

FORENSICS AND MITOCHONDRIAL DNA: Applications, Debates, and Foundations*

Bruce Budowle,¹ Marc W. Allard,² Mark R. Wilson,³
and Ranajit Chakraborty⁴

¹Laboratory Division, FBI, Washington, DC 20535; email: bbudowle@fbi.gov

²Biology Department of Biological Science, George Washington University, Washington, DC 20052; email: mwallard@gwu.edu

³Laboratory Division, FBI Academy, Quantico, Virginia 22135;
email: mwilson@fbiacademy.edu

⁴Center for Genome Information, Department of Environmental Health, University of Cincinnati, Ohio 45267; email: ranajit.chakraborty@uc.edu

Key Words heteroplasmy, nomenclature, paternal inheritance, population data, recombination

■ **Abstract** Debate on the validity and reliability of scientific methods often arises in the courtroom. When the government (i.e., the prosecution) is the proponent of evidence, the defense is obliged to challenge its admissibility. Regardless, those who seek to use DNA typing methodologies to analyze forensic biological evidence have a responsibility to understand the technology and its applications so a proper foundation(s) for its use can be laid. Mitochondrial DNA (mtDNA), an extranuclear genome, has certain features that make it desirable for forensics, namely, high copy number, lack of recombination, and matrilineal inheritance. mtDNA typing has become routine in forensic biology and is used to analyze old bones, teeth, hair shafts, and other biological samples where nuclear DNA content is low. To evaluate results obtained by sequencing the two hypervariable regions of the control region of the human mtDNA genome, one must consider the genetically related issues of nomenclature, reference population databases, heteroplasmy, paternal leakage, recombination, and, of course, interpretation of results. We describe the approaches, the impact some issues may have on interpretation of mtDNA analyses, and some issues raised in the courtroom.

INTRODUCTION

No application-oriented field has embraced the tools of molecular biology more than forensic science. For more than 15 years, DNA typing analytical methods have been used worldwide to resolve identity issues in violent crimes, lesser crimes, acts

*The U.S. Government has the right to retain a nonexclusive, royalty-free license in and to any copyright covering this paper.

of terrorism, missing persons cases, and mass disasters. Methods include restriction fragment length polymorphism typing of variable number tandem repeat (VNTR) loci (22, 47, 67, 68, 123), polymerase chain reaction (PCR)-based systems to analyze single nucleotide polymorphisms (SNPs) (25, 33, 100), VNTR loci (23, 70), short tandem repeat (STR) loci (39), and direct sequencing of mitochondrial DNA (mtDNA) (60, 62, 108, 109, 120, 121).

While implementation and use of DNA typing technologies have been quite successful, the forensic field has a particular constraint not routinely encountered in other scientific disciplines, namely the law. In addition to science, the law plays a role in the review of the technology. While science and the law both are interested in the truth, they do not obtain the truth in the same manner. It is a tenet of science to continuously question one's beliefs and findings. Using the scientific method, hypotheses are proposed and experiments are carried out to test hypotheses. If the data do not refute a hypothesis, the hypothesis gains more support and through constructive incremental steps the hypothesis becomes grounded and accepted as reasonable and reliable. The law operates using an adversarial approach, at times attempting to create doubt and not always requiring any supporting data to reason such doubts. Typically, a defense attorney is placed in the position or has the responsibility to create doubt, even if the attorney is personally convinced of the client's guilt. However, at times, a prosecuting attorney may want to create doubt. In the courtroom, one may exploit the standard practice of science "to question" as lack of consensus, even if most, if not all, agree that the approach is reliable. However, the admissibility of scientific evidence is not always challenged. It is more likely to be challenged when the evidence does not support one's position.

Such challenges often occur in a pretrial admissibility hearing. The federal court and about half the states apply the factors listed in *Daubert v. Merrell Dow Pharmaceuticals* (36a), in which the judge is the gatekeeper. The admissibility hearing establishes that the reasoning and methodology underlying the technique or analysis conducted in the case is scientifically valid and will assist the jury in determining a fact(s) in issue. There are four Daubert factors the court may use to consider admissibility: (a) whether or not the technique can be or has been tested; (b) whether or not the method of analysis has been subjected to peer review and publication; (c) whether or not the technique has an acceptable rate of error, and whether or not there are standards that control the technique's operation; and (d) whether or not the method is generally accepted within the scientific community. The court's role is to attempt to screen out inappropriate and misleading (i.e., termed junk) science. Often the government is in court because the DNA forensic evidence supports the prosecution's hypothesis. Even if the result is objectively obtained, it will support one side's beliefs over that of the other. We submit that the scientific arena and testability have been and are better avenues than the courtroom for determining the validity of forensic DNA methods (although other scientific endeavors might be evaluated more expeditiously with a legal admissibility hearing).

DNA typing methodologies are continuously subjected to scientific and legal scrutiny. Most of these are methods that have been dedicated to typing nuclear

DNA. However, the eucaryotic cell contains two distinct genomes—the nuclear genome and the mitochondrial genome. Forensic mtDNA typing has been used successfully for many years (3, 17, 36, 48, 60, 77, 98, 101, 109, 120, 121) (<http://www.promega.com/ussymp9proc/default.htm>). Although the same analytical methods used in other fields are used in forensic analyses and the basic foundations of the science are the same as other established forensic DNA methods, questions about the admissibility of mtDNA evidence have arisen and are likely to continue to occur in the courtroom. This paper describes the use of mtDNA in forensic analyses, some genetic issues to be considered when establishing an effective methodology, and some examples of challenged topics, so that the reader can gain a better appreciation of science and forensics.

MITOCHONDRIAL DNA

Mitochondria are subcellular organelles that contain an extrachromosomal genome separate and distinct from the nuclear genome. The mtDNA is a histone-free, double-stranded circular molecule. It is a compact genome that encodes 13 polypeptides of approximately 80 protein subunits involved in oxidative phosphorylation, in addition to two ribosomal RNAs and 22 transfer RNAs. There is a noncoding region approximately 1,100 base pairs long, the control region. A mitochondrion contains between 2–10 copies of mtDNA, and there can be as many as 1000 mitochondria per somatic cell.

The mtDNA genome has been completely sequenced. One strand is purine-rich (termed the heavy strand) and one strand is pyrimidine-rich (termed the light strand). Nucleotide positions in the mtDNA genome are numbered according to the convention of Anderson et al. (7) with minor modification (6). An arbitrary position on the heavy strand begins the numerical designation of each base pair, continuing around the molecule for approximately 16,569 base pairs. The low fidelity of mtDNA polymerase and the apparent lack of mtDNA repair mechanisms have led to a higher mutation rate in the mitochondrial genome compared with the nuclear genome. Some regions of the mtDNA genome appear to evolve at rates 5–10 times higher than that of single-copy nuclear genes. These regions are of interest for human identity testing because of their hypervariability consequent of their higher mutation rate. Most of the sequence variation between individuals is found within two specific segments of the control region (50): the hypervariable region 1 (HV1) and hypervariable region 2 (HV2). There is an average of 8 nucleotide differences between Caucasian individuals and 15 differences between individuals of African descent within the HV1 and HV2 (26, 116). Typically, HV1 spans the region 16,024 to 16,365 and HV2 encompasses positions 73 to 340. The small size of each region allows for amplification by PCR and, hence, HV1 and HV2 are routinely typed for forensic testing purposes.

Unlike nuclear DNA, mtDNA is maternally inherited (29, 45, 64). Barring mutation, the mtDNA sequence of siblings and all maternal relatives is identical.

This characteristic can be helpful in forensic cases, such as analyzing the remains of a missing person, where known maternal relatives can provide reference samples for direct comparison to the questioned mtDNA type (48, 60). Because of a lack of recombination, maternal relatives several generations apart from the source of the evidence (or biological material) can serve as reference samples. Nuclear DNA markers, other than those on the nonrecombinant portion of the Y chromosome, cannot provide this feature.

The haploid and monoclonal nature of mtDNA simplifies interpretation of DNA sequencing results (84, 85). Though most individuals are operationally homoplasmic, heteroplasmy at occasional sites may be encountered (13, 14, 32, 46, 66, 122). A person is considered heteroplasmic if he/she carries more than one detectable mtDNA type. Careful analysis and direct comparisons between multiple known samples and a questioned sample should, in most cases, alleviate interpretational difficulties that may arise due to the presence of heteroplasmy (4, 11, 21, 24, 28, 61, 112) (<http://www.promega.com/ussymp10proc/default.htm>).

Because of the high copy number of mtDNA molecules in a cell (16), typing mtDNA is particularly advantageous, compared with nuclear DNA, for certain kinds of forensic analyses. In cases where the amount of extracted DNA is very small or degraded, it is more likely that a DNA typing result can be obtained by typing mtDNA than by typing polymorphic markers found in nuclear DNA.

