cryogenic grinding systems, the Freezer/Mill® 6770 (SPEX SamplePrep, Metuchen, NJ, USA) produced bone powder with a finer structure (and, on average, more detected alleles per profile) than the Freezer/Mill® 6750 (SPEX SamplePrep), while the TissueLyser (Qiagen; with a liquid nitrogen cooled stainless steel grinding jar...
General discussion

set) was not capable of grinding the archaeological bone samples [152] at all (L. Schoenmakers & A.A. Westen, unpublished results). Thus, we regard the Freezer/ Mill® 6770 as the optimal grinding system, and optimised the grinding process on this machine.

2.2 Reduction of stutter peak height

During amplification of STRs, slipped strand displacement may occur resulting in stutter products that are visible in the DNA profiles as peaks at one repeat length before (or after) the actual allele. These stutter peaks resemble real alleles and may hamper the analysis of unequal mixtures. DeepVent® DNA polymerase is a polymerase adapted for use with difficult templates, and we tested its effect on the formation of stutter products. The results were compared to AmpliTaq™ Gold (which is the standard DNA polymerase) in SGM Plus™ DNA profiles. Both DNA polymerases produced stutter peaks of similar heights, and the use of DeepVent® caused a decrease in the inter-locus balance. As no positive effect was seen from the use of DeepVent®, it was decided not to proceed to tests involving unequal mixtures.

2.3 Multiple displacement amplification (MDA)

Whole genome amplification (WGA) can be used to generate manifolds of DNA from relatively low amounts of starting template. This could facilitate multiple analyses for samples otherwise allowing only one or few analyses. Many WGA methods exist which can be generally divided in two categories: those based on variations of PCR amplification (including techniques to attach random primers, degenerate primers and adaptors to genomic DNA fragments to create universal priming regions) and those with isothermal DNA amplification utilising multiple strand displacement amplification. Since the quality of the DNA in human identification and forensic analyses is usually low, fragmenting the DNA to create universal priming regions may lead to DNA that is too much fragmented to allow subsequent STR analysis. Multiple displacement amplification (MDA) methods generally require high quality DNA to be successful. An exception is the REPLI-g FFPE kit (Qiagen) that ligates all the short fragments before MDA takes place. Although the manufacturer recommends a minimum DNA input of several nanograms, we tested this kit with a DNA input of 500 pg to examine its performance for forensic use. Using ALU quantification [40], an increase in DNA concentration was measured. However, DNA profiles generated from this MDA-amplified DNA showed strong preferential amplification, which led to allele drop-out on the one hand and over-amplification of alleles on the other
hand within the same profile. Due to these extremely imbalanced DNA profiles we feel that this method is not fit for forensic practice, but we look forward to future developments that may comply with forensic use.

2.4 Reusing a DNA sample

At the NFI, the standard PCR volume is 25 µL of which 1 µL is analysed by CE. The remaining 24 µL are stored (for possible re-analysis of the PCR product) and discarded after a few months (as the instability of the fluorescent groups increases with time). This stored PCR product still contains the original genomic DNA and we tested whether this DNA could serve as template used for a second amplification with a distinct STR kit. When performing a second amplification, it is important to remove all primers used during the first PCR from the mixture (as these will induce amplification again, otherwise), together with the fluorescent labels from the PCR products (or the fluorescently labelled PCR products themselves). MiniFiler™ was chosen as first and Yfiler™ as second genotyping kit as these present very distinct profiling results. For primer removal ExoSAP-IT™ was tested that contains Exonuclease I, which can degrade single-stranded primers and DNA. Indeed it removed the primers from the PCR mixture, but at the same time also original DNA templates were degraded (possibly because these became single-stranded from the multiple denaturation steps). After purifying the PCR product using in-house prepared Sephadex G-50 or G-100 columns (with an exclusion limit of 20 or 25 bp, respectively) no loss of DNA was detected and (MiniFiler™) PCR primers seemed removed, as after the second (Yfiler™) PCR the Yfiler™ alleles were visible (which did not occur without applying the Sephadex columns). Nevertheless, the MiniFiler™ loci became further amplified, indicating that the MiniFiler™ primers were not removed completely. In addition, when using the G-50 and G-100 columns drop-ins occurred more frequently than in the control reactions (Yfiler™ reactions from untreated genomic DNA). As we did not find a good manner to fully dispose of the primers of the first PCR (in 2010), we decided not to continue this research line. However, with the latest experience with the size-selective AMPure® beads purification, as described in Chapter 7 (published as [7]), this line of research may be revisited.

