CHAPTER

14

Mitochondrial DNA Analysis

Conventional STR typing systems do not work in every instance—even with the development of miniSTR assays mentioned in Chapter 10. Ancient DNA specimens or samples that have been highly degraded often fail to produce results with nuclear DNA typing systems. However, recovery of DNA information from damaged DNA is sometimes possible with mitochondrial DNA (mtDNA). While a nuclear DNA test is usually more valuable, a mtDNA result is better than no result at all.

Because there are hundreds if not thousands of copies of mtDNA in each cell, the probability of obtaining a DNA typing result from mtDNA is higher than that of polymorphic markers found in nuclear DNA, particularly in cases where the amount of extracted DNA is very small, as in tissues such as bone, teeth, and hair. When remains are quite old or badly degraded, often bone, teeth, and hair are the only biological sources left from which to draw a sample.

This chapter will review the characteristics of mitochondrial DNA, the steps involved in obtaining results in forensic casework, and issues important to interpreting mtDNA results.

CHARACTERISTICS OF mtDNA

The primary characteristic that permits mitochondrial DNA (mtDNA) recovery from degraded samples is the higher copy number of mtDNA in cells relative to the nuclear DNA from which STRs are amplified. In short, though nuclear DNA contains much more information, there are only two copies of it in each cell (one maternal and one paternal) while mtDNA provides a bit of useful genetic information hundreds of times per cell. Because of their higher numbers, some mtDNA molecules are more likely to survive than nuclear DNA. Table 14.1 contains a comparison of some basic characteristics of nuclear DNA and mitochondrial DNA.

Location and Structure of mtDNA

The vast majority of the human genome is located within the nucleus of each cell (see Table 14.1). However, there is a small, circular genome found within the mitochondria, the

Advanced Topics in Forensic DNA Typing: Methodology.
DOI: 10.1016/B978-0-12-374513-2.00014-2

TABLE 14.1 Comparison of Human Nuclear DNA and Mitochondrial DNA Markers.

Characteristics	Nuclear DNA	Mitochondrial DNA (mtDNA)
Size of genome	≈3.2 billion bp	≈16,569 bp
Copies per cell	2 (1 allele from each parent)	Can be >1000
Percent of total DNA	99.75%	0.25% content per cell
Structure	Linear; packaged in chromosomes	Circular
Inherited from	Father and mother	Mother
Chromosomal pairing	Diploid	Haploid
Generational recombination	Yes	No
Replication repair	Yes	No
Unique	Unique to individual (except identical twins)	Not unique to individual (same as maternal relatives)
Mutation rate	Low	At least 5–10 times nuclear DNA
Reference sequence	Described in 2001 by the Human Genome Project	Described in 1981 by Anderson and co-workers

energy-producing cellular organelle residing in the cytoplasm. The number of mtDNA molecules within a cell can range from hundreds to thousands. On average there are 4 to 5 copies of mtDNA molecules per mitochondrion with a measured range of 1 to 15 (Satoh & Kuroiwa 1991). Because each cell can contain hundreds of mitochondria (Robin & Wong 1988), there can be up to several thousand mtDNA molecules in each cell as in the case of ovum or egg cells. However, the average has been estimated at about 500 in most cells (Satoh & Kuroiwa 1991). It is this large number of mtDNA molecules in each cell that enables greater success (relative to nuclear DNA markers) with biological samples that may have been damaged with heat or humidity.

Mitochondrial DNA has approximately 16,569 base pairs with the total number of nucleotides in a specific mtDNA genome (mtGenome) varying due to small insertions or deletions. For example, there is a dinucleotide repeat at positions 514 to 524, which in most individuals is ACACACACAC or $(AC)_5$ but has been observed to vary from $(AC)_3$ to $(AC)_7$ (Bodenteich et al. 1992, Szibor et al. 1997). Note that with two copies of nuclear DNA (3.2 billion bp from each parent) and even assuming that there are 1000 copies of mtDNA (16,569 bp per mtDNA) in a cell, mtDNA makes up only about 0.25% of the total DNA content of a cell.

Most of the mtGenome codes for 37 gene products used in the oxidative phosphorylation process or cellular energy production (Figure 14.1). The 37 transcribed "genes" of mtDNA found in the "coding region" include 13 proteins, 2 ribosomal RNAs (rRNA), and 22 transfer RNAs (tRNA). There is also a 1122bp "control" region that contains the origin of replication for one of the mtDNA strands but does not code for any gene products and is therefore referred to sometimes as the "non-coding" region.