The utility, application, and validity of forensic mtDNA are well documented (3, 17, 36, 46, 57, 60, 62, 66, 76, 77, 98, 101, 105, 108, 109, 120, 121) (<http://www.promega.com/ussymp9proc/default.htm>). Besides bones and teeth as sources for analysis, hair shafts contain mtDNA (3, 4, 98, 101, 115, 121). Results have been obtained routinely from as little as 1–2 cm of a single hair shaft. mtDNA sequencing is also the primary analytical tool used to identify individuals from bones recovered from American war casualties (60). Victim remains from the World Trade Center tragedy of the September 11, 2001 act of terrorism are being characterized using mtDNA sequencing technology. In addition, mtDNA analysis has been used in evolutionary studies (5, 18, 27, 31, 42–44, 54, 58, 63, 72, 79, 97, 99, 111, 116) and anthropological studies (57, 78, 89–92, 94, 96, 117), including analyses of 7000-year-old brain tissue (93) and the remains of a Neanderthal man (74). One of the most well-known lineage studies using mtDNA sequencing resulted in verification of Tsar Nicholas II's bones. Using mtDNA obtained from living maternal relatives (Countess Xenia Cheremeteff-Sfiri and the Duke of Fife), a comparison was made with the sequence of mtDNA extracted from the putative bones of the Tsar. The sequences were similar and the data supported the hypothesis that the putative remains were those of Tsar Nicholas II (46, 66).

There is no doubt that mtDNA is generally accepted as a tool for forensic identity testing and evolutionary studies. However, one needs to appreciate mtDNA's various features, and its application, in order to properly interpret results from evidence. These include issues such as nomenclature, reference population databases, heteroplasmy, paternal leakage, and recombination.

NOMENCLATURE

The naming of DNA sequences may seem obvious, simplistic, and trivial. However, complications arise if proper consideration is not afforded to nomenclature. Listing more than 600 bases to describe results from HV1 and HV2 would be cumbersome and unwieldy. Thus, an alternate approach was developed that identifies only differences from a reference sequence. Anderson et al. (7) described the first entire human mtDNA sequence. The published sequence is used as a reference standard and is termed the Cambridge Reference Sequence (CRS) (6). The sequence is displayed as the light strand sequence and is compared directly with a sequence from the sample(s). Only differences between the aligned sample(s) and the CRS are noted; all other positions in the sequence are understood to be the same as the CRS. For example, position number 16311 in the CRS is listed as a thymidine, or T. However, in some individuals, that position could be substituted with a cytosine, or C. This polymorphism is designated as 16311 C, and all other positions in the region sequenced are understood to be identical to the CRS. Insertions are designated by placing a period after the last aligned base in the Anderson sequence and listing the insertion with the appropriate nucleotide. For example, if bases beyond position 16192 were out of register by one base due to the insertion of a C (i.e., 16193 onward are aligned with the reference sequence) the inserted base is designated as 16192.1 C. If two Cs were inserted instead of one, they are designated as 16192.1 C, 16192.2 C. Deletions are recorded as the number of the base or bases missing with respect to the CRS (for example, 249 D or 249-). Bases that cannot be unambiguously determined are coded using an N.

Naming mtDNA sequences by referring to a standard sequence provides a common language and an easy tool for describing the variation observed in human populations. The above general nomenclature guidelines for recording sequence differences have been described elsewhere (11, 24, 28, 112) (<http://www.promega.com/ussymp10proc/default.htm>) and are used throughout the forensic community. However, there are some situations where different alignments may be proffered. If these are not standardized, identical sequences may be (and have been) aligned and listed differently and inconsistently [see (118)]. Thus, when determining the number of times the evidentiary sequence is observed in a reference population database, the count may be underestimated. One should be cautious not to infer solely phylogenetic interpretations to describe alignments because such a treatment may use information from other samples in a data set. In forensics, the evidentiary sample's ethnohistory is unknown, and this is exacerbated with mtDNA, which may not unequivocally reflect the lineage of nuclear genes. To be fair to a defendant, one should attempt to list any sequences in a reference database that could result in the same alignment as the same sequence when estimating how common or rare an evidentiary mtDNA sequence is.

Most ambiguities in alignment/nomenclature arise with insertion and deletion placement. Wilson et al. (118, 119) (<http://www.fbi.gov/hq/lab/fsc/backissu/oct2002/index.htm>) developed an approach to attempt to standardize alignments.

It is based on a phylogenetic context using differential weighting of transitions, transversions, insertions, and deletions. However, the alignment and nomenclature may not reflect the true biological mechanism that generated the biological variation. Instead, the nomenclature approach has the goal of database stability. The following recommendations address most scenarios: *a*) characterize the variant(s) using the least number of differences (i.e., substitutions, insertions, deletions) from the reference sequence; *b*) if there is more than one possible alignment, each having the same number of differences with respect to the reference sequence, a prioritization is made first for insertions/deletions (indels), then transitions, and lastly transversions; *c*) insertions and deletions should be placed 3' with respect to the light strand (when possible, insertions and/or deletions should be combined); and *d*) gaps are combined together only if they can all be placed in the most 3' position while maintaining the same number of differences from the reference. Bases designated with an N should not affect the recommendations. The recommendations listed above are hierarchical. Recommendation *a* takes precedence, followed by recommendations *b* and *c*, which to some degree are arbitrary. Recommendation *d* is a clarification to further facilitate alignment interpretations. Several examples have been identified where alternative alignment strategies are possible and have been described elsewhere (118, 119) (<http://www.fbi.gov/hq/lab/fsc/backissu/oct2002/index.htm>). The Appendix (below) provides examples for conceptualizing the recommended approach.

Example 2 in the Appendix (below) shows three possible alignments. Suppose that three analysts in a laboratory are generating population data and each one names Example 2 type differently. Now consider hypothetically that there are 9 samples in a set of 100 individuals that have the same type as in Example 2 and the other 91 samples have unique types in the data set. If each analyst interpreted 3 of the 9 samples, no more than three individuals (3%) would share the same type in the data set. Yet with a standard nomenclature, nine individuals (9%) would share the same type. In a forensic case, the weight of evidence is primarily based on the number of times a profile is observed in a reference data set; thus, one can appreciate the potential underestimation that could occur if a standard nomenclature is not applied. The hypothetical illustrates a nomenclature issue that might occur within a laboratory, and this would be exacerbated across laboratories. The recommendations can help clarify inconsistencies in nomenclature, provide higher quality databases, and enable better, more reliable profile searches for evaluating the weight of evidence. As more samples are typed, there may be scenarios that require refinement or additional rules. It is important for the forensic community to adopt a standard that resolves ambiguities that may reduce the weight of the evidence. By promulgating such recommendations, the issue of a consistent nomenclature is brought to the forefront so that eventually we may adopt a community-wide standard. Perhaps haplogroup nomenclature recommendations also should be promulgated to avoid possible misinterpretations.

REFERENCE POPULATION DATABASES

When a mtDNA sequence from an evidence sample and one from a known reference sample cannot be excluded as originating from the same source, it is desirable to convey some information about the mtDNA profile's rarity. The current practice is to count the number of times a particular sequence is observed in a population database(s) (24, 26, 61) (<http://www.promega.com/ussymp10proc/default.htm>). Because of the uncertainty involved in all population database samplings, a confidence interval can be placed on the observation (24) (<http://www.promega.com/ussymp10proc/default.htm>). Only the upper bound estimate is provided to the fact finder.

Typically, the population databases used in forensics are comprised of convenience samples (88). The databases tend to represent the general major population groups of the potential contributors of evidence. The relevance and representativeness of these databases should be considered for forensic applications. Demonstrating this is more demanding than with nuclear markers. Because of the lack of recombination, larger mtDNA data sets are needed to assess the degree of variation compared with nuclear markers. Pair-wise comparison of haplotypes and genetic diversity have been used to assess the relevance and representativeness of these databases (26, 81–83, 106). Though such data are of some value, a better approach is to use phylogenetic assessments (1, 2).

Determining the relationships among human mtDNA haplotypes and haplogroups provides more detailed information about the structure of human genetic variation. A haplotype is a specific sequence observed in the population sample; a haplogroup refers to a cluster of haplotypes that share common variable characters that define them as having a shared ancestry. Thus, there are larger samplings within a database of haplogroups than haplotypes. Phylogenetic methods have been used to identify the major human haplogroups and the important variable SNPs that define these groups. Based on the nomenclature alignment recommendations described above, phylogenetic analyses were performed using the Scientific Working Group on DNA Analysis Methods (SWGDM) mtDNA population data sets (86) (<http://www.fbi.gov/hq/lab/fsc/backissu/april2002/index.htm>). There are several alternate methods for constructing a phylogenetic tree of the mtDNA sequences, including neighbor-joining, maximum likelihood, minimum spanning networks, and maximum parsimony. These phylogenetic methods are well characterized by the systematic community (56, 59, 95). The maximum parsimony method was used because of the relative ease of interpretation of character changes on the branches that lead to defining and characterizing major human lineages. This phylogenetic method has also been used extensively to examine and characterize human mtDNA control region variation, and thus a large body of data is available for comparison (1, 2, 42, 58, 72, 97). In addition, the assumption of shared variation due to shared ancestry is preferred over that due to homoplasy (i.e., independent gains or secondary losses) unless there is alternate evidence to suggest otherwise. Allard et al. (1, 2) performed parsimony analyses of human mtDNA variation in the SWGDAM

data set utilizing the programs Winclada and Nona (49). If the same major lineages supported by the same relevant SNPs are observed in other reference data sets, additional support can be offered for the representativeness and relevance of the SWGDAM data set.