2.5 Repair of DNA damage

DNA repair enzyme cocktails, such as Restorase™ and PreCR™, are developed to repair several kinds of DNA damage (e.g. abasic sites, nicks and thymine dimers) that may block processivity during PCR. Repairing damaged
DNA (as encountered in human identification research) may be advantageous and increase genotyping success. To simulate single strand breaks, a model system was developed using cassettes of synthetic oligos (A.A. Westen, B. Erkamp, C. Vervat, D. Mourik & T. Sijen, unpublished results). Each cassette was designed to represent double-stranded DNA of 110 bp with a single nucleotide gap in each strand. The region between the two gaps varied in length and was either 4, 8, 12, 16 or 20 nt. To achieve such a cassette, four oligos were annealed; two for the sense strand and two for the antisense strand. Both Restorase™ and PreCR™ showed promising results and repaired the templates having a region of 8, 12, 16 or 20 nt between the gaps; both kits could not repair the templates with a region between the gaps of 4 nt, which seems to act as a double strand break (for which is known that these kits cannot repair it). When these kits were used on UV-degraded human DNA and analysed using an SGM Plus™ PCR, variable results were obtained. Sometimes the repaired DNA showed additional alleles compared to unrepaired samples, but other times less alleles could be retrieved. These results are not robust enough to be confident that forensic samples will benefit from use of these repair cocktails. As the amount of sample is often limited, it seems not worthwhile to take the risk that genotyping may turn out less efficient. The results of this study were compared to other methods for the analysis of degraded DNA (standard SGM Plus™ profiling, MiniFiler™ and SNPs) and are described in Chapter 4 (published as [4]).

2.6 Follow-up on the tri-allelic SNP research

In August 2010, the Beijing Genomics Institute (BGI) kindly provided information on the potential tri- and tetra-allelic SNPs from the low-coverage (2-4X) study on 60 CEU, 60 CHB-JPT and 60 YRI samples from the 1000-genomes project. Almost 19,000 potential tri- and tetra-allelic SNPs were assessed, but within the provided dataset not one SNP appeared to be tri-allelic in all three population groups, although a few SNPs showed three alleles for two population groups. The discriminating power was assessed for a combined SNP set comprising the SNPs that were tri-allelic in two population groups and the 10 SNPs that were shown to be tri-allelic in our earlier study (as described in Chapter 3 (published as [3]), and this discriminating power did not reach a value powerful enough for human identification. Thus, we have not further tested the newly discovered tri-allelic SNPs. Since the low-coverage study has now expanded from 180 to 1,092 individuals [153] and will be extended to around 2500 samples in the near future, analysis of the new data may reveal additional SNPs that are tri-allelic in the major population groups worldwide.
2.7 Post-mortem interval determination based on molecular markers

