

Forensic Mitochondrial DNA Analysis: Two Years of Commercial Casework Experience in the United States

Terry Melton, Kimberlyn Nelson

Mitotyping Technologies, LLC, State College, Pa, USA

Aim. To describe mitochondrial DNA (mtDNA) forensic casework experience in a commercial laboratory in the United States.

Methods. Frequency statistics were kept for two years on all aspects of mtDNA forensic cases, including types of clientele, types of samples, levels of sample success and failure, site heteroplasmy, length heteroplasmy, contamination, rates of failures to exclude, and match statistics using a mtDNA sequence database.

Results. Low sample failure rate was observed, especially since an "ancient DNA" approach was used for samples with degraded DNA. Levels of contamination were low, and the observed site and length heteroplasmy did not confound the interpretation of results. The data collected from mtDNA haplotypes developed in casework showed extremely high diversity of haplotypes consistent with other formally developed databases.

Conclusions. MtDNA forensic analysis in the private sector was successfully applied to many different types of samples overall, with minimal rates of complication due to sample handling challenges (degraded DNA, minimal samples, contamination) and sequence-specific phenomena (site and length heteroplasmy).

Key words: *forensic medicine; laboratories; DNA, mitochondrial; private sector; United States*

Mitochondrial DNA (mtDNA) analysis is a growing area of forensic testing in many countries (1). Crime scene investigators, law enforcement, and prosecuting attorneys are just beginning to appreciate that this form of testing, while not as discriminating as nuclear DNA testing, may successfully advance the investigation and prosecution of cases with limited biological evidence, such as telogen hairs and degraded skeletal remains. Defense attorneys are also increasingly requesting testing of samples that may aid in exonerating their clients. In fact, one of the most exciting applications of mtDNA testing is in the area of post-conviction relief, as microscopically examined hairs that originally included a presumed perpetrator, even decades ago, are being revisited with mtDNA analysis, and in some cases are found to exclude the original suspect as a contributor.

Mitotyping Technologies, located in State College, Pennsylvania, USA, is a private company that exclusively offers mtDNA forensic analysis. It was established in October 1998 and began forensic casework testing in February 1999; as such, it has over two years of casework experience. Excluding both the Armed Forces DNA Identification Laboratory and the Federal Bureau of Investigation Unit II Laboratory (mtDNA unit), who were pioneers in forensic mtDNA testing, Mitotyping Technologies was the third US commercial laboratory to offer forensic mtDNA anal-

ysis. This private enterprise grew out of independently offered mtDNA testing services in the Department of Anthropology of Penn State University between 1989 and 1998. With our experience to date, Mitotyping Technologies has considerable data on the breadth of casework, limitations and benefits of testing, and outcomes, including match statistics and the incidence of site heteroplasmy, length heteroplasmy, and contamination. This information should be useful to groups investigating the possibility of starting a forensic mtDNA testing service.

Current staff at Mitotyping Technologies consists of two forensic examiners with a PhD degree, who testify in court, and two forensic technicians with an MS degree, who do not testify. There are no additional support or administrative personnel. During the first year of operation, the 84 square meter facility was not operating at capacity, but casework demand in the third year has been such that the company is hiring an additional technician, bringing the total staff to five. The statistics presented here will therefore be reflective of caseload capacity and potential throughput for other laboratories considering the setup of a mtDNA laboratory. Mitotyping Technologies is also accredited by the American Society of Crime Laboratory Directors/Laboratory Accreditation Board, and as such follows the comprehensive guidelines established by the DNA Advisory Board for forensic DNA

laboratories (2). Complying with these guidelines requires an additional two to four weeks annually that are devoted to maintaining internal quality control and assurance requirements, including external proficiency testing.

Overview of Protocols at Mitotyping Technologies

A comprehensive general description of forensic mtDNA analysis is provided by Holland and Parsons (3). In mtDNA testing, there are ongoing issues relative to the prevention of contamination and to optimization of recovery of minimal and degraded DNA because of the types of samples typically encountered, such as telogen hairs and skeletal remains. Whereas most laboratories providing mtDNA analysis have very similar protocols, those that are specifically mentioned below have been developed by Mitotyping Technologies to address difficult samples often encountered during testing. Two goals prevail: protection of the integrity of the evidence, including prevention of contamination at any stage of testing, and collection of the maximum amount of available mtDNA data inherent to any sample.

