

# Considerations by the European DNA profiling (EDNAP) group on the working practices, nomenclature and interpretation of mitochondrial DNA profiles<sup>☆</sup>

G. Tully<sup>\*</sup>, W. Bär, B. Brinkmann, A. Carracedo, P. Gill, N. Morling,  
W. Parson, P. Schneider

*The Forensic Science Service, Trident Court, Solihull Parkway, Birmingham Business Park, Solihull B37 7YN, UK*

Received 20 December 2000; received in revised form 5 July 2001; accepted 30 July 2001

## 1. Introduction

Mitochondrial DNA (mtDNA) analysis is widely used in forensic analysis. This paper considers nomenclature and interpretation issues further to the guidelines from the DNA Commission of the International Society for Forensic Genetics [1]. Since the technique is very sensitive to low levels of DNA, laboratory conditions must be arranged to minimise the possibility of contamination. Accordingly, we consider working practice; provided that good laboratory practice is followed, reliable results will be achieved. Additionally, standardisation of nomenclature is important.

MtDNA has a much higher mutation rate than nuclear DNA; mutations and heteroplasmy are relatively common occurrences. This means that the evidential and reference samples may legitimately contain differences even when the two are in reality from the same individual or lineage. If two samples do not match, yet are purported to have the same origin, the strength of the evidence is reduced, but how best to report such an example?

### 1.1. Characteristics of mitochondrial DNA

MtDNA is a double stranded closed circular molecule of 16,569 base pairs in length. The proportions of the four nucleotides are unequal; in addition, one strand contains more guanine and adenosine bases (the heavy strand or H-strand) whilst the other strand (the light strand or L-strand) is rich in cytosines and thymines.

The control region of human mtDNA, which is the region most commonly analysed for forensic purposes, does not

code for proteins, tRNAs or rRNAs. However, it contains the L- and H-strand promoters, transcriptional regulatory elements, binding sites for mitochondrial transcription factors, the origin of H-strand replication, and the termination associated sequence (TAS) [2–14]. These elements would be expected to be under greater selective pressure than areas with no function.

The most common type of mutation within the control region is the single base substitution, and transitions outnumber transversions by approximately 40:1. Small insertions and deletions are common in the two homopolymeric regions (poly(C) regions between 302 and 310 and between 16,183 and 16,194); the chance of observing insertions and deletions at homopolymeric regions increases as the length of the uninterrupted homopolymer stretch increases.

Over evolutionary timescales, some base positions appear to be very stable while others are highly mutable [15–22]. Furthermore, some body tissues, such as hairs, tend to show more variability in their mtDNA sequence. This could be due to differential segregation of pre-existing heteroplasmic variants, to accumulation of new somatic mutations or to a combination of both phenomena.

MtDNA is primarily maternally inherited. Although recent and somewhat controversial papers have suggested that there may be some transfer of paternal mtDNA into the fertilised egg, with recombination leading to hybrid mtDNA molecules [23,24], the effect, if it occurs, appears to be minimal in the control region on time scales relevant to forensic comparisons [25].

These characteristics are important for the interpretation of mtDNA results.

With interpretation, there will be several different methods that could be used. Research continues to improve our understanding of the mutation and population genetics issues, hence these considerations will need to be expanded at a later date.

<sup>☆</sup> On behalf of the European DNA profiling (EDNAP) group.

<sup>\*</sup> Corresponding author. Tel.: +44-121-329-5431;

fax: +44-121-622-2051.

E-mail address: gtu@fss.org.uk (G. Tully).

## 2. Working practice and quality assurance

In order to minimise the occurrence of contamination, and to ensure that contamination does not cause erroneous results to be reported, a strict anti-contamination practice must be in place. We consider the following to be important:

- Pre- and post-PCR areas should be separated.
- Crime and reference samples should be analysed separately. In smaller laboratories, where space is limited and crime and reference samples cannot be analysed in separate rooms, analysis of crime samples should be carried out prior to analysis of reference samples.
- Positive displacement pipettes or aerosol-resistant tips should be used for all pre-PCR manipulations.
- To the extent possible, plasticware used for pre-PCR operations should be rendered DNA-free before use, preferably by UV cross-linking. Such treatment is good practice even when the manufacturers supply their products as DNA-free, since studies have shown that contaminating DNA can still be detected [26].
- To the extent possible, solutions used for pre-PCR operations should be rendered DNA-free before use, by UV cross-linking where applicable.
- Negative controls should be included for extraction and PCR stages, and these controls should be carried through the sequencing process.
- A positive control should be introduced at the PCR stage, and should be carried through the sequencing process.
- Where sufficient sample is available, duplicate extraction and analysis should be performed; such repetition may take the form of analysis of two sections of a single hair, or analysis of two different hairs from a single item, if the assumption that these hairs have a common donor can be made. Furthermore, analysis of both strands of the mtDNA is considered to be the best standard for ensuring that the sequence determination is accurate. It is the opinion of this group that in the event of duplication failing, or where insufficient sample is present, the report should clearly state the limitations of an unconfirmed result. It is possible to have apparently clear negative controls and yet to obtain an erroneous result due to contamination. This has been demonstrated in studies of non-human samples undertaken as part of validation exercises for mtDNA [27,28].
- The sequences of laboratory personnel should be determined. It is likely that in a proportion of cases, the sequence of the operator will match that of the questioned sample; this does not invalidate the result if the analysis has been duplicated, but alerts the scientist to the possibility of contamination. Ideally, different operators should carry out the duplicate extractions, but this may not be feasible for small laboratories.
- Competency testing of operators and proficiency testing programmes to test the efficacy of the procedures in place are highly desirable.