In order to be able to determine whether a body could belong to a missing person, determination of the late post-mortem interval (PMI) could be of great value. Zimmermann et al. [132] analysed the nucleoside content of 5 to 40 year-old museum moth specimens by mass spectrometry. Promising results were obtained showing a time-dependent loss of deoxyguanosine (dG); the dA/dT ratio remained fairly constant, while the dG/dC ratio decreased gradually with time [132]. We performed a pilot study with a similar setup using human samples [154]. Femur and soft tissue samples were obtained from persons that were buried 2 to 40 years earlier. After sampling, the samples were cooled to 4 °C and upon arrival at the NFI directly stored in a -80 °C freezer. DNA was extracted, amplified and analysed according to standard NFI procedures. No signification correlation between DNA profiling success, lengths of amplified fragments and PMI was detected in these samples. After digestion of the DNA, separation of the nucleosides by high-pressure liquid chromatography (HPLC) and detection by tandem mass spectrometry (MS/MS), no decrease in the dG/dC ratio could be demonstrated. It is unknown what caused the differences in results of both studies. To our knowledge, no independent conformation of the findings by Zimmermann et al. [132] have been reported, which may indicate that either the detection method is technically challenging, or the findings do not apply generally. In addition, our sample set may have an unfavourable amount of heterogeneity because the samples were exposed to more variable conditions than the moth samples that were stored in a museum. Thus, we infer that this approach is not usable for late PMI determination in human remains.

Choice of marker type

Autosomal STRs have been the workhorses in human identification and forensic DNA analysis over the last two decades, and (national) DNA databases for missing persons and forensic casework mainly consist of autosomal STR profiles. These STR profiles can be used both for one-to-one comparisons with reference data and for pedigree analysis for which specialised software tools, such as Bonaparte [67,68], have been developed. When DNA is degraded to fragment lengths shorter than the STR amplicons, the resulting DNA profiles will have lower peak heights or allele drop out at the longer loci. To enable analysis of these loci in case of degraded DNA, mini-STRs were developed that have primers closer to the actual STR (to obtain amplicon sizes preferably below 200 bp). With the introduction of five new European standard set
General discussion

(ESS) markers and the entry of a third commercial company that produces forensic STR kits (Qiagen, since 2010), various new STR kits became available, engineered to have most STRs in the mini- and mid-size range [69]. The diverse kits contain different markers as a mini-STR, and using another kit may retrieve information that was missing upon amplification with the first kit. At the NFI, the AmpFISTR® NGM™ Kit (AB) [52], was selected for standard use and part of the in-house NGM™ validation is described in Chapter 5 (published as [5]), with focus on the determination of stochastic thresholds, back and forward stutter filters and low template DNA strategies. For specific cases, such as pairwise relationship testing in deficient pedigrees (e.g. grandparent-grandchild and avuncular relationships, or great grandparent-great grandchild and cousins), it can be helpful to type additional autosomal STR markers [70]. Nine of the 13 markers in the Investigator HDplex™ Kit (Qiagen) are uncommon to forensic analyses and traditionally used for bone marrow transplantation monitoring [6,71,72]; we determined Dutch allele frequencies for these loci as described in Chapter 6 (published as [6]). In addition, we calculated that, on average, the power of discrimination per marker is higher for HDplex™ than for NGM™ or Identifiler™, and no linkage disequilibrium was detected between the three kits (Chapter 6 (published as [6])). These additional markers, however, are not present in the current DNA databases and can therefore only be used in direct comparisons.

In case not enough STR information can be recovered from the DNA due to the high degree of degradation, it might still be possible to obtain genotyping information using single nucleotide polymorphisms (SNPs, Box 3) [73,74] and/or deletion-insertion polymorphisms (DIPs or InDels) [75,76,77]. Since only one or few nucleotides are analysed, SNP and DIP amplicons can be smaller than mini-STRs, although many more SNP or DIP markers need to be analysed due to the lower discrimination power per marker (the estimated number of markers required to reach a probability of identity comparable to 12-20 STRs lies between 45 and 65 [78,79,80,81,76]). Other advantages are that SNPs and DIPs are widely spread throughout the genome, possess low mutation rates, do not produce stutter artefacts, have high multiplexing capabilities and are amenable to automation. At the moment, no commercial kit is available to type the sequence information for SNPs (a commercial chip may be available soon [82], but this tool has not yet been tested by independent laboratories). Many different SNP typing techniques exist (Box 3), and although single base extension (SBE) is the most widely used method in forensic settings, there is no consensus on the method or the SNP set to be used between forensic laboratories. Recently, a commercial kit was introduced for the analysis of DIPs (i.e. Investigator® DIPplex™ Kit, Qiagen). Since DIPs are characterised by length-differences (and not sequence differences), they can be analysed in a similar manner as STRs, with a multiplex PCR followed by CE detection, which makes detection and (mixture) analysis easier than for SNPs [76,83]. When used for human identification, both SNPs and DIPs should ideally have similar
Box 3. Single nucleotide polymorphisms (SNPs)