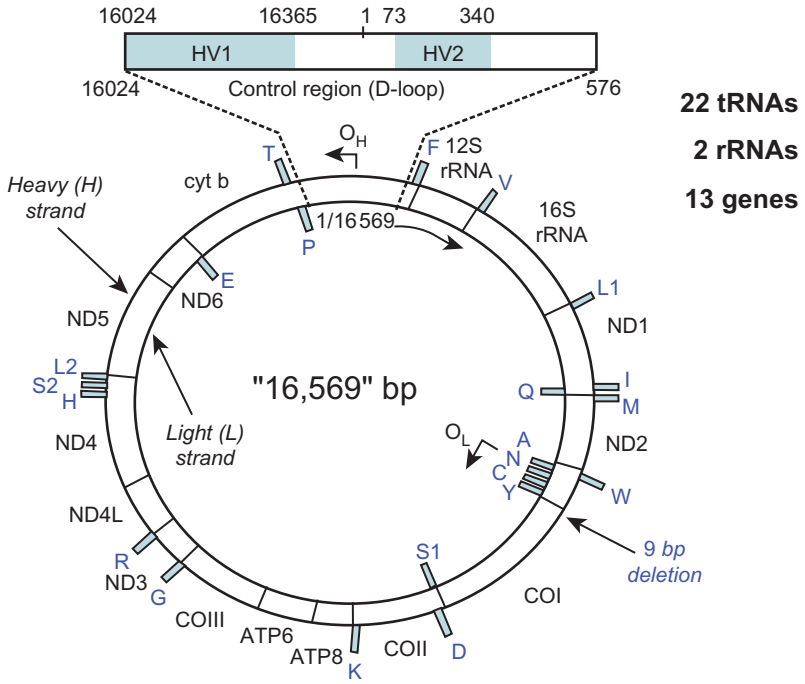

FIGURE 14.1 Illustration of the circular mitochondrial DNA genome (mtGenome). The heavy (H) strand is represented by the outside line and contains a higher number of C-G residues than the light (L) strand. The 37 RNA and protein coding gene regions are abbreviated around the mtGenome next to the strand from which they are synthesized (see Table 14.2). Most forensic mtDNA analyses presently examine only HV1 and HV2 (and occasionally HV3) in the non-coding control region or displacement loop (D-loop) shown at the top of the figure. Due to insertions and deletions that exist around the mtGenome in different individuals, it is not always 16,569 bp in length.

The nucleotide positions for each coding and non-coding segment of the mtGenome are indicated in Table 14.2. Note that the genes are very tightly packed with only 55 nucleotides in the 15,447 bp of the coding region *not* being used to transcribe a protein, rRNA, or tRNA molecule. Thus, the genes within mtDNA are economically packaged with no introns and none or only a few non-coding nucleotides between the coding regions.

An asymmetric distribution of nucleotides in the mtGenome gives rise to a "light" and "heavy" strand when mtDNA molecules are separated in alkaline CsCl gradients (Scheffler 1999). The "heavy" or H-strand contains a greater number of guanine nucleotides, which have the largest relative molecular mass of the four nucleotides (A, T, C, and G), than the "light" or L-strand. Replication of mtDNA begins with the H-strand in the non-coding "control region," also known as the displacement loop or D-loop (Figure 14.1). A total of 28 gene products are encoded from the H-strand while the L-strand transcribes eight tRNAs and an enzyme called ND6 (Table 14.2).

Since the D-loop does not code for gene products, the constraints are fewer for nucleotide variability, and polymorphisms between individuals are more abundant than in

TABLE 14.2 Mitochondrial DNA Information and Genes.