These procedures will not completely eliminate the occurrence of contamination, but will minimise its occurrence and, more importantly, the policy of duplication will ensure that erroneous results are not reported.

## 3. Nomenclature

### 3.1. General principles

The universal standard for mtDNA nomenclature is derived from the first published complete mtDNA sequence [29]. Each of the bases in the mitochondrial genome was given a consecutive number beginning near the origin of heavy strand (H-strand) replication at number 1 and ending at 16,569. It is referred to simply as the “Anderson sequence” or as the “Cambridge Reference Sequence (CRS)”, although it does not originate from a single individual. An analysed sample can be reported as a list of differences from the Anderson sequence. The base on the cytosine-rich light strand (L-strand) is always quoted. For example position 280 is C in the Anderson sequence; if the sequenced base is T then the base difference is denoted and the nomenclature is 280T. If Anderson base 281 is identical to that of the sample analysed, then there is no need to record the fact — i.e. it is implicit that the sequence of the sample is the same as the Anderson sequence unless noted. Although the Anderson sequence is not representative of most sequences (notably at positions 263 and 315.1) and contains several sections of bovine and HeLa sequence, its use is considered to be optimal for control region sequence, as a published and readily available reference. Future use of mtDNA sequence outside the control region is being developed for forensic use. As previous cases would be unaffected and convention is less established outside the control region, it is the opinion of the authors that the “Revised Cambridge Reference Sequence” [30] would be the optimal reference sequence for non-control region sequence. The following are more detailed considerations of this group.

### 3.2. Insertions and deletions

If deletions are observed relative to the Anderson sequence, they would be reported as the position deleted (e.g. if the base between 245 and 247 was deleted, this would be listed as 246d). If an insertion is observed, it would be designated with a “.1” after the lower numbered base of the two between which it has inserted (e.g. if an additional adenosine base was observed between 245 and 246, it would be designated 245.1A). If an insertion occurs within a homopolymeric tract (i.e. several tandem C bases), the exact location of the insertion is unknown. The standard generally adopted is to assume that the insertion has occurred at the highest numbered end of the stretch. For example, an insertion in the C-stretch between positions 302 and 310 would be designated 309.1C; two insertions here would be listed as 309.1C and 309.2C.

### 3.3. Heteroplasmy

#### 3.3.1. Point heteroplasmy

If clear heteroplasmy were observed such that two bases are present at approximately equal intensities, the appropriate IUB designation may be used (e.g. C and T at position 152 would be designated 152Y). Alternatively, the designation  $T \sim C$  could be used.

However, distinguishing heteroplasmy from background 'noise' or from sequencing artefacts is not always straightforward. The optimal way in which to ensure that high quality sequence is obtained is by sequencing both strands of the mtDNA. If two bases, clearly above background levels, are visible on both strands of sequence, the position can be confirmed as heteroplasmic. If the level of heteroplasmy is comparable in the two reactions and one base is at a substantially higher proportion than the other, notation of the type  $C > T$  may be employed. If the presence of two bases cannot be confirmed by a second sequencing reaction, the position would be designated as ambiguous (N) in the relevant extract.

#### 3.4. Length heteroplasmy

A similar approach could be employed for defining whether or not length heteroplasmy is present. This is a particular concern in the polycytosine stretches in both HV1 (between 16,183 and 16,194) and HV2 (between 302 and 310). If 'out of register' sequence (indicating the presence of more than one length variant) is observed after a C-stretch in one sequencing reaction, the second sequencing reaction would confirm the observation. In the case of polycytosine stretches, the second sequencing reaction would often be of the same strand as the first because of the commonly encountered difficulty in sequencing through long polycytosine tracts. If the number of cytosines present can thus be confirmed, nomenclature of the type 309.1 > 309.2 could be employed. However, if it were not possible to confirm the presence of a mixed number of cytosines by two amplification and sequencing reactions, the number of cytosine bases would be reported as ambiguous.

#### 3.5. Alternative approaches

Hühne et al. [31] have described an approach in which a specific threshold level for heteroplasmy is defined. However, setting a realistic threshold which is applicable across different laboratories employing different sequencing chemistries would be extremely difficult due to differences between sequencing chemistries in their ability to detect heteroplasmy [32]. We, therefore, favour interpreting each sequence individually, taking into account the background 'noise' over the entire sequence and using all available knowledge and experience of the characteristics of the sequencing chemistry in the hands of the individual laboratory.