A set of around 50 SNPs (based on the SNPforID assay by Sanchez et al. [73]) was broadly tested by the forensic community (e.g. [155,156,157,158,159,160]) and resulted in the development of an assay named GenPlex™ [161]. This assay was subjected to an inter-laboratory study [162,163], and although this study created support for GenPlex™ in many forensic laboratories, Applied Biosystems unfortunately decided to withdraw the chemistry on which it was based from the market. Several other assays have been developed, most of which are based on a multiplex PCR followed by different methods to detect the SNPs, such as single base extension (SBE) [73], Amplification Refractory Mutation System (ARMS) combined with Universal reporter primers (URP) [164], Matrix Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF MS) [165], Invader assays [166], and micro-arrays [167]. Some of these techniques need high (above 25 ng) amounts of template DNA, while other techniques, such as SBE, also work well with DNA inputs of 1 ng or less, which makes them more suitable for human identification in DVI or forensic settings. One such SBE multiplex that is shown to be sensitive (with full profiles down to 78 pg/µL) is based on SNPs that are located inside nucleosomic regions, which were chosen as there is growing evidence that the histone-DNA complexes found in nucleosomes might offer protection from DNA degradation processes [168]. When regarding the average percentage of markers successfully detected, the nucleosome SNPs performed slightly better than the two SNPforID SBE multiplexes, and all these three SNP-based assays performed better than MiniFiler™ when degraded casework samples were tested [168]. Another study that showed a higher percentage of detected alleles for SNP-based assays compared to MiniFiler™, is described in Chapter 4 (published as [4]). The two SNP assays that outperformed MiniFiler™ were the GenPlex™ assay (based on bi-allelic SNPs) and an SBE assay based on the tri-allelic SNPs as described in Chapter 3 (published as [3]). Tri-allelic SNPs can assist in the recognition of mixed samples due to the third allele that can be present, while for bi-allelic SNPs only a change in peak height ratios can give an indication of a mixture (and as a result of the unequal peak heights for the different nucleotide-dyes in SBE analyses, ratio changes are more difficult to recognise than for STRs) [3]. Although it would be interesting to increase the number of tri-allelic SNPs available for human identification and forensic purposes [3,169,170], their discovery has proven difficult (box 2.6) as they often have a low frequency and are then regarded to be sequencing errors.
allele frequencies for different population groups. If (huge) variation between groups is present, they can be interesting as ancestry informative markers [3,84,85,86,87,88]. Due to their high mutational stability (when compared to STRs), SNPs and DIPs can also be informative (additional) markers in kinship analysis and (when several thousands SNPs are used) in distant pairwise relationship tests [70,83]. The use of thousands of SNPs (each with a low minor allele frequency) has also been proposed as a method to effectively identify the number of contributors in complex mixtures with multiple (up to 10) contributors [89].