Nucleotide Position	Strand Transcribed	Abbreviation	Description	Size (bp)	Number of Non-coding Nucleotides
16024–16569, 1–576		D-loop	Control region	1122	1122
16104–16569, 1–191		OH	Replication origin (H-strand)	658	
16158–16172			D-loop termination signal	15	
531–568			H-strand transcription promoter	38	
577–647	H	F	tRNA phenylalanine	71	
648–1601	H	12S	12S rRNA	954	
1602–1670	H	V	tRNA valine	69	
1671–3229	H	16S	16S rRNA	1559	
3230–3304	H	L1	tRNA leucine 1	75	
3305–4263	H	ND1	NADH dehydrogenase 1	959	
4263–4331	H	I	tRNA isoleucine	69	
4329–4400	L	Q	tRNA glutamine	72	
4401		—	Non-coding	1	1
4402–4469	H	M	tRNA methionine	68	
4470–5511	H	ND2	NADH dehydrogenase 2	1042	
5512–5579	H	W	tRNA tryptophan	68	
5580–5586		—	Non-coding	7	7
5587–5655	L	A	tRNA alanine	69	
5656		—	Non-coding	1	1
5657–5729	L	N	tRNA asparagine	73	
5730–5760		OL	L-strand origin	31	31
5761–5826	L	C	tRNA cysteine	66	
5826–5891	L	Y	tRNA tyrosine	66	
5892–5900		—	Non-coding	9	9
5901–7445	H	COI	Cytochrome c oxidase I	1545	
7445–7516	L	S1	tRNA serine 1	72	
7517		—	Non-coding	1	1
7518–7585	H	D	tRNA aspartic acid	68	
7586–8294	H	COII	Cytochrome c oxidase II	709	
8295–8364	H	K	tRNA lysine	70	

(*Continued*)

TABLE 14.2 Mitochondrial DNA Information and Genes. (*Continued*)

Nucleotide Position	Strand Transcribed	Abbreviation	Description	Size (bp)	Number of Non-coding Nucleotides
8365–8572	H	ATP8	ATP synthase 8	208	
8527–9207	H	ATP6	ATP synthase 6	681	
9207–9990	H	COIII	Cytochrome c (oxidase III)	784	
9991–10058	H	G	tRNA glycine	68	
10059–10404	H	ND3	NADH (dehydrogenase 3)	346	
10405–10469	H	R	tRNA arginine	65	
10470–10766	H	ND4L	NADH (dehydrogenase 4L)	297	
10760–12137	H	ND4	NADH (dehydrogenase 4)	1378	
12138–12206	H	H	tRNA histidine	69	
12207–12265	H	S2	tRNA serine 2	59	
12266–12336	H	L2	tRNA leucine 2	71	
12337–14148	H	ND5	NADH (dehydrogenase 5)	1812	
14149–14673	L	ND6	NADH (dehydrogenase 6)	525	
14674–14742	L	E	tRNA glutamic acid	69	
14743–14746		—	Non-coding	4	4
14747–15887	H	cyt b	Cytochrome b	1141	
15888–15953	H	T	tRNA threonine	66	
15954		—	Non-coding	1	1
15955–16023	L	P	tRNA proline	69	

similarly sized portions of the coding region. More simply, there can be differences in the D-loop region because the sequences do not code for any substances necessary for the cell's function.

Most of the focus in forensic DNA studies to date has involved two hypervariable regions within the control region commonly referred to as HVI (HV1) and HVII (HV2). Occasionally a third portion of the control region, known as HV3, is examined to provide more information regarding a tested sample.

The numbering system for human mtDNA nucleotide positions is arbitrarily based on the L-strand from an *MboI* restriction enzyme site within the control region as defined in the original paper describing the mtGenome sequence (Anderson et al. 1981). Thus, position 1 is not the origin of replication. As can be seen in Figure 14.1, position 1 falls between hypervariable region 1 (HV1) and hypervariable region 2 (HV2).

Human mtDNA Reference Sequence(s)

Human mtDNA was first sequenced in 1981 in the laboratory of Frederick Sanger in Cambridge, England (Anderson et al. 1981). For many years, the original "Anderson" sequence (named after the first author listed in alphabetical order from the Sanger research group) was the reference sequence (GenBank accession: M63933) to which new sequences were compared. The Anderson sequence is also referred to as the Cambridge Reference Sequence (CRS). In 1999, the original placental material used by Anderson and co-workers to generate the CRS was re-sequenced (Andrews et al. 1999).

The 1981 sequence was derived primarily from a single individual of European descent; however, it also contained some HeLa and bovine sequences to fill in gaps resulting from early rudimentary DNA sequencing procedures (Anderson et al. 1981). With improvements in DNA sequencing technology over the intervening two decades, it was felt that any original errors should be rectified to enable robust use of this reference sequence in the future.