Previous analyses of European Caucasian genetic variation outlined the major branches of the tree of West Eurasian mtDNA haplogroups, as well as the SNPs that define these haplogroups (42, 58, 79). Comparing the SWGDAM European Caucasian data sets with other European analyses shows a consistent phylogenetic structure and defining variants (1). Approximately 99% of the known European and U.S. Caucasian mtDNA variation can be categorized into 10 major haplogroup lineages: H, I, J, K, M, T, U, V, W, and X (42, 58, 111). Haplogroup H is the most commonly observed mtDNA type, occurring at a frequency of 45.7% of the SWGDAM data. The next most common haplogroups in the data set are U (15.6%), T (10.5%), J (10.0%), and K (8.9%), and these values are similar to other published analyses (42, 58, 111). For the Caucasian SWGDAM data set, 229 parsimony informative SNPs were observed, of which 72 SNPs defined clades of 10 or more individuals. After removing some of the redundant sites from closely associated SNPs, a minimum set of 32 SNPs define the 10 haplogroups (Figure 1). The SWGDAM reduced list of SNPs identifies all clusters in the data set with 10 or more individuals and largely overlaps that of previous descriptions of the important SNPs for European Caucasians.

When Allard et al. (2) analyzed the SWGDAM African American mtDNA data, there were not any published analyses of African American mtDNA control region genetic variation, but data from African populations were available (5, 15, 31, 97, 99). The phylogenetic analyses show that the SWGDAM African American data set contains variation consistent with that described in other African populations. Sixteen of the 18 haplogroups previously observed in African populations were observed in the SWGDAM reference data set and included L1a, L1b, L1c, L1e, L2a, L2b, L2c, L3*, L3b, L3d, L3e1*, L3e1a, L3e2a, L3e2b, L3e3, and L3e4. Haplogroup L2a is the most commonly observed haplogroup (19%) in the African American data set. The next most common haplogroups are L1c (11%), L1b (9%), and L3b (8%). Approximately 12% of the haplogroups observed within African Americans were common in European Caucasians or Asians; these were H, T, J, K, U6, W/X, V, A, C, and M, respectively. There were 217 parsimony informative SNPs observed in two or more individuals within the African American SWGDAM data set, and 71 SNPs defined clades of 10 or more individuals. A minimum set of 34 SNPs could partition most of the African American haplogroups (Figure 1). The most important SNPs for defining major haplogroups are consistent with previous studies (97).

These data lend additional support that the forensic data sets are useful for applying weight to an observed mtDNA match and are consistent with other assessments of population variation. Additional analyses are underway to assess variation in far East Asians, and the preliminary results suggest that the structure of variation in the SWGDAM data sets is consistent with the variation in other reference population studies.

CRS site	Af	Am	Cauc	CRS site	Af	Am	Cauc	CRS site	Af	Am	Cauc	CRS site	Af	Am	Cauc
16048	G	A		16215	A	G		16327	C	T	G	189	A	C/G	G
16051	A		G	16222	C		T	16343	A		G	194	C		T
16069	C		T	16223	C	T	T	16355	C	T		195	T		C
16086	T	C		16224	T		C	16356	T		C	198	C	T	
16093	T		C	16230	A	G		16360	C	T		199	T		C
16114	C	A/G/T		16231	T		C	16362	T	C	C	200	A	G	
16124	T	C		16256	C	T	T	16390	G	A	A	204	T	C	C
16126	T	C	C	16261	C	T	T	16399	A	G	G	207	G	A	A
16129	G	A	A/C	16263	T		C	16519	T	C	C	215	A		G
16145	G		A	16264	C	T		16527	C			217	T		C
16148	C	T		16265	A	C/G/T		64	C	T		225	G		A
16153	G		A	16270	C	T	T	72	T		C	226	T		C
16162	A		G	16271	T		C	73	A	G	G	228	G		A
16163	A		G	16278	C	T	T	93	A	G	G	236	T	C	
16168	C	T		16286	C	A/G/T		95	A	G		239	T		C
16172	T	C	C	16291	C		T	119	T		C	242	C		T
16183	A	C/G	C	16292	C	T	T	143	G	A		247	G	A	
16185	C	T		16293	A	G/T	G	146	T	C	C	250	T		C
16186	C		T	16294	C	T	T	150	C	T	T	263	A		G
16187	C	T		16295	C	T		151	C	T	T	295	C		T
16188	C	G		16296	C		T	152	T	C	C	297	A	G	
16189	T	C	T	16298	T		C	153	A		G	316	G	A	
16192	C	T	C	16304	T		C	182	C	T		325	C		C
16193	C		T	16309	A	G		183	A	G		357	A	T	G
16207	A			16311	T	C	C	185	G	A/T	A	418	C	T	
16209	T	C	C	16319	G	A		186	C	A		477	T		C
16213	G	A		16320	C	T	T	188	A		G	489	T	C	C
												497	C		T
												499	G	A	

Figure 1 SNPs determined by phylogenetic analysis of the African American and Caucasian mitochondrial DNA (mtDNA) control region sequences in the Scientific Working Group on DNA Analysis Methods (SWGDM) database. The numbers refer to the revised Cambridge Reference Sequence (CRS) nomenclature system for mtDNA sites. Light gray and black shading designate single nucleotide polymorphisms (SNPs) that defined groups with 10 or more individuals in the respective data set. Black shading indicates the most informative SNPs based on phylogenetic analysis and examination of the evolution of character data on a tree. This partly entailed removal of some of the redundant characters that define groups. More than one character listed in a shaded area indicates observed multiple states at a site and/or the presence of reversals.

HETEROPLASMY

Heteroplasmy is operationally defined as detecting more than one mtDNA type within an individual. (There are two classes of heteroplasmy: length and point substitution; we concentrate on heteroplasmy at specific points). Heteroplasmy manifests in different ways: *a*) an individual may show more than one mtDNA type in a single tissue, *b*) an individual may be heteroplasmic in one tissue sample and homoplasmic in another tissue sample, and *c*) an individual may exhibit one mtDNA type in one tissue and a different type in another tissue (24, 105) (<http://www.promega.com/ussymp10proc/default.htm>). Of the three possible heteroplasmic scenarios, the last one is least likely to occur. In the majority of forensic analyses involving HV1 and HV2 of the control region of the mtDNA genome, heteroplasmy is not observed. However, when heteroplasmy is found the mtDNA

species in an individual usually differ at a single base (based on HV1 and HV2 typing). Heteroplasmy at two, and even possibly three, sites also may occur, but at much lower rates.

A decade ago most individuals were thought to be homoplasmic (19, 84, 85). However, with better technology to detect minor components in a sample [i.e., see (114)], and more samples being typed, researchers have observed many examples of heteroplasmy in humans. For example, heteroplasmy was observed at position 16169 of the mtDNA control region in the putative remains of Tsar Nicholas II of Russia (46, 66). The mtDNA sequence analysis of the remains of the Tsar's brother, Grand Duke of Russia Georgij Romanov, also demonstrated heteroplasmy at position 16169 (66). Comas et al. (32) detected heteroplasmy at two positions (16293 and 16311) in the mtDNA from an anonymous donor's plucked hair. Wilson et al. (122) observed a family that carried a mtDNA heteroplasmic state at position 16355. In this family (a mother and two children), blood and buccal swab samples demonstrated the heteroplasmy, and in some cases individual hairs carried either a C or a T at position 16355. Thus, some hairs from a single individual appeared homoplasmic, and some hairs differed from one another at only one nucleotide position. The fact that heteroplasmy occurs more often than originally observed and the mechanism and rate of heteroplasmy are not well defined are often raised in admissibility challenges in an attempt to exclude mtDNA evidence. But with careful evaluation, one can avoid erroneous interpretations.