## 4. Interpretation

### 4.1. General principles

The role of mtDNA analysis is to provide evidence where nuclear DNA fails to give a result, or when only maternally related references are available. Typically, tissues almost devoid of nuclear DNA are utilised. The evidence types most commonly analysed are bone and hair. Evidential material (Q) is compared with a known sample (K). As a rule, mtDNA is maternally inherited without recombination, which means that the donors of Q and K may be different individuals, e.g. mother and son. This type of analysis is often used to identify decomposed bodies. Several generations may have passed between the originator of Q and the donor of the reference sample (K). Although leakage of paternal mtDNA into the egg and recombination with maternal mtDNA may occur [23,24], maternal passage of mtDNA has been demonstrated through multiple extended lineages with no evidence of paternal contribution, e.g. [33]. For practical purposes in identification cases, therefore, maternal inheritance can be assumed. The second kind of analysis is to determine whether or not there is evidence of association between material recovered from a crime scene and reference material from a suspect. In this case, the donor of K is the putative donor of Q.

The aim of a mtDNA analysis is to provide evidence to support one of two alternative propositions:

1. the contention that the evidential sample (Q) originates from the suspect (the donor of K) or a maternally linked relative;  
or:
2. the contention that the evidential sample (Q) and the suspect's sample (K) originate from different individuals (of different maternal lineage).

If Q and K match, then this is evidence to support the first contention. If two samples do not match then this is evidence to support the second contention. However, defining whether or not Q and K could have originated from the same maternal lineage is not always straightforward if there is only a slight mismatch between them.

### 4.2. Criteria for inclusion and exclusion

It is well established that the mitochondrial mutation rate is substantially higher than that encountered with nuclear DNA. Consequently, it is not uncommon for differences to be observed in the DNA sequence when comparing close maternal relatives (such as mother and child) [33]. Substitution has also been observed in somatic tissues, presumably due to segregation of existing heteroplasmy within the individual. This means that differences may be observed between different hairs/tissues within an individual [34–36]. Consequently, if there are mismatches between Q and K this does not *automatically* exclude,

although the strength of the evidence is less than that of a match.

Mutations are passed between generations in varying ratios and segregate during development and later life; mutations also accumulate and segregate during the lifetime of an individual (reviewed in [37]). This results in mixtures of mtDNA molecules that characteristically differ from each other at one or more bases. This is known as heteroplasmy. Heteroplasmy probably exists in all individuals, although it is often at such a low level that it cannot be detected by the routine sequencing techniques presently used. In order for a mutation to be detected by sequencing, it must be present at a level approaching 20% to be distinguished from background. In addition, the chance of detection of heteroplasmy is dependent upon the sequencing chemistry used. Furthermore, detection may be more efficient at certain nucleotide positions than at others, and differences in detection may be also observed between the two DNA strands.

Once a mutation has led to significant heteroplasmy in the germ line, the offspring will either be similar to the mother, or else fixation will occur, so that heteroplasmy can no longer be detected, i.e. heteroplasmy is an intermediate state.

If Q and K do not match at a base position, it is clear that the strength of the evidence will be dependent upon the inherent mutability of that base. The current level of knowledge regarding substitution rates at each position within the mtDNA non-coding region is limited. Nevertheless, estimation of substitution rates in the control region is assisted by:

- observations of both heteroplasmy and substitutions from germline and intra-individual mutation studies [31–36,38,39];
- observations of the occurrence of heteroplasmy in case-work and research;
- inference from phylogenetic studies [15–22].

For example, long polycytosine stretches (between positions 302 and 310 and between 16,183 and 16,194) are extremely mutable since substitutions are observed in at least one out of 10 individuals (unpublished observations). At the opposite end of the spectrum, position 73 is relatively stable [20]. Thus, a difference at position 73 may provide stronger evidence for exclusion of K as a possible the source of Q than two base differences between K and Q if these are at positions 309.2 (homopolymeric region) and 16,093 (an apparent mutation “hotspot” that has been observed to vary in studies from several laboratories) [32].

However, despite such estimates, precise values for substitution rates are difficult to determine for the following reasons:

1. segregation of mutations will occur at different rates in different tissues;
2. there may be sequence context specific variations in substitution rates;

3. paternal inheritance and recombination, if they occur at appreciable rates, may bias inference from phylogenetic studies.

In applying a likelihood ratio approach, e.g. [37], the assessment of match criteria is made in the numerator:

probability of obtaining SQ and SK| evidential sample originated from suspect where SQ is the sequence of the evidential sample and SK is the sequence of the reference sample from the suspect.

If there are many differences between SQ and SK, this probability is effectively zero, and so an exclusion can be reported. Conversely, if SQ and SK match, the probability approaches 1. In the case of sequences that differ at one or two base positions, the value of the numerator is intermediate.

Note that if Q and K both have the same heteroplasmy, then this will increase the strength of the evidence for inclusion purposes, being a rather rare event if the donors of Q and K are unrelated [40].

#### 4.3. Assessing evidential strength when an inclusion is reported

Because the mitochondrial DNA is inherited as a haplotype, the sequence is interpreted as a single haploid locus; it is invalid to estimate evidential strength by multiplying the population frequencies of each of the bases in the sequence.

Currently most laboratories use the counting method where the number of matching sequences in the database of size ‘n’ is reported. This method presents the evidence in a purely factual way and is a statement of observation, there are no assumptions related to population genetics.