The reanalysis effort confirmed all but 11 of the original nucleotides identified in the original published sequence (Table 14.3). One of these differences was the loss of a single cytosine residue at position 3107. An additional seven nucleotide positions were demonstrated to be accurate but represent rare polymorphisms. These sites were 263A, 311–315CCCCC, 750A, 1438A, 4769A, 8860A, and 15326A. Fortunately, no errors were observed in the widely used control region. Thus, the original Anderson sequence (Anderson et al. 1981) was found to be

TABLE 14.3 Comparison of Nucleotide Differences Observed Between the Original Cambridge Reference Sequence (Anderson et al. 1981) and the Revised Cambridge Reference Sequence (Andrews et al. 1999) Based on Re-sequencing of the Original Placenta Material. The True Sequence at Position 3106–3107 Is Only a Single C Making the Entire mtGenome 16,568 bp Rather than the Originally Reported 16,569 bp. However, to Maintain the Historical Numbering, a Deletion at Position 3107 Is Used to Serve as a Placeholder (Andrews et al. 1999). Note That No Differences Exist Between These Sequences for the Two Hypervariable Regions Most Commonly Used in Forensic Applications That Span Positions 16024 to 16365 and 73 to 340. See MITOMAP for a Fully Annotated Version of the rCRS: http://www.mitomap.org/bin/view.pl/MITOMAP/HumanMitoSeq.

Nucleotide Position	Region of mtGenome	Original CRS	Revised CRS (rCRS)	Remarks
3106–3107	16S rRNA	CC	C	Error
3423	ND1	G	T	Error
4985	ND2	G	A	Error
9559	COIII	G	C	Error
11335	ND4	T	C	Error
13702	ND5	G	C	Error
14199	ND6	G	T	Error
14272	ND6	G	C	Error (bovine sequence inserted)
14365	ND6	G	C	Error (bovine sequence inserted)
14368	ND6	G	C	Error
14766	cyt b	T	C	Error (HeLa sequence inserted)

identical to the revised Cambridge reference sequence (Andrews et al. 1999) across the HV1 and HV2 regions that are widely used in forensic applications.

This revised Cambridge reference sequence (rCRS) is now the accepted standard for comparison and is available in GenBank using NCBI reference sequence NC_012920.1. While the loss of a single C nucleotide at position 3107 means that the reference mtGenome is 16,568 bp rather than the traditionally accepted value of 16,569 bp, historical numbering has been maintained by including an "N" in place of the 3107 deletion. Not including this extra space would have created an unacceptable amount of confusion and inability to easily correlate previous work. Therefore, Andrews and co-workers (1999) recommended that the original numbering be retained in the rCRS with a deletion in the sequence at position 3107 to serve as a place holder. The "16,569 bp" rCRS is available at the MITOMAP website: http://www.mitomap.org/mitoseq.htm.

Maternal Inheritance of mtDNA

Human mitochondrial DNA is considered to be inherited strictly from our mothers. At conception only the sperm's nucleus enters the egg and joins directly with the egg's nucleus. The fertilizing sperm is not believed to contribute other cellular components. When the zygote cell divides and a blastocyst develops, the cytoplasm and other cell parts save the nucleus are consistent with the mother's original egg cell. Mitochondria with their mtDNA molecules are passed directly to all offspring independent of any male influence. Thus, barring mutation, a mother passes along her mtDNA type to her children, and therefore siblings and maternal relatives have an identical mtDNA sequence. Hence, an individual's mtDNA type is not unique to him or her.

An example family pedigree is shown in Figure 14.2 to demonstrate the inheritance pattern of mtDNA. In this example, unique mtDNA types exist solely for individuals 1, 5, 7,

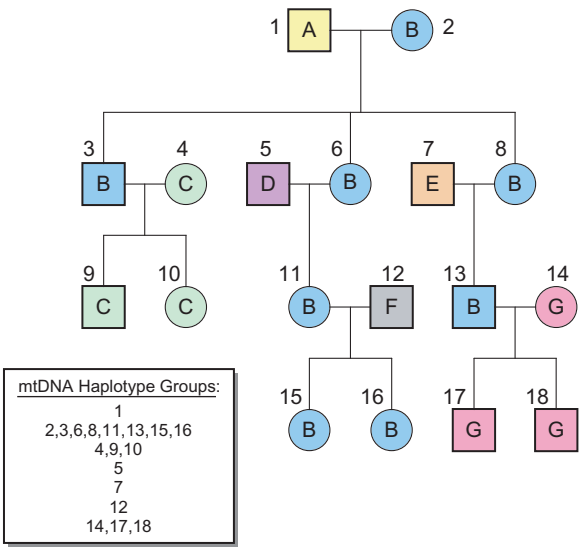

FIGURE 14.2 Illustration of maternal mitochondrial DNA inheritance for 18 individuals in a hypothetical pedigree. Squares represent males and circles females. Each unique mtDNA type is represented by a different letter.

and 12. Note that individual 16 will possess the same mtDNA type as seven of the other represented individuals (i.e., 2, 3, 6, 8, 11, 13, and 15). This can be helpful in solving missing persons or mass disaster investigations but will likely reduce the significance of a match in forensic cases. Since even distantly related maternal relatives should possess the same mtDNA type, this extends the number of useful reference samples that may be used to confirm the identity of a missing person.