All samples, including DNA-abundant blood reference standards, are handled individually from the beginning of DNA extraction through sequencing. There are two primary reasons for this approach: 1) the probability of cross-contamination between samples is virtually negligible (and is difficult to allege); and 2) a high proportion of samples require individualized attention during the polymerase chain reaction (PCR) and sequencing phase either because of degraded DNA or because of sequence-specific variation, such as length heteroplasmy or, much more rarely, site heteroplasmy. In our experience, it is much more efficient to complete one sample at a time than to batch samples at any stage and play "catch-up" with individual samples to keep them in line with each other during mass processing. Our approach results in complete sequence data (~783 base pairs of double stranded sequence data in hyper-variable regions 1 and 2) being obtained for most samples by use of an ABI 310 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA).

Individualized handling is very valuable for samples that fail on the first amplification attempt of a standard ~250 base pair (bp) fragment, which is likely due to the presence of extremely degraded DNA. An "ancient DNA" approach is then used to amplify smaller fragments ranging in size from ~80 bp to ~140 bp that overlap within the HV1 region; this method often recovers abundant DNA in a sample that would otherwise be missed with standard protocols (4). This approach has been used successfully in skeletal remains cases, such as the mtDNA analysis of Neanderthal bones (5,6). At Mitotyping Technologies, this "ancient DNA" approach has also been used on telogen hairs and hair fragments. The disadvantage to this approach is that one must be more vigilant about contamination of reagents, especially primers, due to the enhanced proclivity for closely spaced primers to hybridize to degraded ex-

ogenous DNA from reagents or equipment. Therefore, it is usual for "ancient DNA" primers to be tested for cleanliness before every use on a sample, making this form of analysis more rigorous and expensive.

Casework Results

The statistics presented here represent approximately 24 months of forensic casework experience. Table 1 shows the distribution of clientele requesting mtDNA testing. Within this period, there were 105 laboratory cases completed, comprising 199 questioned (Q) or evidentiary samples, and 137 known (K) reference samples. The number of questioned samples in any case ranged from 1-9, and the number of known samples ranged from 1-5. Table 2 shows the distribution of sample types in questioned and known samples. For all samples, nuclear DNA testing had previously been determined to be an unsuccessful approach to the sample, or mtDNA testing was necessary to compare the sample to another sample that could only undergo mtDNA testing.

Failure to Obtain mtDNA Amplification

The Mitotyping Technologies has had good success with hair fragments as small as 3 mm, however, the success of any hair is dependent not only on hair length but hair diameter. For example, pubic hairs, which are often of substantial diameter, rarely fail in an analysis, whereas head hairs vary more in their success because the fragment supplied for testing may be more proximal to the root end (greater success rate) or more distal to the root (lower success rate).

Table 1. Distribution of clientele that have requested mtDNA analysis between February 1999 and February 2001

Client	No. of cases
Crime lab, law enforcement, prosecuting attorney, coroner or medical examiner	66
Private or public defense attorney	21
Private individual, attorney or investigator	13
Other institution such as parentage lab or academic	5
Total cases	105

Table 2. Types of casework samples encountered during 2 years of work in Mitotyping Technologies (between February 1999 and February 2001)

Tissue	No. of samples	
	questioned	known
Bone/tooth	31	1
Blood (liquid blood or prepared stain card)	4	111
Cigarette butt	5	0
DNAs (extracted in other laboratory)	5	3
Hair or hair fragment	130	14
Maggots	2	0
Saliva	0	5
Semen	1	0
DNA sequence (determined by another laboratory)	0	2
Bloodstain (miscellaneous substrates)	19	0
Tissue (muscle, organ)	2	1
Total	199	137

In 24 months of testing, 17 of 199 questioned samples failed to yield any mtDNA amplification product after both standard and "ancient DNA" analyses (8.5%). Cremated skeletal remains from two cases comprised five of these samples; although these remains retained some recognizable skeletal structure, all of the DNA apparently had been destroyed by high temperatures. Two stains and nine hairs failed to give results. Five of these hairs had a history of exposure to difficult environmental conditions before they were collected for testing: three had been undisturbed in a non-climate-controlled storage locker for 25 years, one was found under black electrical tape on an ax handle in a greenhouse nine years after the crime, and the remaining one was suspected to have been a "hair-weave" hair.