In order to use a likelihood ratio to estimate the strength of evidence from matching mtDNA profiles, the probability of obtaining the match conditional on the suspect not being the source of the evidential hair must be determined:

probability of obtaining SQ and SK| evidential hair did not come from the suspect

Evaluation of this probability requires an estimate of the frequency of occurrence of SQ in the offender’s population. Because mtDNA frequency databases are composed primarily of sequences from blood, this estimate would include the frequency of observation of SQ in a database of sequences from blood samples from the offender’s population, plus an assessment of the probability of obtaining SQ from individuals of blood sequence differing slightly from SQ. For example, if only one individual in a population of 100 has the sequence SQ, but 10 individuals have a sequence that differs from SQ at only position 309.1, then the probability of obtaining SQ and SK if the hair did not come from the suspect is clearly higher than would be estimated from the occurrence of SQ alone in the database. Nevertheless, this probability is substantially less than the combined frequency

of the two sequences, as a mutation must have occurred for any of the 10% of the population with the slightly differing sequence to have left the questioned hair. The ever-increasing body of data on mtDNA mutation rates and segregation will enable the significance of such a correction factor to be assessed. It is likely that the frequency of a sequence in a database of mtDNA sequences from blood samples closely approximates the frequency of that sequence amongst hairs, in which case the correction is unnecessary.

All of the above discussion, however, assumes that the frequency of all mtDNA haplotypes in the relevant reference population is known. This is not the case, and estimation of such population statistics from a database (a sample from the population) requires population genetics assumptions to be made. The following considerations are therefore important.

#### 4.4. Distribution of mtDNA types in populations

The migration rate must be high enough to effectively randomise mitotypes within the general population. There is evidence that this has occurred in European cosmopolitan populations [41] since  $F_{st}$  is low ( $<0.01$ ). On the other hand, in Asian populations  $F_{st}$  can be very high (up to 0.1) [42], this compares to a maximum  $F_{st} = 0.05$  observed in nuclear DNA in unusual inbred populations [43]. Whereas homogeneity assumptions may be reasonable to conclude in cosmopolitan situations, where the population is highly mobile, there is no information on small village populations in Europe where mobility may have been severely restricted over centuries. Most of the existing databases cannot address this question since they were collected over a wide area across a wide range of sub-populations (e.g. the UK database comprises individuals from all over the country). Small, relatively isolated European populations need to be analysed

in order to improve understanding of the population genetics of mitochondrial DNA at a local level. The assumption of homogeneity of mitotypes within the population may not hold for such populations. Provided that the data were available, this could be addressed by applying  $F_{st}$  corrections [44].

#### 4.5. Random sampling of the population

In practice, true random sampling is virtually impossible to achieve. The frequency of a mitotype ( $p$ ) is uncertain in a population because of sampling error. With relatively small databases, and large numbers of haplotypes, it is inevitable that new variants that do not appear in the database will frequently be observed. Balding and Nichols [44] have demonstrated that the uncertainty due to sampling error is approximately achieved by adding both the suspect and the criminal profiles to the database:

$$p = \frac{x + 2}{n + 2} \quad (1)$$

where  $x$  is the number of observations of the haplotype in the database size  $n$ .

An alternative method has been suggested by Holland and Parsons [37] using a 95% confidence limit from zero proportion:

$$p = 1 - \alpha^{1/n} \quad (2)$$

where  $\alpha$  is set to 0.05 for a 95% confidence interval and  $n$  is the sample size.

However, a comparison of the 95% confidence limit from zero proportion and Balding correction reveals very little difference between the two methods (Fig. 1).

When there is one or more samples in the database then use of confidence intervals has been recommended using

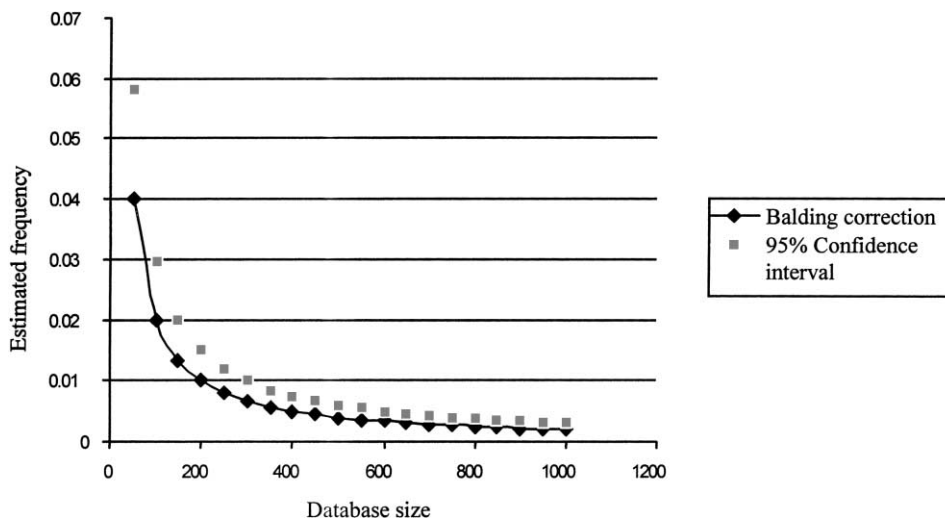


Fig. 1. A comparison of the Balding vs. confidence interval correction when a mitotype has not been observed in the database.