Other proposed markers combine multiple polymorphisms. Haplotype blocks, for instance, contain multiple SNPs that are tightly linked, and when several haplotypes are observed at the population level, a haploblock has a higher discrimination power than the individual SNPs within the block [90]. SNP-STRs, which refers to sequence differences (SNPs) within STRs, can be determined by, for instance, (electrospray ionisation) mass spectrometry (ESI-)MS [91,92] or (next generation) sequencing (by for instance Illumina GAIIx short-read technology [93]). With ESI-MS, additional polymorphisms were identified in 11 of the 13 CODIS core STRs, thereby increasing the discrimination power significantly [92]. Next generation sequencing faced some difficulties in analysing the longer alleles as most reads had a maximum read length of 150 bp [93], but next generation sequencing methods hold great promise for mixture analysis. A big advantage of both techniques is that the obtained data are backwards compatible with existing STR information that is present in the national DNA databases, as the repeat lengths of the STRs can be inferred (next to the additional SNP sequence information). A disadvantage, however, is that they require specialised instruments that are not common in forensic laboratories. Another combined marker is the DIP-STR, which consists of a DIP near an STR (within 500 bp distance) [94]. This marker type has been developed for the analysis of minor components in DNA mixtures by using allele-specific amplifications of DIP-STR haplotypes [94]. An advantage of these markers is that they are sex-independent (unlike Y-STRs which are frequently used to examine male minor components when the major is female); a disadvantage is that relatively large sizes for the amplicons are needed, which makes them less suitable for degraded samples.

Next to autosomal DNA testing, analysis of Y-chromosomal, X-chromosomal and mitochondrial DNA (mtDNA) is of interest for human identity testing. Y-chromosomal markers are inherited from father to son, and therefore provide patrilineal information. Deep ancestral relationships can be analysed with Y-SNPs, as they have very low mutation rates. The Y-STRs currently used in forensic genetics, provide information on (more recent) relationships in the paternal lineage, while the recently described rapidly mutating Y-STRs seem to provide the possibility to distinguish between closely related males [95,96]. Both these types of Y-STRs can provide valuable information in paternity testing and in forensic cases involving a male component in mixtures with
an overabundance of female DNA. In contrast to the Y-chromosome that is inherited by sons, daughters inherit the X-chromosome of their fathers, which make X-STRs interesting markers for the analysis of, for instance, disputed father-daughter relationships [97,98]. Children always inherit their mitochondrial DNA from their mothers, and therefore mtDNA provides matrilineal information. Since several hundreds of copies of the mtDNA are present per cell (compared to two copies for the autosomes), mtDNA analysis is highly sensitive. Telogen hairs lack the hair roots that comprise nuclei and nuclear DNA, but still contain mtDNA up to at least 12 cm from the scalp (L. Clarisse & T. Sijen, unpublished results). Therefore, mtDNA is mostly applied on hairs, but it can be valuable for maternal lineage testing or analysis of any sample type having extremely degraded DNA. The analysis of degraded DNA (from hairs or other specimens) is even further improved when mini-amplicons are used for the mtDNA analysis [99,100].

All these different markers have their specific niche in which they can produce valuable results. Except for the core STRs, none of these markers are stored in the (national) DNA databases (yet), and thus for now their use is limited to one-to-one comparisons.

Other options for human identification beside comparative DNA analysis

Next to comparative DNA analysis, the main characteristics for human identification are fingerprints, dental status, facial recognition and body marks (such as scars, tattoos and medial implants) [101], but when reference data are absent, body parts are missing, or bodies are beyond recognition, these methods fail. When skeletal elements are present, physical anthropologists may be able to determine the sex and estimate the age at death and stature of the person, together with possible diseases he or she has gone through. For example, in an airplane crash in which several members of one family died, this information might help to discriminate between siblings of the same gender, but with different ages (who are not discernable based on their STR profiles without reference samples for direct comparison). Physical anthropology may also aid a facial reconstruction, which can be further supported by DNA-based information from the analysis of SNPs that are indicative of externally visible characteristics such as eye colour, hair colour and skin colour, and for which several multiplexes have been proposed in forensic literature [102,103,104,105,106,107,108]. Biogeographic ancestry markers cannot be used for this purpose, as they only pertain to geography and not to physical appearance [109]. Other externally visible characteristics that are being investigated include freckles, hair morphology, boldness, body height and facial
morphology [110], but further developments are needed before these can be used in human identification research. Especially body height (which is in addition strongly affected by non-genetic nurture effects) and facial morphology seem influenced by many genes [111,112] and can probably not be predicted by a few markers. Technologies that can detect many SNPs in one sample, such as micro-arrays and next generation sequencing, have the drawback that they generally need high amounts of input DNA (which is usually not present in human identification or forensic samples). Though, also in this area advances are being made as exemplified by the Identitas v1 Forensic Chip, which allows simultaneous analysis of appearance traits, biogeographic ancestry, relatedness and gender by analysing a total of 201,173 autosomal, X-chromosomal, Y-chromosomal and mtDNA SNPs, and for which positive results (genotype call rate >90 %) were reported for a sample with an input of 175 pg [82].