Other Interesting Differences Between mtDNA and Nuclear DNA

Mitochondrial DNA uses a different genetic code than nuclear DNA (Scheffler 1999). For example, the codon for mitochondrial-transcribed amino acid tryptophan is UGA while the universal (nuclear) genetic code for UGA is a stop codon. In the mtDNA genetic code, AUA codes for methionine instead of isoleucine and AGA and AGG both code for stops rather than arginine.

Fewer DNA repair mechanisms exist in mitochondria thereby leading to higher mutation rates compared to nuclear DNA. In addition, lack of proofreading capabilities in the mtDNA polymerase increases mutations during replication. The 10-fold higher mutation rate (relative to nuclear DNA) helps introduce more variability in samples from identical maternal lineages that otherwise would not vary. This increased variability is a good thing for most applications in human identity testing although mutations can sometimes be a hindrance when trying to definitely establish familial relationships (e.g., when comparing remains to reference samples from distant maternal relatives).

The circular nature of mtDNA makes it less susceptible than genomic DNA to exonucleases that break down DNA molecules. The encapsulation of mtDNA in a two-walled organelle also enhances the mtDNA survival rate.

Various Applications for mtDNA Testing

Mitochondrial DNA variation is extensively studied in several disciplines besides forensic science. Medical scientists have linked a number of diseases to mutations in mtDNA (see Wallace et al. 1999). Evolutionary biologists examine human mtDNA sequence variation relative to other species in an effort to determine relationships. A good example of this application is the determination that Neanderthals are not the direct ancestors of modern humans based on control region sequences determined from ancient bones (Krings et al. 1997). Molecular anthropologists study differences in mtDNA sequences from various global population groups to examine questions of ancestry and migration of peoples throughout history (Relethford 2003). Hundreds of papers have been published in these fields over the past few decades. Genetic genealogists are now using mtDNA and Y-chromosome markers in an attempt to trace ancestry where paper trails run cold (Brown 2002).

In the past few years a number of interesting historical identifications have been performed with the aid of mtDNA testing. Remains from the Tomb of the Unknown Soldier associated with the Vietnam War have been identified as those of Michael Blassie (Holland & Parsons 1999). Bones discovered in Russia in 1991 were demonstrated to be those of the Tsar Nicholas II (Gill et al. 1994, Ivanov et al. 1996, Coble et al. 2009). The claims of Anna Anderson Manahan as the Russian princess Anastasia were proven false (Gill et al. 1995). The remains of

the outlaw Jesse James were linked to living relatives putting to rest a myth that he had somehow escaped death at the hands of Robert Ford (Stone et al. 2001).

Different Methods for Measuring mtDNA Variation

Over the past three decades, methods for measuring mtDNA variation have progressed in their ability to separate unrelated and closely related maternal lineages. The first studies with mtDNA in the 1980s involved low-resolution restriction fragment length polymorphism (RFLP) analysis using five or six restriction enzymes (see Richards and Macaulay 2001). Higher-resolution restriction analysis involved polymerase chain reaction (PCR) amplification of typically nine overlapping fragments followed by digestion with 12 or 14 restriction enzymes. These restriction endonucleases included *Alu*I, *Ava*II, *BamH*I, *Dde*I, *Hae*II, *Hae*III, *Hha*I, *Hinc*II, *Hinf*I, *Hpa*I, *Msp*I, *Mbo*I, *Rsa*I, and *Taq*I (Torroni et al. 1996).

Genetically different population types or haplotypes have been defined in the literature based on site losses or site gains with the various restriction enzymes. For example, haplogroup A, which is found in Asians and Native Americans, is defined by a site gain at position 663 with *Hae*III (listed as + 663 *Hae*III). Haplogroup B was initially defined as a 9 bp deletion in the intergenic region between the COII and $tRNA^{LYS}$ genes (see Table 14.2). Individuals belonging to haplogroup A may also be defined by control region polymorphisms 16223T, 16290T, and 16319A while haplogroup B individuals differ from the Anderson reference sequence at 16189C and 16217C.