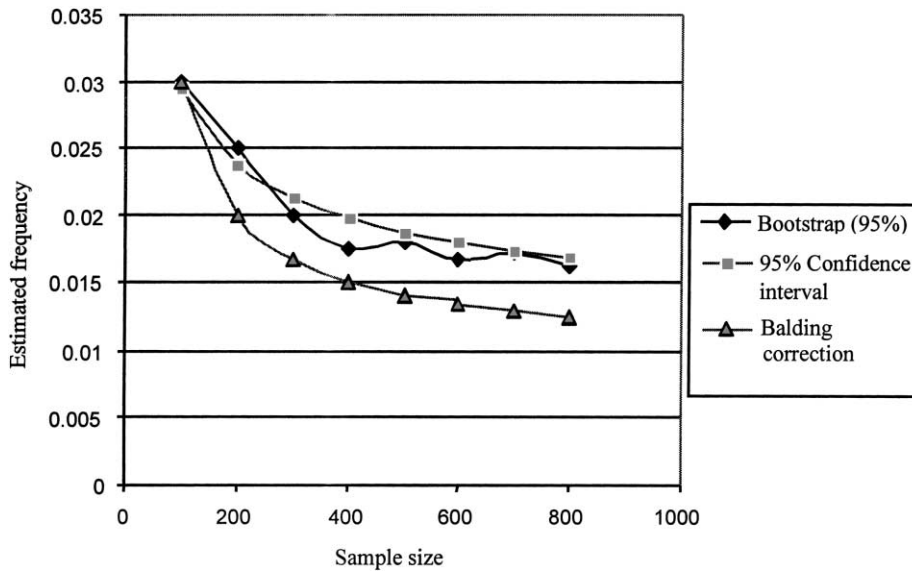


Fig. 2. A comparison of the Balding method (1) and confidence interval correction (3) for a mitotype frequency 0.01 in the database.

the formula:

$$p \pm 1.96 \sqrt{\frac{p(1-p)}{n}} \quad (3)$$

However, normal approximation to the binomial is assumed, which is unlikely to be good for small genotype proportions [45]. In Fig. 2, a comparison of the Balding method (1) versus confidence interval method (3) are shown for a mitotype where frequency = 0.01. The bootstrap analysis simulated a population size of 10,000 with 1000 samples taken with replacement. To calculate the 95 per-

centage of samples where the database most favourable to the defendant is employed, the strength of evidence will always be limited by the size of the smallest of the relevant databases.

#### 4.6. Assessing the evidence using a likelihood ratio framework

Although the current data regarding mtDNA substitution rates and population genetics are limited, the general likelihood ratio formulation is a useful framework upon which an assessment of evidential significance can be made:

$$LR = \frac{\text{Probability of obtaining SQ and SK|Q and K are from the same maternal lineage}}{\text{Probability of obtaining SQ and SK|Q and K are not from the same maternal lineage}}$$

centile, the results were ranked and the 950th sample was used. This gave similar results to the confidence interval method whereas the Nichols and Balding method gave lower estimates.

Thus, when a sample has not been observed in the database, this can be accommodated by either the Balding and Nichols method (Eq. (1)) or by a confidence interval (Eq. (2)). When one or more samples appear in the database, sampling error could be addressed using the Balding and Nichols method (Eq. (1)) or alternatively by a confidence interval, generated using Eq. (3) or using bootstrap analysis.

If suitable correction factors as described above can be applied, then either upper bounds of frequencies of occurrence or likelihood ratios could be quoted. However, the question of the relevance of the databases employed remains an issue. If the conservative approach of quoting from the

Examples of how this framework could be used in commonly encountered situations are given below. These situations are examples only; in each, the LR framework is used to assess the evidence, without exact values for any of the variables being known.

1. The known and questioned sequences match exactly and there are no observations of the same or similar sequences in the database: the findings support the hypothesis that Q and K are from the same individual or maternal lineage. The counting method, a frequency or a likelihood ratio (corrected for sampling error and distribution of mtDNA types in the population if applicable) could be used to help the court in assessing the strength of support for this hypothesis.
2. SQ and SK differ by a single base that is known to mutate frequently; neither sequence has been previously

observed in the database, and there are other sequences in the database differing by a single base: the findings support the hypothesis that Q and K are from the same maternal lineage, but the strength of this support relative to that in example 1 is diminished by a factor equal to the approximate substitution rate of the nucleotide by which SQ and SK differ.

3. SQ and SK differ by a single base at which substitution has previously been observed, but which does not appear to be a mutation “hotspot”; SQ has been observed once in a database of 100 sequences, SK has not been observed before, and there is an additional sequence in the database that differs at a frequently mutating base: taking into account the uncertainties in estimation of substitution rates and frequencies, the likelihood ratio is not likely to be significantly different to 1, so the findings are inconclusive.
4. SQ and SK differ at a position that has not been reported as variable and appears from phylogenetic studies to be stable; SQ and SK have both been observed several times in the database: the findings support the proposition that SQ and SK come from different maternal lineages.