A conceptually difficult task is to differentiate between monozygotic twins, as their appearance is generally similar and their DNA sequence is virtually equal. Differences between fingerprints [113] and earprints [114] of monozygotic twins have been described and can be used when the decomposition status of the body allows collecting them and references for direct comparison are available (though earprints are probably more interesting for forensic burglary cases). Although the genotype of monozygotic twins is very similar, copy number variations between their genotypes were detected [115,116,117], which makes these markers interesting for further examination. Epigenetic differences in the genomic distribution of 5-methylcytosine DNA and histone modifications between monozygotic twins have been shown to accumulate with age (and to affect gene-expression) [118,119], and might be used to discriminate between them. Both for CNVs and epigenetic markers, additional research into the stability for these markers in different tissue types needs to be performed [120], as in forensics DNA from a buccal swab is compared to that derived from, for instance, a bloodstain. Another potential research topic focuses on personal variation for immunoglobulins produced during immune responses. Evidently, individual (disease) history determines to which antigens a person was exposed, but the most prominent source of variation derives from the biological process by which memory cells are generated from naïve B-lymphocytes. This involves the stochastic recombination of V, D, and J elements present on the immunoglobulin heavy (IgH) chain locus, which takes place in maturing B-cells. This recombination process results in extreme variation at a 45 bp region and this region can serve as a marker for human identification [121]. Interestingly, both disease history and VDJ recombination are believed to have inter-person variation, even between monozygotic twins. All these DNA-based markers clearly need further research and optimisation of techniques to enable analysis of low template amounts and/or degraded DNA, but they provide interesting starting points for research on monozygotic twin discrimination. Isotope ratios are non-DNA markers that could potentially be used to discriminate between monozygotic twins.
that do not live in the same area. Drinking water and eating patterns have an influence on hydrogen ($\delta^2$H) and oxygen ($\delta^{18}$O) isotope ratios of organic matter; and the $\delta^2$H and $\delta^{18}$O values in precipitation and tap water vary along geographic gradients [122]. When, for example, hairs are being analysed, information can be obtained about the area in which someone has lived recently [122]. In contrast, bones are formed much slower and might present information about the living area roughly ten years before death, and teeth about the area in which someone has grown up (as dental crowns (except for the third molars) are formed by the age of 8 years) [123]. This information might assist human identification.

Another area of research focuses on age prediction of an unidentified person (deceased victim or perpetrator of a crime). Physical anthropology (combined with odontology, counting of tooth cementum annulations [124] and microscopic analysis of the bone structure [125]) can give a broad estimation of the age at death (depending on the number of markers that could be assessed). In addition, other methods have been explored in search for more accurate or complementary predictions. Age-dependent changes in telomere repeat length and accumulation of mutations and deletions in the mtDNA have been proposed as markers, but did not show the accuracy needed for human identification or forensic purposes (e.g. a standard deviation of 22 years for the telomere approach by Karlsson et al. [126,127]. The abundance of sjTREC molecules (single joint T-cell receptor excision circles) can act as an age indicator from blood sources, but the prediction shows a standard error of ± 9 years [128]. Thus far, the most promising DNA-based age predictions were derived from a model based on the analysis of three DNA methylation markers that showed age dependent changes in saliva samples from donors between 18 and 70 years old, with an average accuracy of 5.2 years [129]. The method used in this paper (bisulfite conversion before PCR amplification and detection by pyrosequencing) though, requires high amounts of template DNA and needs to be redesigned before it can be used in human identification analyses. When teeth of the deceased are available for examination, chemical analysis of the tooth dentin by aspartic acid racemisation can give age at death estimations with an overall absolute error of 5.4 ± 4.2 years, and even better estimations can be obtained from radiocarbon ($^{14}$C) analyses of the enamel with an overall absolute error of 1.0 ± 0.6 years [130]. Thus, advances in age at death estimations are being made and it might be worthwhile to assess the possibilities of a combinatorial approach (if enough sample is available) to obtain higher age estimation accuracies.