In the early 1990s, DNA sequence analysis from portions of the control region came into wide acceptance. Most population data outside of the forensic community continues to be collected for only hypervariable segment I (HVS-I) spanning approximately mtDNA nucleotide positions 16024 to 16365. As will be seen below, the forensic DNA typing community has standardized on specific portions of the control region for most of the data that currently exists.

December 2000 marked the beginning of the mtDNA population genomics era with the publication of 53 entire mtGenomes from a diverse set of individuals representing populations from around the world (Ingman et al. 2000). As of December 2010 over eight thousand complete mtGenomes exist in public DNA databases (Ruiz-Pesini et al. 2004, Irwin et al. 2011).

MITOCHONDRIAL DNA SEQUENCING IN FORENSIC CASEWORK

The following section describes the methodologies used for determining the sequence contained in mitochondrial DNA. Several nice overviews of forensic mtDNA analysis have been published and may be consulted for further information on this topic (Holland & Parsons 1999, Budowle et al. 2003, Isenberg 2004, Edson et al. 2004).

Steps Involved in Obtaining mtDNA Results

The steps involved in performing mitochondrial DNA sequence comparisons are illustrated in Figure 14.3. Extraction of the mtDNA needs to be performed in a very clean

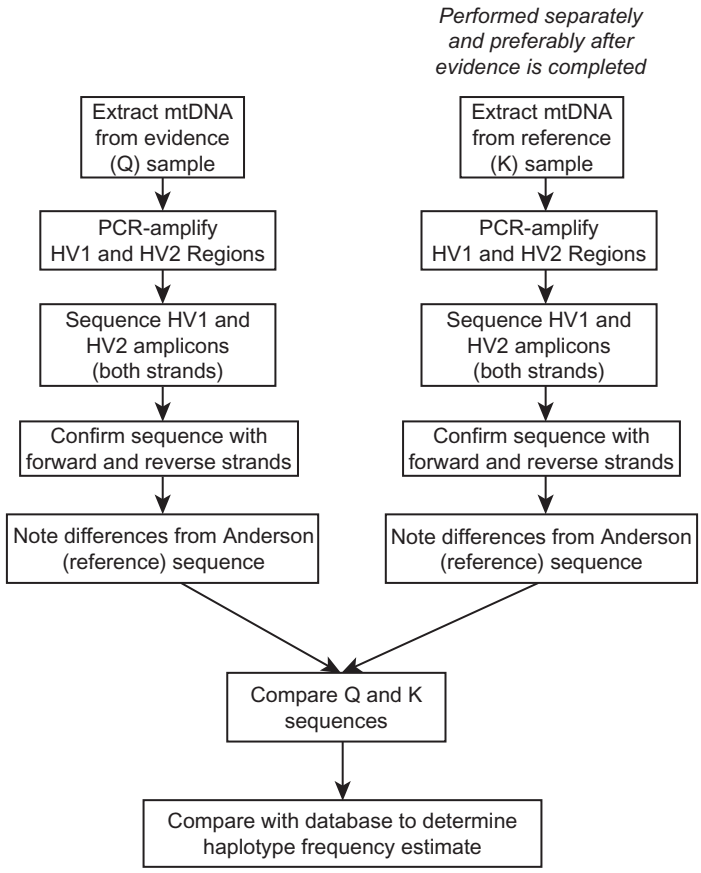

FIGURE 14.3 Process for evaluation of mtDNA samples. The evidence or question (Q) sample may come from a crime scene or a mass disaster. The reference or known (K) sample may be a maternal relative or the suspect in a criminal investigation. In a criminal investigation, the victim may also be tested and compared to the Q and K results.

laboratory environment because the high copy number per cell makes mtDNA more sensitive to contamination than nuclear DNA. Thus, it is preferable to analyze the reference samples after the evidence samples have been completely processed to avoid any potential contamination.

Mitochondrial DNA analysis is commonly performed using the Sanger sequencing chemistry (Sanger et al. 1977, Wilson et al. 1995). This DNA sequencing is performed in both the forward and reverse directions so that the complementary strands can be compared to one another for quality control purposes.

Typically laboratories report results in terms of variation compared to the rCRS. Thus, the observation of a C nucleotide at position 16126, which contains a T in the reference sequence, would be reported as 16126C. If no other nucleotide variants are reported, then it is assumed that the remaining sequence contains the same sequence as the rCRS.