The purpose of this paper is not to provide recommendations, but to discuss the options available to the practitioner. Each case should, of course, be treated on its own merits, using all available data regarding frequency of the sequences in question, mutation rates and structure of the relevant populations. Thus, estimated values for each of the variables in a particular case may be fed into the LR framework and an approximation for the LR calculated. In some instances, where population structure has been well studied, differences between sequences are at well-characterised nucleotide positions and sizeable frequency databases for the relevant population are available, the LR estimate will be sufficiently accurate to be quoted. In instances where the values are estimated with less certainty, the practitioner may opt to use a verbal expression of the strength of evidence as opposed to quoting the actual LR estimate.

## 5. Conclusions

The results of mitochondrial DNA analysis have been accepted as evidence in courts of law in Europe and North America, e.g. [46–48]. Despite diversity of methodologies, the ability to obtain consistent results and to compare results between laboratories has been demonstrated [49]. The present paper provides considerations on nomenclature to aid inter-laboratory comparisons and on working practices to ensure that the results reported are valid. Guidance on the issues to be considered when interpreting mtDNA evidence is given; as more data on mutation and population demography are gathered, it will be possible to refine the guidelines and to assign numerical values to quantities such as

substitution rates. In the absence of such values and of values for correction factors as described above, the counting method remains a valid approach for providing information to the courts. However, if the sequence of interest has not been observed in the database, it is incumbent on the forensic scientist to ensure that the court is not left with the impression that the sequence could be as rare as an STR profile at one in many millions.

## Acknowledgements

The authors wish to thank all of the members of the EDNAP group for useful comments and discussion. A list of contributing member laboratories is given in Appendix A. The EDNAP group worked in the period 1997–2000 within the framework of the STADNAP (Standardisation of DNA Profiling in Europe) consortium, a network project of the European Commission — DG XII programme “Standards, Measurement and Testing”. Contract 97-7506. The authors also wish to thank the anonymous reviewers for helpful comments.

## Appendix A. Contributors

Professor Dr. med. W. Bär, Institut für Rechtsmedizin, Universität Zürich-Irchel, Winthurerstrasse 190, CH-8057 Zürich, Switzerland.

Professor B. Brinkmann, Institut für Rechtsmedizin, Universität Münster, Von Esmarch Strasse 62, D-48149 Münster, Germany.

Professor Angel Carracedo and Dr. Maviky Lareu, Institute of Legal Medicine, San Francisco, s/n, E-15705 Santiago de Compostela, Spain.

Dr. Peter Gill and Dr. Gillian Tully, Forensic Science Service, Trident Court, Solihull Parkway, Birmingham Business Park, Solihull B37 7YN, UK.

Dr. Niels Morling, Department of Forensic Genetics, Institute of Forensic Medicine, University of Copenhagen, Frederik V's Vej 11, DK-2100 Copenhagen, Denmark.

Dr. Walther Parson, Institute of Legal Medicine, University of Innsbruck, Müllerstrasse 44, A-6020 Innsbruck, Austria.

Dr. rer. nat. Peter M. Schneider, Institut für Rechtsmedizin, Universität Mainz, Am Pulverturm 3, D-55131 Mainz, Germany.

Dr. Cristian Capelli, Istituto Medicina Legale, Università Cattolica, Largo Francesco Vito 1, I-00168 Roma, Italy.

Dr. Ate D. Kloosterman, Netherlands Forensic Institute, Volmerlaan 17, NL-2288 GD Rijswijk, The Netherlands.

Dr. Hermann Schmitter, Bundeskriminalamt, Thaerstrasse 11, D-65193 Wiesbaden, Germany.

Dr. Denise Syndercombe Court, Department of Haematology, St. Bartholomew's and The Royal, London School of Medicine and Dentistry, Turner Street, E1 2AD London, UK.

Dr. Maria Conceicao Vide, Instituto de Medicina Legal, Largo da Sé Nova, P-3000-213 Coimbra, Portugal.