Partial Sequencing

In several cases, we could collect only partial sequence data on questioned samples using the previously mentioned "ancient DNA" approach. This commonly occurs because multiple PCR amplifications are necessary but each covers a small region, therefore exhausting the total supply of extraction product. In two of these cases, even partial data (<400 bp of double stranded sequence) were sufficient to make a match where the partial haplotype was either unique in the current database or had been observed only twice. These results greatly enhanced the investigation and prosecution of these two cases.

Contamination

Because each questioned or known sample was accompanied by its own parallel reagent blank negative extraction control throughout processing, data were collected on the frequency of contamination observed in these reagent blank controls. Reagent blank contamination was observed in 29 of 1,218 (2.4%) PCR reactions in casework. Mitotyping Technologies has never observed a reagent blank contaminant throughout all the PCR products generated in a single sample analysis. Reagent blank contaminants instead appeared in one of the amplifications only, suggesting that they were sporadic contaminants present in a disposable tip or tube used for that PCR only. These contaminants were sequenced where possible, although often the amplification product was so minimal that only partial sequence was observed or the sequence was uninterpretable. Mitotyping Technologies has never observed that a contaminant is a staff member's type or the type of a recently handled sample, which also suggests that these contaminants are sporadic and exogenous to the laboratory. In all but two cases where reagent blank contamination was observed, the contaminant did not confound the interpretation of results. In one of these, an "ancient-DNA" case, the reagent blank contaminant was clearly visible in a portion of the questioned sample's DNA sequence, creating an uninterpretable mixture. We also have observed that reagent blank contamination is disproportionately more prevalent with "ancient DNA"-type samples (20 of the 29 observations, whereas ancient DNA samples comprise only a small percentage of the number of samples handled). We have never had an incident of laboratory-wide perva-

sive, uneradicable, and unidentifiable contamination.

Length Heteroplasmy

Length heteroplasmy in the homopolymeric C-stretches is a commonly encountered phenomenon, which requires extra PCR amplifications or sequencing when it is severe, due to the uninterpretable electropherogram traces downstream of the heteroplasmy (7). Hypervariable region 1 (HV1) length heteroplasmy alone was observed 15 times and hypervariable region 2 (HV2) length heteroplasmy alone was observed 77 times, whereas 17 samples had both HV1 and HV2 length heteroplasmy. For these calculations, samples that shared the same type within a case were only counted once. Deletion heteroplasmy, or the presence of two types of mtDNA in a sample when one of the types has a single base-pair deletion, manifests as a mild and readable length heteroplasmy on electropherogram traces. It was observed on three occasions, twice due to a missing nucleotide at position 249 and once due to a missing nucleotide at position 71 within a G-stretch.

Site Heteroplasmy

Site heteroplasmy is defined as the appearance of one position with two nucleotide bases in an otherwise unmixed sequence. It was observed a total of 19 times (5.7% overall, 9.7% in hairs, 1.7% in whole blood, and 8.6% in tissue, bone, or organs) (Table 3). At Mitotyping Technologies site heteroplasmy is confirmed with at least two independent PCR reactions and double strand sequencing to rule out sequencing artifact or mixtures, except when it is observed at positions 16093 and 152. In none of the heteroplasmic samples noted in Table 3 were there indications of mixtures of two or more samples (as suggested by other ambiguous positions). The heteroplasmy itself was clearly visible as at least a 1:4 ratio of the two nucleotides on all strands. In six of seven cases where any additional questioned or known matched samples were available for comparison, either clear-cut substitutions or heteroplasmic observations at these sites in these additional samples bolstered the original call of heteroplasmy in the "proband" sample (Table 3).

Failures to Exclude

Failures to exclude an individual as donor of a biological sample were observed in 57 out of 105 cases (54.3%). Besides criminal casework and missing person investigations, this count includes parentage testing where two individuals believed to be maternally related were tested to confirm that relationship. It also includes five cases in which two different individuals were respectively matched to different questioned samples or one individual was matched to biological evidence from different crime scenes, thereby linking the crimes to a single suspect. The distribution of the count of observations of inclusions within the FBI mitochondrial DNA database in use at the time of testing is shown in Figure 1. For example, of 62 types developed in failures to exclude, 31 types had never been observed in the database at the time of testing (50%). The most common type in the database is observed in