Importance of a Clean Laboratory

The use of higher PCR cycle numbers (e.g., 36 or 42) and the high copy number of mtDNA per cell necessitate great care to avoid contamination. The DNA templates under investigation are often damaged so they may not be as readily amplified as even low amounts of high-quality DNA from laboratory personnel or reference samples. Reference samples from the victim, the suspect, and maternal relatives are typically available as blood stains or buccal swabs and generally contain large amounts of high-quality DNA.

Practices to reduce or minimize contamination often employed by forensic laboratories performing mtDNA testing include use of protective clothing such as disposable lab coats, frequent cleaning procedures with bleach and UV irradiation of hoods and lab bench surfaces, processing the question samples prior to the known samples, multiple glove changes during sample handling, using dedicated equipment for the mtDNA testing, and physically separating the pre- and post-amplification spaces. During an analytical procedure only one item of evidence from a case is opened at a time (Isenberg & Moore 1999).

Some laboratories even control movement of laboratory personnel between spaces. For example, a technician may not be permitted on the same day to return to a pre-amplification area after having entered a post-amplification area. Vigilance on the part of all laboratory personnel is important to keep a forensic mtDNA laboratory clean. Reagent blanks and negative controls are also run to monitor levels of exogenous DNA in reagents, laboratory environment, or instruments.

Sample Extraction for mtDNA Analysis

Mitochondrial DNA analysis typically involves materials where little DNA is present to begin with. Teeth, hair, and bones such as ribs and long bones (e.g., femur and humerus) are often materials used for mtDNA analysis in forensic cases. The mtDNA must be carefully extracted from these materials and often purified away from PCR inhibitors that can be co-extracted (Yoshii et al. 1992).

Because anthropological examination of a bone is often performed in addition to mtDNA testing, care must be taken to remove a section of the bone that will not destroy the physical features of the bone. Thus, an analyst might remove a small section from the middle of the bone without cutting all the way through the bone so that the overall length of the bone is not impacted. The same idea applies for teeth where odontological examinations are performed to aid an investigation. A tabulation of success rates for obtaining reportable mtDNA sequencing results across different skeletal materials found that ribs and femurs work best (Edson et al. 2004). A demineralization extraction protocol has dramatically improved success rates with mtDNA analysis (Loreille et al. 2007). In fact, improved DNA recoveries with new extraction protocols have enabled laboratories to obtain results with STRs or miniSTRs on bone samples where just a few years ago only mtDNA results were possible.

Special Considerations for Hair Evidence

Hair and fiber examiners can perform microscopic comparisons of hairs much more quickly than mtDNA can be analyzed. These comparisons therefore can be used as an effective screening tool to reduce the amount of evidence processed through the steps of mtDNA

sequencing. A correlation of microscopic and mitochondrial DNA hair comparisons found that the techniques can be complementary (Houck & Budowle 2002).

With hair evidence, the physical examination by a hair examiner must be performed prior to the mtDNA testing as the hair is destroyed during the extraction process. Typically for analysis of hair shafts, a tissue grinder is used to break down the keratin structure of the hair and release the mtDNA molecules (Wilson et al. 1995a). Usually 1 cm to 2 cm of hair shaft is ground up after carefully cleaning the outside of the hair (Jehaes et al. 1998). A hair digestion protocol has also been used successfully to release nuclear DNA and mtDNA for analysis (Hellman et al. 2001).

Comparisons of head, pubic, and axillary hair shafts found the highest success rate with head hair shafts (Pfeiffer et al. 1999). The addition of bovine serum albumin (BSA) (Giambernardi et al. 1998) helped reduce the PCR inhibitory effects of melanin previously noted by Yoshii et al. (1992) and Wilson et al. (1995a). A nested PCR amplification approach has successfully recovered mtDNA sequence information from as little as 33 femtograms to 330 femtograms (10 to 100 copies) of mtDNA (Allen et al. 1998).

Estimating mtDNA Quantity

Many laboratories perform a nuclear DNA quantitation assay and then estimate the amount of mtDNA present assuming a fixed ratio between nuclear and mtDNA. For example, in some of the early work 50 pg or 500 pg of DNA template would be used in an mtDNA amplification based on a nuclear quantification result from Quantiblot (Wilson et al. 1995b). Newer approaches involving real-time PCR have been published (Meissner et al. 2000, Andreasson et al. 2002, von Wurmb-Schwark et al. 2002, Alonso et al. 2004) that enable direct characterization of the number of mtDNA molecules in a cell.