## References

- [1] W. Bär, B. Brinkmann, B. Budowle, A. Carracedo, P. Gill, M. Holland, P.J. Lincoln, W. Mayr, N. Morling, B. Olaisen, P.M. Schneider, G. Tully, M. Wilson. DNA Commission of the International Society for Forensic Genetics: guidelines for mitochondrial DNA Typing, *Forensic Sci. Int.* 110 (2000) 79–85.
- [2] Y. Aloni, G. Attardi, Symmetrical in vivo transcription of mitochondrial DNA in HeLa cells, *P.N.A.S.* 68 (1971) 1757–1761.
- [3] W. Murphy, B. Attardi, C. Tu, G. Attardi, Evidence for complete symmetrical transcription in vivo of mitochondrial DNA in HeLa cells, *J. Mol. Biol.* 99 (1975) 809–814.
- [4] P. Cantatore, G. Attardi, Mapping of nascent light and heavy strand transcripts on the physical map of HeLa cell mitochondrial DNA, *N.A.R.* 8 (1980) 2605–2625.
- [5] D. Ojala, J. Montoya, G. Attardi, tRNA punctuation model of RNA processing in human mitochondrial DNA, *Nature* 99 (1981) 470–474.
- [6] D.D. Chang, D.A. Clayton, Precise identification of individual promoters for transcription of each strand of human mitochondrial DNA, *Cell* 36 (1984) 635–643.
- [7] J.E. Hixson, D.A. Clayton, Initiation of transcription from each of the two human mitochondrial promoters requires unique nucleotides at the transcriptional start sites, *P.N.A.S.* 82 (1985) 2660–2664.
- [8] H. Suzuki, Y. Hosokawa, H. Toda, M. Nishikimi, T. Ozawa, Common protein-binding sites in the 5(-flanking regions of human genes for cytochrome  $c_1$  and ubiquinone-binding protein, *J. Biol. Chem.* 265 (1990) 8159–8163.
- [9] H. Suzuki, Y. Hosokawa, M. Nishikimi, T. Ozawa, Existence of common homologous elements in the transcriptional regulatory regions of human nuclear genes and mitochondrial gene for the oxidative phosphorylation system, *J. Biol. Chem.* 266 (1991) 2333–2338.
- [10] K. Ohno, M. Takana, H. Suzuki, T. Obayashi, S. Ikebe, H. Ino, S. Kumar, A. Takahashi, T. Ozawa, Identification of a possible control element, Mt5, in the major noncoding region of mitochondrial DNA by intraspecific nucleotide conservation, *Biochem. Int.* 24 (1991) 263–272.
- [11] R.P. Fisher, D.A. Clayton, A transcription factor required for promoter recognition by human mitochondrial RNA polymerase, *J. Biol. Chem.* 260 (1985) 11330–11338.
- [12] R.P. Fisher, J.N. Topper, J.A. Clayton, Promotor selection in human mitochondria involves binding of a transcription factor to orientation-dependent upstream regulatory elements, *Cell* 50 (1987) 247–258.
- [13] S. Crews, D. Ojala, J. Posakony, H. Nishiguchi, G. Attardi, Nucleotide sequence of a region of human mitochondrial DNA containing the precisely identified origin of replication, *Nature* 277 (1979) 192–198.
- [14] J.N. Doda, C.T. Wright, D.A. Clayton, Elongation of displacement-loop strands in human and mouse mitochondrial DNA is arrested near specific template sequences, *P.N.A.S.* 78 (1981) 6116–6120.
- [15] G.B. Golding, Estimates of DNA and protein sequence divergence: an examination of some assumptions, *Mol. Biol. Evol.* 1 (1983) 125–142.
- [16] T.D. Kocher, A.C. Wilson, Sequence evolution of mitochondrial DNA in humans and chimpanzees: control region and protein coding region, in: S. Osawa, T. Honjo (Eds.), *Evolution of Life: Fossils, Molecules and Culture*, Springer, Tokyo, 1991, pp. 391–365.
- [17] M. Hasegawa, H. Kishino, T. Yano, Dating of the human-ape splitting by a molecular clock of mitochondrial DNA, *J. Mol. Evol.* 32 (1985) 37–42.
- [18] M. Hasegawa, S. Horai, Time of the deepest root for polymorphism in human mitochondrial DNA, *J. Mol. Evol.* 32 (1991) 37–42.
- [19] J. Wakeley, Substitution rate variation among sites in hypervariable region 1 of human mitochondrial DNA, *J. Mol. Evol.* 37 (1993) 613–623.
- [20] H.M. Wilkinson-Herbots, M.B. Richards, P. Forster, B.C. Sykes, Site 73 in hypervariable region II of the human mitochondrial genome and the origin of European populations, *Ann. Hum. Genet.* 60 (1996) 499–508.
- [21] L. Excoffier, Z. Yang, Substitution rate variation among sites in the mitochondrial DNA hypervariable region I of humans and chimpanzees, *Mol. Biol. Evol.* 16 (1999) 1357–1368.
- [22] S. Meyer, G. Weiss, A. von Haeseler, Pattern of nucleotide substitution and rate heterogeneity in the hypervariable regions I and II of human mtDNA, *Genetics* 152 (1999) 1103–1110.
- [23] A. Eyre-Walker, N.H. Smith, J. Maynard Smith, How clonal are human mitochondria, *Proc. R. Soc. Lond. B* 266 (1999) 477–483.
- [24] P. Awadalla, A. Eyre-Walker, J. Maynard Smith, Linkage disequilibrium and recombination in hominoid mitochondrial DNA, *Science* 286 (1999) 2524–2525.
- [25] T.J. Parsons, J.A. Irwin, Questioning evidence for recombination in human mitochondrial DNA, *Science* 288 (2000) 1931.
- [26] T. Schmidt, S. Hummel, B. Herrmann, Evidence of contamination in PCR laboratory disposables, *Naturwissenschaften* 82 (1995) 423–431.
- [27] M.R. Wilson, J.A. DiZinno, D. Polansky, J. Replogle, B. Budowle, Validation of mitochondrial DNA sequencing for forensic casework analysis, *Int. J. Leg. Med.* 108 (1995) 68–74.
- [28] J.M. Morley, J.E. Bark, C.E. Evans, J.G. Perry, C.A. Hewitt, G. Tully, Validation of mitochondrial DNA minisequencing for forensic casework, *Int. J. Leg. Med.* 112 (1999) 241–248.
- [29] S. Anderson, A.T. Bankier, B.G. Barrell, M.H.L. de Bruijin, A.R. Coulson, J. Drouin, I.C. Eperon, D.P. Nierlich, B.A. Roe, F. Sanger, P.H. Schreier, A.J.H. Smith, R. Staden, I.G. Young, Sequence and organization of the human mitochondrial genome, *Nature* 290 (1981) 457–463.
- [30] R.M. Andrews, I. Kubacka, P.F. Chinery, R.N. Lightowers, D.M. Turnbull, N. Howell, Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA, *Nature Genet.* 23 (1999) 147.
- [31] J. Hühne, H. Pfeiffer, B. Brinkmann, Heteroplasmic substitutions in the mitochondrial DNA control region in mother and child samples, *Int. J. Leg. Med.* 112 (1999) 27–30.
- [32] L.A. Tully, T.J. Parsons, R.J. Steighner, M.M. Holland, M.A. Marino, V.L. Prenger, A sensitive DGGE assay reveals a high frequency of heteroplasmy in hypervariable region one of the