Table 3. Site heteroplasmy observed in thirteen cases^a

Tissue	Observed heteroplasmy	Other matched K or Q sample for comparison?	Nucleotide present in comparison sample
Q hair	16093 T/C	K blood	C (substitution from CRS)
Q hair	16093 T/C	K blood	C (substitution from CRS)
Q hair	16166 A/G	K blood	A
K bone	189 A/G	No	-
Q hair	152 T/C	2 other Q hairs, 1 K hair	All 3 T/C heteroplasmy
K blood	207 G/A	No	-
K hair	16093 T/C	No	-
K blood	16286 C/T	No	-
K hair	279 T/G	No	-
Q hair	16093 T/C	No	-
Q hair	16093 T/C	1 other Q hair	T/C heteroplasmy
Q bone	72 T/C	K muscle	T/C heteroplasmy
Q hair	16093 T/C	1 other Q hair, K blood	T/C heteroplasmy in hair, C in blood (substitution from CRS)

^aAbbreviations: Q – questioned; K – known; T – thymine; C – cytosine; CRS – Cambridge reference sequence; A – adenine; G – guanine.

approximately 7% of Caucasians and has a substitution at position 263 and C-stretch variation around HV2 position 310. This type was seen in 3 matches (data not shown).

Laboratory Database

Mitotyping Technologies has an internal laboratory database assembled on the Microsoft Access platform used by the MitoSearch software, which is disseminated to mtDNA practitioners by the FBI (8). This database, at present, contains 116 complete mtDNA profiles (positions 15997-16400 in HV1 and 30-407 in HV2) that have been generated in casework. The profiles were entered into the database individually; in the event of a match between two or more samples within the same case, the matched profile was entered only once. A "Pairwise" analysis on complete profiles comprising bases 16024-16400 and 30-400 was carried out on the database (additional profiles with mostly complete or partial data could not be included in the analysis). The overall genetic diversity of this database was 0.997, whereas the random match probability was 1.17%. There were 21 matches in 6,670 pairwise comparisons, a frequency of 0.00315 or 1 in 318. The average number of nucleotide differences between pairs of haplotypes was 10.3 ± 4.9 . Table 4 shows the distribution of haplotypes observed within the database.

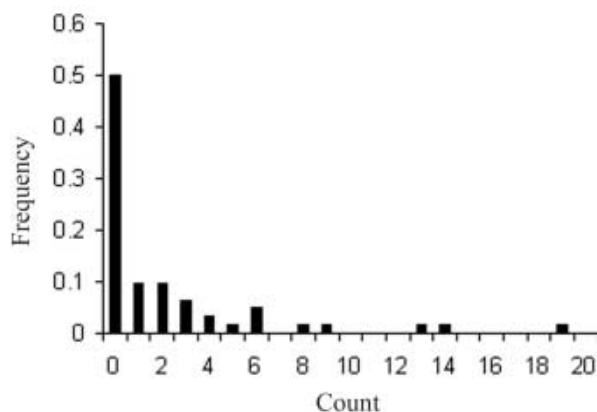


Figure 1. Distribution of match observations within the FBI mitochondrial DNA database in use at the time of testing.

Conclusions

Based on the case load of Mitotyping Technologies over two years, the rate of throughput is approximately one to two cases per staff member per month, with a system currently in place where samples are individually analyzed. It is questionable if this rate of throughput can be increased due to the requirement of many samples to have additional PCR and sequencing reactions when length or site heteroplasmy must be dealt with and when ancient DNA analyses are possible. With this system, we have been overwhelmingly successful in keeping contamination low and in developing nearly complete profiles for the majority of samples with adequate redundancy in PCR and sequencing to confirm the substitutions that characterize each sample.

Distribution of cases among kinds of clientele has remained stable over two years, with approximately a 1:3 ratio of defense to prosecution/law enforcement testing. What remains to be seen is whether a "bulge" of re-examined cold cases from decades past or post-conviction mtDNA testing of hairs will pass through the system over the next few years. A substantial proportion of submissions (exact number not known) have been from cases at least five years old. Investigators in the United States are still learning about the availability of mtDNA testing, so we expect caseload to increase steadily for some time to come.