Another time-related aspect that can be of importance for human identification and forensic research is the post-mortem interval (PMI). Numerous approaches for PMI determination have been suggested of which most relate to the first hours or days after death (e.g. based on body temperature, metabolic changes or mRNA stability). Entomology can be used for PMI determination of weeks to months after death. However, for human identification purposes especially the longer time periods
in the range of months to years are of interest. Although DNA degrades with time, several studies have shown that there is no usable correlation between the amount of DNA or the DNA fragment length and PMI (e.g. Box 2.7 and [131]), which is probably due to the huge influence of environmental factors. A study into the nucleoside content of 5 to 40 year-old museum moth specimens showed promising results with a time-dependent loss of deoxyguanosine, measured by mass spectrometry [132]. We performed a pilot study with a similar setup using human tissue samples of 2 to 40 years old, but did not detect the decrease in deoxyguanosine as described in the study by Zimmermann et al. (Box 2.7). Thus far, no reliable method for late PMI determination has been described and further research into this subject is needed.

A related forensic subject estimates the time since deposition of biological evidence. Many studies have examined the ageing of bloodstains using biochemical or physical methods [133]. For instance, UV-VIS spectrophotometric analysis and hyperspectral imaging of dried bloodstains showed a spectral shift with increasing stain age, and permitted to distinguish between bloodstains that were deposited minutes, hours, days and weeks prior to analysis [134,135,136,137]. The proposed methods are moving from a fully experimental phase to tests on simulated crime scenes, and by using a reference dataset the median relative error is reduced to 13 %. The effects of environmental factors (temperature and humidity) are being assessed, but when unknown it is possible to determine the order of formation of bloodstains [133]. Thus, these non-invasive/non-contact methods are moving slowly towards forensic practise. Not only time since deposition, but also clock time at the moment of deposition might provide forensically relevant information. A pilot study showed that the time of day (or night) that a bloodstain was deposited could be estimated in the order of about 4 to 5 hours, based on the 24-hour concentration fluctuations of circadian hormones [138]. For this study melatonin and cortisol were tested in fresh and aged bloodstains, and evaluation of additional circadian biomarkers in several body fluids might hold promise for a more accurate system to determine the time of deposition.

Should we do everything we technically can?

The introduction of novel markers and innovative technologies will bring new possibilities and opportunities for human identification. However, these advances also pose ethical and legal questions. For instance, when the analysis of large SNP arrays and whole genome sequencing becomes possible on low amounts of template DNA (such as usually present in human identification and forensic samples), possibilities open up to not only analyse the commonly accepted markers for identification, but also the predisposition for diseases and/or behavioural traits that may influence criminal behaviour. In some cases in the US and the UK, in court the defence has asked for
behavioural genetic analyses, which (in combination with other evidence) has led to a reduction of sentence, though not to acquittal, in a few cases. Although it is technically possible to type the behavioural genetic markers that are known, many factors (such as the effects of unknown genes, the interactions between genes and influences of nurture) remain unknown and other aspects (such as differences between population groups) may be socially unacceptable (regarded racism) [139]. These aspects make this kind of research delicate and complex at the very least.