In forensic analysis, the mtDNA types between a known exemplar(s) and an evidence sample(s) are compared, and established interpretation guidelines are used to assist in the evaluation. If the mtDNA sequences are the same at all defined sites, the interpretation is a failure to exclude the samples as possibly having the same origin (or under some circumstances the same maternal lineage). When heteroplasmy is observed, guidelines are in place to effect an interpretation. If the compared samples have mtDNA sequences that are heteroplasmic at the same nucleotide sites, the interpretation is a failure to exclude. If one sample is heteroplasmic and the other homoplasmic, and they share the same bases at the defined sites, the interpretation also is a failure to exclude the samples as possibly having the same origin because these samples share at least one mtDNA species. If both samples yield a homoplasmic mtDNA profile, yet differ typically, for example, at only one site, additional investigation is warranted. If possible, additional reference samples are obtained and processed. If the evidence and reference samples differ by a single nucleotide and there is no evidence of heteroplasmy, the comparison is reported as inconclusive (i.e., there is insufficient information to render an interpretation of inclusion or exclusion). Some have suggested that one-base differences may be further evaluated based on the rate of mutation or evolution (4, 28, 61, 112). This is a reasonable proposition; however, more data may be needed to apply proper weight to the interpretation. In cases where two or more nucleotide differences exist between the two sequences, generally an exclusion is the rendered interpretation, although rarely a false exclusion will occur using this criterion. Heteroplasmy is most often observed in hair samples because genetic drift and bottlenecks are

created due to a hair follicle's semiclonal nature. If an evidentiary hair sample contains one of the two heteroplasmic lineages observed in a reference sample (or vice versa), then an interpretation of an exclusion is incorrect. Lastly, hairs are the most likely tissue to express homoplasmic profiles that at times may differ by one nucleotide within an individual. When the mtDNA sequence from a hair evidence sample differs from a known reference at one nucleotide position, typing additional known hairs may resolve the issue.

The suggestion that the presence of heteroplasmy renders forensic mtDNA analysis invalid may seem heavy-handed, yet it is routinely raised in admissibility hearings. The following is an example of how the existence of heteroplasmy has been used to challenge the use of mtDNA evidence in the courtroom. Based on current knowledge, when heteroplasmy is observed within the HV1 and HV2, sequences typically show evidence of a mixture or differ in sequence at one and, to a much lesser extent, two bases. However, Grzybowski (52, 53) reported much higher levels of mtDNA heteroplasmy in hair samples. Grzybowski sequenced just HV1 in 100 forcibly removed hairs from 35 Polish individuals. Thirteen of the 35 people displayed heteroplasmy, and more than half of the heteroplasmic individuals carried multiple heteroplasmic sites. One individual was reported to have six heteroplasmic positions. This level of heteroplasmy was substantially higher than was previously observed. Though inconsistent with typical observations, Grzybowski's findings (52) were used to challenge the admissibility of mtDNA analysis of forensic samples (21, 37, 38), though to no avail. A single study may be an outlier and/or may not be sufficient evidence to change established tenets. Though such a study might be readily dismissed in the scientific arena, this sort of finding can persist in the courtroom. Regardless, oddities should not be dismissed in forensic DNA analysis without review and due consideration.

Grzybowski (52) used at least three orders of magnitude more template DNA in the PCR, as well as approximately twice the number of cycles of PCR, than routinely used in forensic laboratories. Such analytical conditions are likely to succumb to contamination problems, and Tully & Lareu (113) suggested this could be the basis for Grzybowski's findings. Budowle et al. (20) concurred and carried out a more in-depth review of the Grzybowski data. Since the same hair samples were not available for reanalysis, we analyzed the natural human mtDNA HV1 variation within a reference sample population (1). Using the phylogenetic analyses described above, we compared these analyses with that observed in the Grzybowski (52) study.

First, we found the primary sequence data displayed in the tables and figures of Grzybowski (52) inconsistent with the reported sites. The sequences were misidentified and/or outside the mtDNA regions that were initially amplified. Second, only two of Grzybowski's (52) four defined hot spots (i.e., rapidly changing sites) match the most variable sites (16126 and 16311) known for U.S. and European Caucasians (1). Using the SWGDAM mtDNA population data set (86) (<http://www.fbi.gov/hq/lab/fsc/backissu/april2002/index.htm>), many other sites (16093, 16129, 16172, 16183, 16189, 16192, 16261, 16362) change more rapidly

in HV1 (1) than sites 16294 and 16296, which Grzybowski (52) described as hot spots. In fact, site 16296 has a relatively slow rate of change (1). Sites that change rapidly in the mtDNA HV1 and are not listed as hot spots by Grzybowski (52) include 16129, 16189, 16192, and 16362. Third, Grzybowski (52) reported other heteroplasmic sites that are uncommon in the Caucasian reference data set. For instance, sites 16088, 16105, and 16125 showed no variation when compared to the 1771 Caucasians in the SWGDAM data set (1). Many of the heteroplasmic sites listed by Grzybowski (52) are rare variants in the SWGDAM data set. These include 16024 ($n = 6$), 16090 ($n = 2$), 16140 ($n = 4$), 16167 ($n = 6$), 16243 ($n = 7$), 16288 ($n = 5$), and 16316 ($n = 11$).

These observations suggest that Grzybowski's (52) results are unusual and would be an unlikely combination of variable sites. Review of the data supports that fundamental flaws in the primary data and contamination may explain some, but not all, of the results. The observation of rare types in multiple individuals cannot be explained solely by contamination. It is possible that Poles are genetically different (at mtDNA), although European data do not support this contention. Alternatively, the extreme analytical PCR conditions may enable detection of nuclear pseudogenes. Such genes are found within the nuclear genome and are abundant (12, 87). More studies are needed to determine if some of the odd results are due to amplification of nuclear pseudogenes. Regardless, Grzybowski's analytical methods are not typically recommended for forensic casework. Thus, his results do not apply to other methods of mtDNA analysis and cannot be used to question the reliability of current practices in the forensic arena. One might ask how such a study could be considered a basis for questioning the reliability of mtDNA forensic analysis in light of the fact that the scientific literature is replete with data demonstrating its reliability. As stated earlier, in the legal arena any study that may raise doubt may be considered [see (20, 37, 38)].

Instead of wholly rejecting a method, it might be better for one to consider what the consequence would be if higher levels of heteroplasmy were found to exist. Typically, when more than one site is observed to be heteroplasmic in an individual sample(s), most of the samples show both types at the sites. Only a small portion of samples would be problematic and these would be when two samples present themselves as homoplasmic and differ by two sites. With the current interpretational guidelines, the interpretation would be an exclusion. The forensic community accepts such an interpretation because it would at most result in excluding an accused individual as the source of the sample when he/she is the true source.

PATERNAL LEAKAGE AND RECOMBINATION

The mitochondrial genome is inherited maternally. Although sperm contain a few mitochondria in the neck and tail regions, the male mitochondrial genome is destroyed during or shortly after fertilization. Sperm mitochondria disappear in early

embryogenesis by selective destruction, inactivation, or dilution (34, 35, 103, 104). However, there are a few reported examples of paternal inheritance of the mitochondrial genome in animals, but these do not represent the norm. For example, Gyllensten et al. (51) reported paternal inheritance of mtDNA in lab mice to occur at a rate about 1×10^{-5} to 5×10^{-5} per generation. In humans, the evidence is scant for paternal inheritance. Although rare, the possibility that paternal inheritance may occur has been proffered in the courtroom as a basis for not admitting mtDNA evidence. The argument is that because paternal mitochondrial DNA leakage may occur, and though it may be rare, one does not know if it occurs and if so how often; therefore mtDNA analysis is not ready for forensic applications.

The initial evidence for the possibility of human paternal mtDNA inheritance was reported by Hagelberg et al. (54). They observed in an isolated population from the island of Nguna, in the archipelago Vanuatu in Melanesia, a rare mutation in three unrelated human mtDNA lineages. The variant was a substitution (a C to T) at position 16076, which had been observed only once before in a European sample population. The occurrence of such a rare mutation in three distinct haplogroups in one geographic location is highly unlikely and could only be explained by paternal leakage and recombination. Thus, the Hagelberg et al. (54) study became the basis of a debate in the courtroom and also in the scientific arena to support the notion that recombination may occur within the mitochondrial genome. Although Hagelberg et al. (55) retracted their findings, reporting that the observation was due to an alignment error, the debate continues. Schwartz & Vissing (102) recently identified a 28-year-old man with mitochondrial myopathy, where the causative mtDNA was paternal in origin and accounted for 90% of his muscle's mtDNA. The patient's blood, hair, and cultured fibroblasts carried maternal origin mtDNA.

Fueled by the initial report by Hagelberg et al. (54), Awadallah et al. (10) and Eyre-Walker et al. (40) statistically analyzed mtDNA sequence data for evidence of recombination. They observed a large number of homoplasies and reasoned that these were due either to repeated mutations or recombination between lineages. Since they observed disruption in linkage disequilibrium correlated with distance, recombination seemed a more plausible explanation. Awadallah et al. (10, 40) suggested that the large number of homoplasies most likely are the result of recombination between maternal and paternal lineages or between nuclear pseudogenes and the mitochondrial genome. However, there is no demonstrated case of recombination. But there is evidence of homologous recombination activity within mammalian mitochondrial extracts (possibly involved in mitochondrial DNA repair) (110), and it has been reported that paternal mitochondria can enter the egg (34, 35, 103, 104). In the case of paternal and maternal recombination, recombination would require physical contact between egg and sperm mtDNA, which currently is difficult to envision.