Contamination has been a very minor problem at Mitotyping Technologies. We attribute this to the following factors: 1) individual handling of samples; 2) a clean-room environment with one-way flow of supplies and reagents from pre- to post-PCR areas; 3) restrictions that prohibit anyone besides staff inside all parts of the laboratory; 4) multiple glove changes during sample handling, abundant use of ultraviolet irradiation of surfaces, and frequent 10% bleach cleaning of surfaces; 5) handling of tubes during extraction with a new disposable wipe with every manipulation; 6) use of individually wrapped sterile Eppendorf tubes for extractions; and 7) use of bottled sterile tissue culture water for all reagents that come into contact with samples. Of the above, individual sample

Table 4. Distribution of 116 complete profiles in the laboratory database

No. of profiles	Frequency	Total	Haplotypes ^a
1	5	5	263 G
2	3	6	16362C 239C 263G 16126C 16163G 16186T 16189C 16294T 73A 152C 195C 263G
5	2	10	16192T 263G 16162G 16172C 16292T 16304C 73G 152C 249D 263G 16189C 16192T 16270T 16398A 73G 150T 263G 16069T 16126C 16193T 16278T 73G 150T 152C 263G 152C 263G
95	1	95	Not shown
Total	103	116	

^aProfiles do not include variation in the HV2 homopolymeric C-stretch.

handling is probably most responsible for the low observed levels of contamination.

MtDNA control region site heteroplasmy has been described in recent publications (9,10) and is sometimes alleged to be a potential confounding factor in the forensic use of mtDNA. On the contrary, we have not found it to be difficult either to detect or interpret. Detection is, of course, reliant on the sensitivity of instrumentation and may also be dependent on the particular chemistry used for sequencing. Therefore, it is possible that different laboratories may have different abilities to detect site heteroplasmy (10,11). To the extent that it is detectable, the single most important factor in detection is sequence quality. With little to no background artifact in electropherogram traces, site heteroplasmy is usually very obvious when the ratio of peak heights is at least 1:4. Hence, mtDNA laboratories must be vigilant about keeping sequence quality high.

Although site heteroplasmy at positions 16093 and 152 is that most frequently observed in biological samples (Rebecca Reynolds, unpublished data), additional heteroplasmy was observed at Mitotyping Technologies at positions 16166, 16286, 72, 189, 207, and 279. Of all these sites, 16093, 16166, 152, 189, and 207 have been previously identified as "fast sites", meaning they have higher mutation rates than others in the control region (12,13). Consequently, it is not surprising that these sites may be observed in a transitional heteroplasmic state in biological samples. In about half the cases observed, matched samples either had additional evidence of heteroplasmy or had substitutions at those locations. The sites noted above do not include the frequently observed T to C transition at position 16189, which sometimes manifests as a site heteroplasmy. It is also a "fast site" and often causes length heteroplasmy as well.

Site heteroplasmy is unlikely to routinely confound the interpretation of sample matches for two reasons: 1) it is somewhat uncommon, as shown by an observed frequency of 5.7%; and 2) on average, most samples contributed by different individuals differ at an average of 8.9 substitutions, based on the current MitoSearch database. Observations of samples within the same case that differ by only one base pair are rare. In fact, we have only observed this once, in a non-heteroplasmy case. The interpretation of this outcome is always that the results are inconclusive. Keeping in mind that the 5.7% frequency observed here includes those samples matching the proband samples that themselves were heteroplasmic, the ac-

tual rate of heteroplasmy in unmatched randomly selected samples is even lower.

The diversity that Mitotyping Technologies has observed in the North American mtDNA types developed from biological samples is very high. It is comparable to that observed in population studies both at the sequence level and with methods that have examined variation at only some informative control region sites (14,15). Within this relatively small database, the probability of selecting two samples at random that match is only 1 in 318, which compares favorably with the probability of 1 in 392 in the MitoSearch database that at present contains 4,142 sequences. Both the Mitotyping Technologies internal database and MitoSearch database contain multiple racial and ethnic groups. The samples in the Mitotyping Technologies internal database come from approximately 35 states and Canada. In addition, the most common type observed in Caucasians (with an A to G transition at 263 and variation in the HV2 C-stretch) was observed five times (4.3%), which falls near the frequency previously reported by others (3). This most-common-type frequency has remained stable in all issues of the MitoSearch database, through successively larger number of sequences. In the Mitotyping Technologies database, fully 81.9% of the sequences are unique, having been observed only once. Although the diversity of the Mitotyping Technologies database is not really relevant for our casework searches, since searches are done with the full MitoSearch database, it is however interesting to compare with the larger available compilation of sequences.