- human mitochondrial DNA control region, *Am. J. Hum. Genet.* 67 (2000) 432–443.
- [33] T.J. Parsons, D.S. Muniec, K. Sullivan, N. Woodyatt, R. Alliston-Greiner, M.R. Wilson, D.L. Berry, K.A. Holland, V.W. Weedn, P. Gill, M.M. Holland, A high observed substitution rate in the human mitochondrial DNA control region, *Nature Genet.* 15 (1997) 363–368.
- [34] K.M. Sullivan, R. Alliston-Greiner, F.I.A. Archampong, R. Piercy, G. Tully, P. Gill, C. Lloyd-Davies, A single difference in mtDNA control region sequence observed between hair shaft and reference samples from a single donor, in: *Proceedings of the Seventh International Symposium on Human Identification*, Scottsdale (AZ), Promega Corporation, 1996, pp. 126–129.
- [35] M.R. Wilson, D. Polanskey, J. Repogle, J.A. DiZinno, B. Budowle, A family exhibiting heteroplasmy in the human mitochondrial DNA control region reveals both somatic mosaicism and pronounced segregation of mitotypes, *Hum. Genet.* 100 (1997) 167–171.
- [36] C.D. Calloway, R.L. Reynolds, G.L. Herrin, W.W. Anderson, The frequency of heteroplasmy in the HVII region of mtDNA differs across tissue types and increases with age, *Am. J. Hum. Genet.* 66 (2000) 1384–1397.
- [37] M.M. Holland, T.J. Parsons, Mitochondrial DNA sequence analysis — validation and use for forensic casework, *Forensic Sci. Rev.* 11 (1999) 21–50.
- [38] S. Siguroardottir, A. Helgason, J.R. Gulcher, K. Stefansson, P. Donnelly, The mutation rate in the human mtDNA control region, *Am. J. Hum. Genet.* 66 (2000) 1599–1609.
- [39] J. Hühne, H. Pfeiffer, K. Waterkamp, B. Brinkmann, Mitochondrial DNA in human hair shafts — evidence of intra-individual differences, *Int. J. Leg. Med.* 112 (1999) 172–175.
- [40] P.L. Ivanov, M.J. Wadhams, R.K. Roby, M.M. Holland, V.W. Weedn, T.J. Parsons, Mitochondrial DNA sequence heteroplasmy in the Grand Duke of Russia Georgii Romanov establishes the authenticity of the remains of Tsar Nicholas II, *Nature Genet.* 12 (1996) 417–420.
- [41] T. Melton, M. Wilson, M. Batzer, M. Stoneking, Extent of heterogeneity in mitochondrial DNA of European populations, *J. Forensic Sci.* 42 (1997) 437–446.
- [42] T. Melton, M. Stoneking, Extent of heterogeneity in mitochondrial DNA of ethnic Asian populations, *J. Forensic Sci.* 41 (1996) 591–602.
- [43] L.L. Cavalli-Sforza, W.F. Bodmer, in: D. Kennedy, R.B. Park (Eds.), *The Genetics of Human Populations*, Freeman, San Francisco, 1971.
- [44] D.J. Balding, R.A. Nichols, DNA profile match probability calculation: how to allow for population stratification, relatedness, database selection and single bands, *Forensic Sci. Int.* 64 (1994) 125–140.
- [45] I.W. Evett, B.S. Weir, *Interpreting DNA Evidence — Statistical Genetics for Forensic Scientists*, Sinauer, MA, 1998, pp. 142–146.
- [46] *The People of Tennessee versus Paul Ware*, 11 Judicial Dist. 203757, Div. 1 Criminal Ct. (1996).
- [47] *The Chattanooga Times*, Vol. CXXVII, no. 226, 4 Sept 1996, Ware found guilty could get death.
- [48] Sumario 1/93-Audiencia de Valencia against Migual Ricart Tarrega.
- [49] A. Carracedo, E. D'Aloja, B. Dupuy, A. Jangblad, M. Karjalainen, C. Lambert, W. Parson, H. Pfeiffer, H. Pfitzinger, M. Sabatier, D. Syndercombe Court, C. Vide, Reproducibility of mtDNA analysis between laboratories: a report of the European DNA profiling group (EDNAP), *Forensic Sci. Int.* 97 (1998) 165–170.