There have been several criticisms of the Awadallah et al. (10, 40) analysis (8, 65, 69, 73, 75, 80). The critiques suggest that Eyre-Walker et al. (a) analyzed

data with sequencing errors and scoring inconsistencies, which is another reason to establish nomenclature rules as discussed earlier; *b*) constructed their phylogenetic tree assuming that hypothetical ancestral types at nodes of the tree were extinct, which is probably an invalid assumption that resulted in an increased number of observed homoplasies (i.e., an artifact of analysis); *c*) used the measure r_2 , and this created a bias in the analysis (i.e., an artifact of analysis); *d*) rejected that mutational hot spots exist; and *e*) chose sites that by chance gave rise to a negative correlation. Awadallah et al. (10, 41) rebutted these arguments by analyzing better quality data sets and by asserting that their statistical approach was appropriate. We suggest that direct and indirect signatures of recurrent mutations are a more plausible explanation than recombination (30, 71).

The debate on paternal leakage and recombination will likely continue. Family studies have yet to document recombination (75), so recombination would seem unlikely or at best rare. However, disproving the negative is impossible. Furthermore, and rightly so, Eyre-Walker et al. (41) point out that paternal inheritance is rarely tested in family studies, and SNPs have not been considered as possibly being due to recombination. Yet, maternal inheritance patterns do not seem discrepant, and for the HV1 and HV2 on average Caucasians differ at eight sites and Africans differ at 15 sites. Regardless, even if paternal leakage and recombination were to occur, it would seem to be infrequent; so it may be better to ask what the implications are, particularly for forensics.

Perhaps paternal leakage may make a small contribution to mtDNA variation. Therefore, if recombination occurs, and it would have to be more than rarely, it may affect the evolutionary clock estimates and times of divergence may be greater for mtDNA. It would not impact evolutionary estimates based on other markers. The forensic impact would be negligible, if any at all. First, forensic estimates of mtDNA profile rarity do not rely on evolutionary estimates; relevant reference population data sets are from current population groups. Second, most comparisons are made from evidence to putative direct sources. Thus, the mechanism that generated an individual's genetic constitution does not need to be understood to effect a comparison of evidence with an exemplar. However, when believed maternal relatives are used as a reference, one must always consider mutations. If paternal inheritance and/or recombination had occurred, the result, if not considered, could be a false exclusion. This will likely be a rare circumstance. In the case of identifying remains, mtDNA data, alone, rarely are sufficient to render an absolute identification. Other data, such as anthropological, geo-location, markings, time of death, etc., are considered. In those unlikely situations where all data, except mtDNA, support an identification, it may be worth considering typing an informative paternal lineage sample although we would not recommend typing paternal sources routinely. It is difficult to envision that paternal inheritance or recombination would result in a false positive. It does not matter how the genetic constitution of an individual is generated; it matters whether or not he/she carries a mtDNA type in common with the evidence. Lastly, if

recombination occurs, the current forensic practice of determining the weight of evidence (by counting the number of observed sequences in a data set) will be more conservative than believed. We suggest that current practices are adequate and reliable. Indeed, Eyre-Walker's position on forensics is consistent with ours—"if the sequences are identical the chances are good that's the woman's son or daughter" and "If you get a one-base-pair mismatch, do you say 'This is not your child?'" (107).

SUMMARY/CONCLUSION

Although the scientific community has spent more than ten years developing, validating, and laying the foundation for forensic use of mtDNA analysis, debate on the validity and reliability of scientific methods often arises in the courtroom. When the government is the proponent of the evidence, it is the defense's responsibility to vigorously challenge such evidence. In some instances such challenges are warranted; sometimes they are frivolous. Regardless, one must not dismiss but instead recognize and properly characterize issues raised in the courtroom. Moreover, prior to implementation of mtDNA sequencing of forensic biological evidence, foundations should be laid that address genetically related issues, e.g., nomenclature, reference population databases, heteroplasmy, paternal leakage, recombination, and interpretation of results. Challenges to mtDNA analysis often focus on the genetic issue of heteroplasmy, and to a lesser degree paternal leakage and recombination. However, the premise that "unless one has the sum total knowledge of everything, one should not proceed" is not acceptable or a tenet in any scientific endeavor. These challenges on mtDNA in the courtroom have been overwhelmingly unsuccessful. The techniques are well grounded and reviewed by the scientific community. Moreover, the scientific community is well aware of those factors that can impact the interpretation of results and does not find them to affect the reliability of mtDNA profiling. Forensic scientists take appropriate steps and evaluate new issues as they arise for their potential impact. We have described some of the issues raised in the courtroom and the impact some of those issues may have on interpretation. Space limitations do not permit us to go into greater detail on some genetic issues that are important to the application of forensic mtDNA analysis. Hopefully the reader gains an appreciation of some of the genetic aspects to consider for use of a forensic application such as mtDNA analysis.

ACKNOWLEDGMENT

This is publication number 03-01 of the Laboratory Division of the Federal Bureau of Investigation. Names of commercial manufacturers are provided for identification only, and inclusion does not imply endorsement by the Federal Bureau of Investigation.

The Annual Review of Genomics and Human Genetics is online at
<http://genom.annualreviews.org>

LITERATURE CITED

1. Allard MW, Miller K, Wilson MR, Monson KL, Budowle B. 2002. Characterization of the Caucasian haplogroups present in the SWGDAM forensic mtDNA data set for 1771 human control region sequences. *J. Forensic Sci.* 47:1215–23
2. Allard MW, Miller K, Wilson MR, Monson KL, Budowle B. 2003. Characterization of 1257 human control region sequences for the African American and African haplogroups in the SWGDAM forensic mtDNA data set. *Forensic Sci. Int.* Submitted
3. Allen M, Engstrom AS, Myers S, Handt O, Saldeen T, et al. 1998. Mitochondrial DNA sequencing of shed hairs and saliva on robbery caps: sensitivity and matching probabilities. *J. Forensic Sci.* 43:453–64
4. Alonso A, Salas A, Albarran C, Arroyo E, Castro A, et al. 2002. Results of the 1999–2000 collaborative exercise and proficiency testing program on mitochondrial DNA of the GEP-ISFG: an interlaboratory study of the observed variability in the heteroplasmy level of hair from the same donor. *Forensic Sci. Int.* 125:1–7
5. Alves-Silva J, Santos M, Guimaraes P, Ferrekrá A, Bandelt H, et al. 2000. The ancestry of Brazilian mtDNA lineages. *Am. J. Hum. Genet.* 67:444–61
6. Andrews R, Kubacka I, Chinnery P, Lightowlers R, Turnbull D, Howell N. 1999. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.* 23:147
7. Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, et al. 1981. Sequence and organization of the human mitochondrial genome. *Nature* 290:457–65
8. Arctander P. 1999. Mitochondrial recombination? *Science* 284(5423):2090–91
9. Awadalla P, Eyre-Walker A, Maynard Smith J. 1999. Linkage disequilibrium and recombination in hominid mitochondrial DNA. *Science* 286:2524–25
10. Awadalla P, Eyre-Walker A, Maynard Smith J. 2000. Questioning evidence for recombination in human mitochondrial DNA. *Science* 288:1931a
11. Bar W, Brinkmann B, Budowle B, Carracedo A, Gill P, et al. 2000. DNA Commission of the International Society for Forensic Genetics: guidelines for mitochondrial DNA typing. *Int. J. Leg. Med.* 113:193–96
12. Bensasson D, De-Xing Z, Hartl DL, Hewitt GM. 2001. Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends Ecol. Evol.* 16:314–21
13. Bendall KE, Macaulay VA, Sykes BC. 1997. Variable levels of a heteroplasmic point mutation in individual hair roots. *Am. J. Hum. Genet.* 61:1303–8
14. Bendall KE, Sykes BC. 1995. Length heteroplasmy in the first hypervariable segment of the human mtDNA control region. *Am. J. Hum. Genet.* 57:248–56
15. Bendelt H, Alves-Silva J, Guimaraes P, Santos M, Brehm A, et al. 2001. Phylogeography of the human mitochondrial haplogroups L3e: a snapshot of African prehistory and Atlantic slave trade. *Ann. Hum. Genet.* 65:549–63
16. Bogenhagen D, Clayton DA. 1974. The number of mitochondrial deoxyribonucleic acid genomes in mouse L and human HeLa cells. *J. Biol. Chem.* 249:7791–95
17. Boles TC, Snow CC, Stover E. 1995. Forensic DNA typing on skeletal remains from mass graves: a pilot project in Guatemala. *J. Forensic Sci.* 40:349–55
18. Brown WM, Prager EM, Wang A, Wilson AC. 1982. Mitochondrial DNA sequences