Thus far, casework experience has been quite straightforward at Mitotyping Technologies. Protocols were developed at the outset that optimized recovery of DNA and minimized sample contamination. In no small part we attribute the evolution of these successful protocols to our prior experience in academic labs that focused solely on mtDNA investigations, including ancient DNA challenges. Combining that experience with forensic DNA handling protocols has allowed for the synthesis of a successful commercial lab, and a challenging, gratifying environment for serving the criminal justice community.

Acknowledgments

We would like to thank Gloria Dimick, Bonnie Higgins, and Kim Seebart for expert technical assistance. Thanks also go to Dr Keith Monson of the FBI Academy for assisting with the Pairwise analysis of the laboratory database.

References

- 1 Carrecedo A, Bär W, Lincoln P, Mayr W, Morling N, Olaisen B, et al. DNA Commission of the International Society for Forensic Genetics: guidelines for mitochondrial DNA typing. *Forensic Sci Int* 2000;110:79-85.
- 2 U.S. Department of Justice. Federal Bureau of Investigation. DNA Advisory Board. Quality assurance standards for forensic DNA testing laboratories. 1998 Oct 1. *Forensic Science Communications* 2000; 2(3). Available at: <http://www.fbi.gov/hq/lab/fsc/backissu/july2000/codis2a.htm>. Accessed: March 13, 2001.
- 3 Parsons TJ, Holland MM. Mitochondrial DNA sequence analysis – validation and use in forensic casework. *Forensic Science Review* 1999;11:21-50.
- 4 Gabriel MN, Huffine EF, Ryan JH, Holland MM, Parsons TJ. Improved mtDNA sequence analysis of forensic remains using a “mini-primer set” amplification strategy. *J Forensic Sci* 2001;46:247-53.
- 5 Krings M, Stone A, Schmitz RW, Krainitski H, Stoneking M, Pääbo S. Neanderthal DNA sequences and the origin of modern humans. *Cell* 1997;90:1-20.
- 6 Krings M, Geisert H, Schmitz RW, Krainitski H, Stoneking M, Pääbo S. DNA sequence of the mitochondrial hypervariable region II from the Neanderthal type specimen. *Proc Natl Acad Sci USA* 1999;96:5581-5.
- 7 Stewart J, Fisher C, Aargard P, Wilson M, Eisenberg A, Polansky D, et al. Length variation in HVII of the human mitochondrial DNA control region. *J Forensic Sci*. In press 2001.
- 8 MitoSearch: mtDNA analysis software. Windows 95/98/NT. Version 2.1. Quantico (VA): FBI Academy.
- 9 Bendall KE, Macaulay VA, Baker JR, Sykes BC. Heteroplasmic point mutations in the human mtDNA control region. *Am J Hum Genet* 1996;59:1276-87.
- 10 Calloway CD, Reynolds RL, Herrin GL, Anderson WW. The frequency of heteroplasmy in the HVII region of mtDNA differs across tissue types and increases with age. *Am J Hum Genet* 2000;66:1384-97.
- 11 Holland M. Heteroplasmy: interlaboratory study on heteroplasmy. In: Gillian Tully, editor. *Mitochondrial DNA Workshop. Proceedings of the First International Conference on Forensic Human Identification in the Millenium*; 1999 Oct 27-28; London, UK. Birmingham (UK): Forensic Science Service; 1999. p. 4.
- 12 Wakeley J. Substitution rate variation among sites in hypervariable region 1 of human mitochondrial DNA. *J Mol Evol* 1993;37:613-23.
- 13 Meyer S, Weiss G, von Haeseler A. Pattern of nucleotide substitution and rate heterogeneity in the hypervariable regions I and II of human mtDNA. *Genetics* 1999;152:1103-10.
- 14 Budowle B, Wilson MR, DiZinno JA, Stauffer C, Fasano MA, Holland MM, et al. Mitochondrial DNA regions HVI and HVII population data. *Forensic Sci Int* 1999; 103:23-35.
- 15 Melton T, Clifford S, Kayser M, Nasidze I, Batzer M, Stoneking M. Diversity and heterogeneity in mitochondrial DNA of North American populations. *J Forensic Sci* 2001;46:46-52.

Received: March 26, 2001

Accepted: April 19, 2001

Correspondence to:

Terry Melton
Mitotyping Technologies, LLC
1981 Pine Hall Drive
State College, PA 16801, USA
twm107@mitotyping.com