When markers became known for the analysis of someone’s biogeographic ancestry, questions were raised on whether this information would be used rightfully to focus investigations and not misused to justify the targeting of certain racial groups [110]. In fact, an example of the opposite situation has been obtained in the (Dutch) case of the murder on Marianne Vaatstra, where asylum seekers (from Northern Africa and the Middle East) were initially suspected of the crime, while ancestry information pointed towards a perpetrator of Western European descent (and later on indeed a Caucasian suspect was matched to the evidentiary trace profiles). The Netherlands is currently the only country in which determination of biogeographic ancestry and externally visible traits (gender and eye-colour; at present) is explicitly allowed by law. Another research possibility that has recently been allowed by Dutch law is familial searching. In the aforementioned murder case, such a familial search within the national forensic DNA database did not provide new investigative leads, and based on tactic information the decision was made to perform a mass DNA screening (with the aim to find investigative leads towards the perpetrator through participating family members). 8080 men, who lived in the vicinity of the crime scene at that time, were asked to voluntarily participate in this research and the high attendance rate of 89 % shows the social involvement of people in that area. Based on a Y-chromosomal and autosomal match with biological material found on the crime scene, this research has led to the arrest of a suspect (and his subsequent confession, although the case has not yet been closed at the time of writing) thirteen years after the crime was committed, which shows the impact that changes in legislation can have.

The above-described case has renewed the discussion on whether everyone (in the Netherlands) should be in the national DNA database. Technically speaking this is possible since from almost every child (around 500 births per day [140]) FTA blood cards are sampled to test a variety of genetic diseases, and (in theory) it should be possible to collect one additional card for DNA databasing-purposes. Proponents of this idea bring forward that more crimes could be solved and that there is no reason for fear if you are innocent. Opponents refute that even if you are innocent there is a chance that your DNA matches the DNA found on a crime scene, which may put you in a position in which you need to prove your innocence (which is in sharp contrast with the presumption of innocence [141]). A similar situation occurs when during crime scene investigations unintentionally non-crime-related pieces of evidence are
collected as well. Another argument is that the people carrying out the research for the DNA database have access to your DNA without your consent, which could be seen as an intrusion of personal privacy [142]. Next to the considerable costs entailed with establishing a national DNA database, aspects like the storage time of such samples need to be considered; should they only be used to produce an STR profile and be destroyed afterwards, or should the samples be stored to make an upgrade to new markers possible in the future (the current legal storage time for certain samples in the DNA database is 80 years)? When the latter option should be chosen, misuse of such a database by future (malicious) politicians and health insurance companies might be a risk. On top of that, currently, the DNA profiles in the national DNA database are being exchanged daily with other European countries under the Prüm Convention, and it should be considered whether all DNA profiles are exchanged (which makes them also known to international authorities) or only those being crime-related.

In addition to the ethical and legal questions that may arise from marker development and technical advancements in DNA research, there can also be questions from a criminalistic point of view. Although techniques may become so sensitive that reliable DNA profiles can be made from single (or few) cells, an important question remains to be answered: "What is the criminalistic value of a single cell on a crime scene?" The mere fact that someone’s DNA is found on a crime-related objected does not make him or her the perpetrator of the crime; biological material may be left at the scene before the crime was committed, or may have been deposited indirectly. In forensic literature, the latter is termed secondary (or even tertiary) transfer [143,144]. Based on, for instance, presumptive testing or mRNA profiling, the cellular source of a sample can be elucidated [145], and it is evident that the finding of certain cell types (e.g. semen) may be more incriminating than others. Thus, when the sensitivity of forensic analyses is raised to the single cell level (assuming that advances have been such that single cell typing does not suffer from quality loss), the holistic or integrated interpretation of the evidence may become highly complex. Forensic scientists have the responsibility to not only point out the possibilities, but also the limitations of the available techniques irrespective whether these reside at the technical or interpretation level.
References


General discussion


M. de Pagter, Exploring copy number variation as a forensic marker system (MSc thesis). (2009).