- of primates: tempo and mode of evolution. *J. Mol. Evol.* 18:225–39
19. Budowle B, Adams DE, Comey CC, Merrill CR. 1990. Mitochondrial DNA: a possible genetic material suitable for forensic analysis. In *Advances in Forensic Science*, ed. H Lee, RE Gaensslen, 76–97. Chicago, IL: Year Book Med. Publ. 278 pp.
 20. Budowle B, Allard MW, Wilson MR. Critique of interpretation of high levels of heteroplasmy in the human mitochondrial DNA hypervariable region I from hair. *Forensic Sci. Int.* 126:30–33
 21. Budowle B, Allard MW, Wilson MR. 2002. Characterization of heteroplasmy and hypervariable sites in HV1: critique of D'Eustachio's interpretations. *Forensic Sci. Int.* 130:68–70
 22. Budowle B, Baechtel FS. 1990. Modifications to improve the effectiveness of restriction fragment length polymorphism typing. *Appl. Theor. Electrophoresis* 1:181–87
 23. Budowle B, Chakraborty R, Giusti AM, Eisenberg AJ, Allen RC. 1991. Analysis of the VNTR locus D1S80 by the PCR followed by high-resolution PAGE. *Am. J. Hum. Genet.* 48:137–44
 24. Budowle B, DiZinno JA, Wilson MR. 1999. *Interpretation guidelines for mitochondrial DNA sequencing*. Presented at Int. Symp. Hum. Identif., 10th, Madison, Wis.
 25. Budowle B, Lindsey JA, DeCou JA, Koons BW, Giusti AM, Comey CT. 1995. Validation and population studies of the loci LDLR, GYPA, HBGG, D7S8, and Gc (PM loci), and HLA-DQ α using a multiplex amplification and typing procedure. *J. Forensic Sci.* 40:45–54
 26. Budowle B, Wilson MR, DiZinno JA, Stauffer C, Fasano MA, Holland MM. 1999. Mitochondrial DNA regions HVI and HVII population data. *Forensic Sci. Int.* 103:23–35
 27. Cann RL, Stoneking M, Wilson AC. 1987. Mitochondrial DNA and human evolution. *Nature* 325:31–36
 28. Carracedo A, Bar W, Mayr W, Morling N, Olaisen B, et al. 2000. DNA commission of the international society for forensic genetics: guidelines for mitochondrial DNA typing. *Forensic Sci. Int.* 110:79–85
 29. Case JT, Wallace DC. 1981. Maternal inheritance of mitochondrial DNA polymorphisms in cultured human fibroblasts. *Somat. Cell Genet.* 7:103–8
 30. Chakraborty R, Jin L, Deka R, Kimmel M, Budowle B. 2001. *Signatures of recurrent mutations at single nucleotide polymorphism sites in the hypervariable domains of the mitochondrial control region and their implications for evolutionary studies*. Presented at Am. Assoc. Phys. Anthropol., St. Louis, Mo.
 31. Chen Y, Olekers A, Schurr T, Kogelnik A, Huoponen K, Wallace D. 2000. MtDNA variation in the South African Kung! and Khwe and their genetic relationships to other African populations. *Am. J. Hum. Genet.* 66:1362–83
 32. Comas D, Paabo S, Bertranpetit J. 1995. Heteroplasmy in the control region of human mitochondrial DNA. *Genes Res.* 5:89–90
 33. Comey CT, Budowle B. 1991. Validation studies on the analysis of the HLA-DQ alpha locus using the polymerase chain reaction. *J. Forensic Sci.* 36:1633–48
 34. Cummins JM, Wakayama T, Yanagimachi R. 1997. Fate of microinjected sperm components in the mouse oocyte and embryo. *Zygote* 5:301–8
 35. Cummins JM, Wakayama T, Yanagimachi R. 1998. Fate of microinjected spermatid mitochondria in the mouse oocyte and embryo. *Zygote* 6:213–22
 36. Daoudi Y, Morgan M, Diefenbach C, Ryan J, Johnson T, et al. 1998. *Identification of the Vietnam tomb of the unknown soldier; the many roles of mitochondrial DNA*. Presented at Int. Symp. Huml Identif., 9th, Madison, Wis.
 - 36a. Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993)
 37. D'Eustachio P. 2002. Letter to the editor.

- High levels of mitochondrial DNA heteroplasmy in human hairs by Budowle et al. *Forensic Sci. Int.* 130(1):63–67
38. D'Eustachio P. 2000. People v. Edmund Ko, Indictment No. 2449/98, New York County, May 10, pp. 528–31
 39. Edwards A, Civitello A, Hammond HA, Caskey CT. 1991. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *Am. J. Hum. Genet.* 49:746–56
 40. Eyre-Walker A, Smith N, Maynard Smith J. 1999. How clonal are human mitochondria? *Proc. R. Soc. London Ser. B* 266:477–83
 41. Eyre-Walker A, Smith N, Maynard Smith J. 1999. Reply to Macaulay et al. (1999): mitochondrial DNA recombination—reasons to panic. *Proc. R. Soc. London Ser. B* 266:2041–42
 42. Finnila S, Lehtonen MS, Majamaa K. 2001. Phylogenetic network for European mtDNA. *Am. J. Hum. Genet.* 68:1475–84
 43. Foran D, Hixson JE, Brown WM. 1988. Comparisons of ape and human: sequences that regulate mitochondrial DNA transcription and D-loop DNA synthesis. *Nucleic Acids Res.* 17:5841–61
 44. Fucharoen G, Fucharoen S, Horai S. 2001. Mitochondrial DNA polymorphisms in Thailand. *J. Hum. Genet.* 46:115–25
 45. Giles RE, Blanc H, Cann HM, Wallace DC. 1980. Maternal inheritance of human mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 77:6715–19
 46. Gill P, Ivanov PL, Kimpton C, Piercy R, Benson N, et al. 1994. Identification of the remains of the Romanov family by DNA analysis. *Nat. Genet.* 6:130–35
 47. Gill P, Jeffreys AJ, Werrett DJ. 1985. Forensic application of DNA fingerprints. *Nature* 318:577–79
 48. Ginther C, Issel-Tarver L, King MC. 1992. Identifying individuals by sequencing mitochondrial DNA from teeth. *Nat. Genet.* 2:135–38
 49. Goloboff P. 1994. NONA: a tree search program. ftp.unt.edu.ar/pub/parsimony and www.cladistics.org
 50. Greenberg BD, Newbold JE, Sugino A. 1983. Intraspecific nucleotide sequence variability surrounding the origin of replication in human mitochondrial DNA. *Gene* 21:33–49
 51. Gyllensten U, Wharton D, Josefsson A, Wilson AC. 1991. Paternal inheritance of mitochondrial DNA in mice. *Nature* 352:255–57
 52. Grzybowski T. 2000. Extremely high levels of human mitochondrial DNA heteroplasmy in single hair roots. *Electrophoresis* 21:548–53
 53. Grzybowski T. 2000. Erratum. *Electrophoresis* 21
 54. Hagelberg E, Goldman N, Lio P, Whelan S, Schiefenhover W, et al. 1999. Evidence for mitochondrial DNA recombination in a human population of island Melanesia. *Proc. R. Soc. London Ser. B* 266:485–92
 55. Hagelberg E, Goldman N, Lio P, Whelan S, Schiefenhover W, et al. 2000. Evidence for mitochondrial DNA recombination in a human population of island Melanesia: correction. *Proc. R. Soc. London Ser. B* 267:1595–96
 56. Hall BG. 2001. *Phylogenetic Trees Made Easy: A How-To Manual For Molecular Biologists*. Sunderland, MA: Sinauer Assoc. 179 pp.
 57. Hanni C, Laudet V, Sakka M, Begue A, Stehelin D. 1990. Amplification of mitochondrial DNA fragments from ancient human teeth and bones. *C. R. Acad. Sci. III* 310:365–70
 58. Helgason A, Hickey E, Goodacre S, Bosnes V, Stefansson K, et al. 2001. mtDNA and the islands of the North Atlantic: estimating the proportions of Norse and Gaelic ancestry. *Am. J. Hum. Genet.* 68:723–37
 59. Hillis DM, Moritz C, Mable BK. 1996. *Molecular Systematics*. Sunderland, MA: Sinauer Assoc. 655 pp. 2nd ed.
 60. Holland MM, Fisher DL, Mitchell LG,

- Rodriguez WC, Canik JJ, et al. 1993. Mitochondrial DNA sequence analysis of human skeletal remains: identification of remains from the Vietnam War. *J. Forensic Sci.* 38:542–53
61. Holland MM, Parsons TJ. 1999. Mitochondrial DNA sequence analysis—validation and use for forensic casework. *Forensic Sci. Rev.* 11:21–50
 62. Hopgood R, Sullivan KM, Gill P. 1992. Strategies for automated sequencing of human mitochondrial DNA directly from PCR products. *Biotechniques* 13:82–92
 63. Horai S, Murayama K, Hayasaka K, Matsubayashi S, Hattori Y, et al. 1996. mtDNA polymorphism in east Asian populations with special reference to the peopling of Japan. *Am. J. Hum. Genet.* 59:579–90
 64. Hutchinson CA III, Newbold JE, Potter SS, Edgell MH. 1974. Maternal inheritance of mammalian mitochondrial DNA. *Nature* 251:536–38
 65. Innan H, Nordborg M. 2002. Recombination or mutational hot spots in human mtDNA? *Mol. Biol. Evol.* 19:1122–27
 66. Ivanov PL, Wadhams MJ, Roby RK, Holland MM, Weedn VW, Parsons TJ. 1996. Mitochondrial DNA sequence heteroplasmy in the Grand Duke of Russia Georgij Romanov establishes the authenticity of the remains of Tsar Nicholas II. *Nat. Genet.* 12:417–20
 67. Jeffreys AJ, Wilson V, Thein SL. 1985. Hypervariable minisatellite regions in human DNA. *Nature* 314:67–73
 68. Jeffreys AJ, Wilson V, Thein SL. 1985. Individual-specific fingerprints of human DNA. *Nature* 316:76–79
 69. Jorde LB, Bamshad M. 2000. Questioning evidence for recombination in human mitochondrial DNA. *Science* 288:1931a
 70. Kasai K, Nakamura Y, White R. 1990. Amplification of a variable number of tandem repeat (VNTR) locus (pMCT118) by the polymerase chain reaction (PCR) and its application to forensic science. *J. Forensic Sci.* 35:1196–200
 71. Kimmel M, Bobrowski A, Wang N, Budowle B, Chakraborty R. 2001. *Non-homogeneous infinite sites model under demographic changes of population size: application to mitochondrial DNA data.* Presented at Am. Assoc. Phys. Anthropol., St. Louis, Mo.
 72. Kivisild T, Tolk H, Parik J, Wang Y, Papiha S, et al. 2002. The emerging limbs and twigs of the East Asian mtDNA tree. *Mol. Biol. Evol.* 19:1737–51
 73. Kivisild T, Villems R. 2000. Questioning evidence for recombination in human mitochondrial DNA. *Science* 288:1931a
 74. Krings M, Stone A, Schmitz RW, Krainitzki H, Stoneking M, Pääbo S. 1997. Neandertal DNA sequences and the origin of modern humans. *Cell* 90:19–30
 75. Kumar S, Hedrick P, Dowling T, Stoneking M. 2000. Questioning evidence for recombination in human mitochondrial DNA. *Science* 288:1931a
 76. Lord WD, DiZinno JA, Wilson MR, Budowle B, Taplin D, Meinking TL. 1998. Isolation, amplification, and sequencing of human mitochondrial DNA obtained from human crab louse, *Phthirus pubis* (L.), blood meals. *J. Forensic Sci.* 43:1097–100
 77. Lutz S, Weisser HJ, Heizmann J, Pollak S. 1996. mtDNA as a tool for identification of human remains. *Int. J. Leg. Med.* 109:205–9
 78. Kurosaki K, Matsushita T, Ueda S. 1993. Individual DNA identification from ancient human remains. *Am. J. Hum. Genet.* 53:638–43
 79. Macaulay V, Richards M, Hickey E, Vega E, Cruciani F, et al. 1999. The emerging tree of West Eurasian mtDNAs: a synthesis of control region sequences and RFLPs. *Am. J. Hum. Genet.* 4:232–49
 80. Macaulay V, Richards M, Sykes B. 1999. Mitochondrial DNA recombination—no need to panic. *Proc. R. Soc. London Ser. B* 266:2037–39
 81. Melton T, Clifford S, Kayser M, Nasidze I, Batzer M, Stoneking M. 2001. Diversity and heterogeneity in mitochondrial DNA

- of North American populations. *J. Forensic Sci.* 46:46–52
82. Melton T, Ginther C, Sensabaugh G, Soodyall H, Stoneking M. 1997. Extent of heterogeneity in mitochondrial DNA of sub-Saharan African populations. *J. Forensic Sci.* 42:582–92
 83. Melton T, Stoneking M. 1996. Extent of heterogeneity in mitochondrial DNA of ethnic Asian populations. *J. Forensic Sci.* 41:591–602
 84. Monnat RJ, Loeb LA. 1985. Nucleotide sequence preservation of human mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 82:2895–99
 85. Monnat RJ, Reay DT. 1986. Nucleotide sequence identity of mitochondrial DNA from different human tissues. *Gene* 43:205–11
 86. Monson K, Miller K, Wilson M, DiZinno J, Budowle B. 2002. The mtDNA population database: an integrated software and database resource for forensic comparison. *Forensic Sci. Com.* 4(2)
 87. Mourier T, Hansen AJ, Willersev E, Arcander P. 2001. The human genome project reveals a continuous transfer of large mitochondrial fragments to the nucleus. *Mol. Biol. Evol.* 18:1833–37
 88. Natl. Res. Council. II Rep. 1996. *The Evaluation of Forensic Evidence*. Washington, DC: Natl. Acad. Press
 89. Oota H, Kurosaki K, Pookajorn S, Ishida T, Ueda S. 2001. Genetic study of the Paleolithic and Neolithic Southeast Asians. *Hum. Biol.* 73:225–31
 90. Oota H, Saitou N, Matsushita T, Ueda S. 1999. Molecular genetic analysis of remains of a 2,000-year-old human population in China—and its relevance for the origin of the modern Japanese population. *Am. J. Hum. Genet.* 64:250–58
 91. Oota H, Saitou N, Matsushita T, Ueda S. 1995. A genetic study of 2,000-year-old human remains from Japan using mitochondrial DNA sequences. *Am. J. Phys. Anthropol.* 98:133–45
 92. Pääbo S. 1989. Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. *Proc. Natl. Acad. Sci. USA* 86:6196–200
 93. Pääbo S, Gifford JA, Wilson AC. 1988. Mitochondrial DNA sequences from a 7000-year-old brain. *Nucleic Acids Res.* 16:9775–78
 94. Pääbo S, Higuchi R, Wilson AC. 1989. Ancient DNA and the polymerase chain reaction. The emerging field of molecular archaeology. *J. Biol. Chem.* 264:9707–12
 95. Page RDM, Holmes EC. 1998. *Molecular Evolution: A Phylogenetic Approach*. Oxford, UK: Blackwell Sci.
 96. Parr RL, Carlyle SW, O'Rourke DH. 1996. Ancient DNA analysis of Fremont Amerindians of the Great Salt Lake Wetlands. *Am. J. Phys. Anthropol.* 99:507–18
 97. Pereira L, Macaulay V, Torroni A, Scozzari R, Prata J, Amorim A. 2001. Prehistoric and historic traces in the mtDNA of Mozambique: insights into the Bantu expansions and the slave trade. *Am. J. Hum. Genet.* 65:439–56
 98. Pfeiffer H, Huhne J, Ortmann C, Waterkamp K, Brinkmann B. 1999. Mitochondrial DNA typing from human axillary, pubic and head hair shafts—success rates and sequence comparisons. *Int. J. Leg. Med.* 112:287–90
 99. Rando J, Pinto F, Gonzales A, Hernandez M, Larruga J, et al. 1998. Mitochondrial DNA analysis of Northwest African populations reveals genetic exchanges with European, Near-Eastern and sub-Saharan populations. *Am. J. Hum. Genet.* 62:531–50
 100. Saiki RK, Walsh PS, Levenson CH, Erlich HA. 1989. Genetic analysis of amplified DNA with immobilized sequence-specific oligonucleotide probes. *Proc. Natl. Acad. Sci. USA* 86:6230–34
 101. Schneider PM, Seo Y, Rittner C. 1999. Forensic mtDNA hair analysis excludes a dog from having caused a traffic accident. *Int. J. Leg. Med.* 112:315–16
 102. Schwartz M, Vissing J. 2002. Paternal

- inheritance of mitochondrial DNA. *N. Engl. J. Med.* 347:576–80
103. Shitara H, Hayashi JI, Takahama S, Kaneda H, Yonekawa H. 1998. Maternal inheritance of mouse mtDNA in interspecific hybrids: segregation of leaked paternal mtDNA followed by the prevention of subsequent paternal leakage. *Genetics* 148:851–57
 104. Shitara H, Kaneda H, Sato A, Inoue K, Ogura A, et al. 2000. Selective and continuous elimination of mitochondria microinjected into mouse eggs from spermatids, but not from liver cells, occurs throughout embryogenesis. *Genetics* 156:1277–84
 105. Stewart JEB, Fisher CL, Aagaard PJ, Wilson MR, Isenberg AR, et al. 2001. Length variation in HV2 of the human mitochondrial DNA control region. *J. Forensic Sci.* 46:862–70
 106. Stoneking M, Hedgecock D, Higuchi RG, Vigilant L, Erlich HA. 1991. Population variation of human mtDNA control region sequences detected by enzymatic amplification and sequence specific oligonucleotide probes. *Am. J. Hum. Gen.* 48:370–82
 107. Strauss E. 1999. MtDNA shows signs of paternal influence. *Science* 286:2436
 108. Sullivan KM, Hopgood R, Lang B, Gill P. 1991. Automated amplification and sequencing of human mitochondrial DNA. *Electrophoresis* 12:17–21
 109. Sullivan KM, Hopgood R, Gill P. 1992. Identification of human remains by amplification and automated sequencing of mitochondrial DNA. *Int. J. Leg. Med.* 105:83–86
 110. Thyagarajan B, Padua RA, Campbell C. 1996. Mammalian mitochondria possess homologous DNA recombination activity. *J. Biol. Chem.* 271:27536–43
 111. Torroni A, Huoponen K, Francalacci P, Petrozzi M, Morelli L, et al. 1996. Classification of European mtDNAs from an analysis of three European populations. *Genetics* 114:1835–50
 112. Tully G, Bar W, Brinkmann B, Carracedo A, Gill P, et al. 2001. Considerations by the European DNA profiling (EDNAP) group on the working practices, nomenclature and interpretations of mitochondrial DNA profiles. *Forensic Sci. Int.* 124:83–91
 113. Tully G, Lareu M. 2001. Letter to the editor: Controversies over heteroplasmy. *Electrophoresis* 22:180–81
 114. Tully LA, Parsons TJ, Steighner RJ, Holland MM, Marino MA, Prenger VL. 2000. A sensitive denaturing gradient-Gel electrophoresis assay reveals a high frequency of heteroplasmy in hypervariable region I of the human mtDNA control region. *Am. J. Hum. Genet.* 67:432–43
 115. Vigilant L, Pennington R, Harpending H, Kocher TD, Wilson AC. 1989. Mitochondrial DNA sequences in single hairs from a southern African population. *Proc. Natl. Acad. Sci. USA* 86:9350–54
 116. Vigilant LM, Stoneking M, Harpending H, Hawkes K, Wilson AC. 1991. African populations and the evolution of human mitochondrial DNA. *Science* 253:1503–7
 117. Wang L, Oota H, Saitou N, Jin F, Matsushita T, Ueda S. 2000. Genetic structure of a 2,500-year-old human population in China and its spatiotemporal changes. *Mol. Biol. Evol.* 17:1396–400
 118. Wilson MR, Allard MW, Monson KL, Miller K, Budowle B. 2002. Recommendations for consistent treatment of length variants in the human mitochondrial DNA control region. *Forensic Sci. Int.* 129:35–42
 119. Wilson MR, Allard MW, Monson KL, Miller K, Budowle B. 2002. Further discussions of the consistent treatment of length variants in the human mitochondrial DNA control region. *Forensic Sci. Com.* 4(4)
 120. Wilson MR, DiZinno JA, Polansky D, Replogle J, Budowle B. 1995. Validation of mitochondrial DNA sequencing for forensic casework analysis. *Int. J. Leg. Med.* 108:68–74
 121. Wilson MR, Polansky D, Butler J,

- DiZinno JA, Replogle J, Budowle B. 1995. Extraction, PCR amplification, and sequencing of mitochondrial DNA from human hair shafts. *Biotechniques* 18:662–69
122. Wilson MR, Polanskey D, Replogle J, DiZinno JA, Budowle B. 1997. A family exhibiting heteroplasmy in the human mitochondrial DNA control region reveals both somatic mosaicism and pronounced segregation of mitotypes. *Nat. Genet.* 100:167–71
123. Wyman AR, White R. 1980. A highly polymorphic locus in human DNA. *Proc. Natl. Acad. Sci. USA* 77:6754–58

APPENDIX

Example 1

The sequences of the sample and the Cambridge Reference Sequence (CRS) are from nucleotide positions 244–253 in HV1. The sample has a deletion.

ATTGATGTC	Sample
ATTGAATGTC	CRS

One possible alignment places a deletion at nucleotide position 248 (as shown below).

ATTG–ATGTC	Sample
ATTG <u>A</u> ATGTC	CRS

Another possible alignment places the deletion between the A and T residues (as shown below).

ATTGA–TGTC	Sample
ATTGA <u>A</u> TGTC	CRS

Because both alignments are possible based on a single deletion, recommendations 1 and 2 cannot resolve the choice of alignments; recommendation 3 is needed. The deletion is placed at the 3' end with respect to the light strand. Hence, the latter alignment is selected and the sequence is listed as 249D.

Example 2

Homopolymeric stretches of Cs with a T residing within the stretch can be found in some individuals at nucleotide positions 16180–16193. In this example (positions 16180–16198), T is found at nucleotide position 16186 rather than nucleotide position 16189. Also, the total number of bases between positions 16180 and 16193 is less than the 14 in the CRS.

AAAACCTCCCCCATGCT	Sample
AAAACCCCTCCCCATGCT	CRS

One possible alignment (shown below) would list the sequence as 16186T and 16189D.

```

AAAACCTCC-CCCCATGCT      Sample
AAAACCCCCCTCCCCATGCT    CRS

```

Another possible alignment (shown below) requires three changes to describe it in relation to the CRS.

```

AAAACC-TCCCCCATGCT      Sample
AAAACCCCCCTCCCCATGCT    CRS

```

A third alignment (shown below) also requires three changes, i.e., two transitions and a deletion.

```

AAAACCTCCCCC-ATGCT
AAAACCCCCCTCCCCATGCT

```

The first alignment is preferred; it requires the least number of differences to describe it.



CONTENTS

GENETICS OF HUMAN LATERALITY DISORDERS: INSIGHTS FROM VERTEBRATE MODEL SYSTEMS, <i>Brent W. Bisgrove, Susan H. Morelli, and H. Joseph Yost</i>	1
RACE, ANCESTRY, AND GENES: IMPLICATIONS FOR DEFINING DISEASE RISK, <i>Rick A. Kittles and Kenneth M. Weiss</i>	33
GENE ANNOTATION: PREDICTION AND TESTING, <i>Jennifer L. Ashurst and John E. Collins</i>	69
THE DROSOPHILA MELANOGASTER GENOME, <i>Susan E. Celniker and Gerald M. Rubin</i>	89
FORENSICS AND MITOCHONDRIAL DNA: APPLICATIONS, DEBATES, AND FOUNDATIONS, <i>Bruce Budowle, Marc W. Allard, Mark R. Wilson, and Ranajit Chakraborty</i>	119
CREATIONISM AND INTELLIGENT DESIGN, <i>Robert T. Pennock</i>	143
PEROXISOME BIOGENESIS DISORDERS, <i>Sabine Weller, Stephen J. Gould, and David Valle</i>	165
SEQUENCE DIVERGENCE, FUNCTIONAL CONSTRAINT, AND SELECTION IN PROTEIN EVOLUTION, <i>Justin C. Fay and Chung-I Wu</i>	213
MOLECULAR PATHOGENESIS OF PANCREATIC CANCER, <i>Donna E. Hansel, Scott E. Kern, and Ralph H. Hruban</i>	237
THE INHERITED BASIS OF DIABETES MELLITUS: IMPLICATIONS FOR THE GENETIC ANALYSIS OF COMPLEX TRAITS, <i>Jose C. Florez, Joel Hirschhorn, and David M. Altshuler</i>	257
PATTERNS OF HUMAN GENETIC DIVERSITY: IMPLICATIONS FOR HUMAN EVOLUTIONARY HISTORY AND DISEASE, <i>Sarah A. Tishkoff and Brian C. Verrelli</i>	293
HUMAN NONSYNDROMIC SENSORINEURAL DEAFNESS, <i>Thomas B. Friedman and Andrew J. Griffith</i>	341
ENZYME THERAPY FOR LYSOSOMAL STORAGE DISEASE: PRINCIPLES, PRACTICE, AND PROSPECTS, <i>Gregory A. Grabowski and Robert J. Hopkin</i>	403
NONSYNDROMIC SEIZURE DISORDERS: EPILEPSY AND THE USE OF THE INTERNET TO ADVANCE RESEARCH, <i>Mark F. Leppert and Nanda A. Singh</i>	437

THE GENETICS OF NARCOLEPSY, *Dorothee Chabas, Shahrad Taheri, Corinne Renier, and Emmanuel Mignot* 459

INDEXES

Subject Index 485
Cumulative Index of Contributing Authors, Volumes 1–4 503
Cumulative Index of Chapter Titles, Volumes 1–4 505

ERRATA

An online log of corrections to *Annual Review of Genomics and Human Genetics* chapters (if any) may be found at <http://genom.annualreviews.org/>