By incorporating a dual real-time nuclear and mtDNA quantitation assay into their workflow, the University of Innsbruck mtDNA group were able to reduce their re-amplifications from 18% down to 7% over a two-year period of examining some 12,000 casework samples (Niederstätter et al. 2007).

PCR Amplification

PCR amplification of mtDNA is usually done with 34 to 38 cycles. Protocols for highly degraded DNA specimens even call for 42 cycles (Gabriel et al. 2001a). Sometimes excess Taq is added to overcome PCR inhibitors such as melanin (Wilson et al. 1995a). It is important to keep in mind that sensitivity is maximized with mtDNA testing because it is usually only turned to as a last resort in efforts to obtain DNA results from a sample. The higher the sensitivity of any assay, the greater the chance for contamination and thus greater care is usually required with mtDNA work than with conventional STR typing.

The most extensive mtDNA variations between individuals in the human population are found within the control region, or displacement loop (D-loop). Two regions within the D-loop known as hypervariable region I (HV1, HVI, or HVS-I) and hypervariable region II (HV2, HVII, or HVS-II) are normally examined by PCR amplification followed by sequence analysis. Approximately 610 bp are commonly evaluated—342bp from HV1 (Figure 14.4a) and 268 bp from HV2 (Figure 14.4b).

The DNA sequence for each sample between nucleotide positions 16024 and 16365 in HV1 and 73 and 340 in HV2 is determined and then compared to the Anderson or the revised Cambridge Reference Sequence (as mentioned earlier, these reference sequences are equivalent for the control region). Differences are noted and reported with the nucleotide position and the altered base. Sometimes a third hypervariable region (HVIII) is examined that is 137 bp long and spans nucleotide positions 438 to 574. Additional polymorphic sites within HVIII can sometimes help resolve indistinguishable HVI/HVII samples (Lutz et al. 2000, Bini et al. 2003).

A number of different PCR and sequencing primers have been used to generate the DNA sequence data for HV1 and HV2. Some of these primer combinations will be discussed later in the chapter. The mtDNA control region has been estimated to vary only about 1% to 2% (7 to 14 nucleotides out of the 610 bases examined are different) between unrelated individuals (Budowle et al. 1999). This variation is scattered throughout the HV1 and HV2 regions and is therefore best measured with DNA sequence analysis.

However, there are "hotspots" or hypervariable sites and regions where most of the variation is clustered (Stoneking 2000). Several methods for rapidly screening mtDNA variation have been developed that may be used for excluding samples that do not match. These methods often focus on measuring variation at the hypervariable hotspots and include using sequence-specific oligonucleotide probes (Stoneking et al. 1991), mini-sequencing (Tully et al. 1996), and denaturing gradient gel electrophoresis (Steighner et al. 1999) as well as a restriction digest assay for HV1 amplicons (Butler et al. 1998a) and a reverse dot blot or linear array assay approach (Comas et al. 1999, Gabriel et al. 2003).

In order to track work with a specific mtDNA sequence, results are commonly reported in comparison to the rCRS reference sequence. Nucleotide positions within the mtDNA molecule are numbered from 1 to 16569 using the L-strand sequence with position 1 arbitrarily coming from a restriction enzyme site found in the control region (Anderson et al. 1981). The HV1 region commonly used in forensic labs spans positions 16024 to 16365, or 342 bp, while HV2 covers positions 73 to 340, or 268 bp. Thus, use of both HV1 and HV2 provides examination of 610 bp of mtDNA sequence.

DNA Sequencing

The Sanger method for DNA sequencing was first described over 30 years ago (Sanger et al. 1977). This Nobel Prize winning sequencing technique is still widely used. The process involves the polymerase incorporation of dideoxyribonucleotide triphosphates (ddNTPs) as chain terminators followed by a separation step capable of single nucleotide resolution. There is no hydroxyl group at the 3′-end of the DNA nucleotide with a ddNTP and therefore chain growth terminates when the polymerase incorporates a ddNTP into the synthesized strand. Extendable dNTPs and ddNTP terminators are both present in the reaction mix so that some portions of the DNA molecules are extended. At the end of the sequencing reaction, a series of molecules are present that differ by one base from one another.

In the Sanger sequencing process, each DNA strand is sequenced in separate reactions with a single primer. Often either the forward or reverse PCR primers are used for this purpose. Four different colored fluorescent dyes are attached to the four different ddNTPs. Thus, ddTTP (thymine) is labeled with a red dye, ddCTP (cytosine) is labeled with a blue

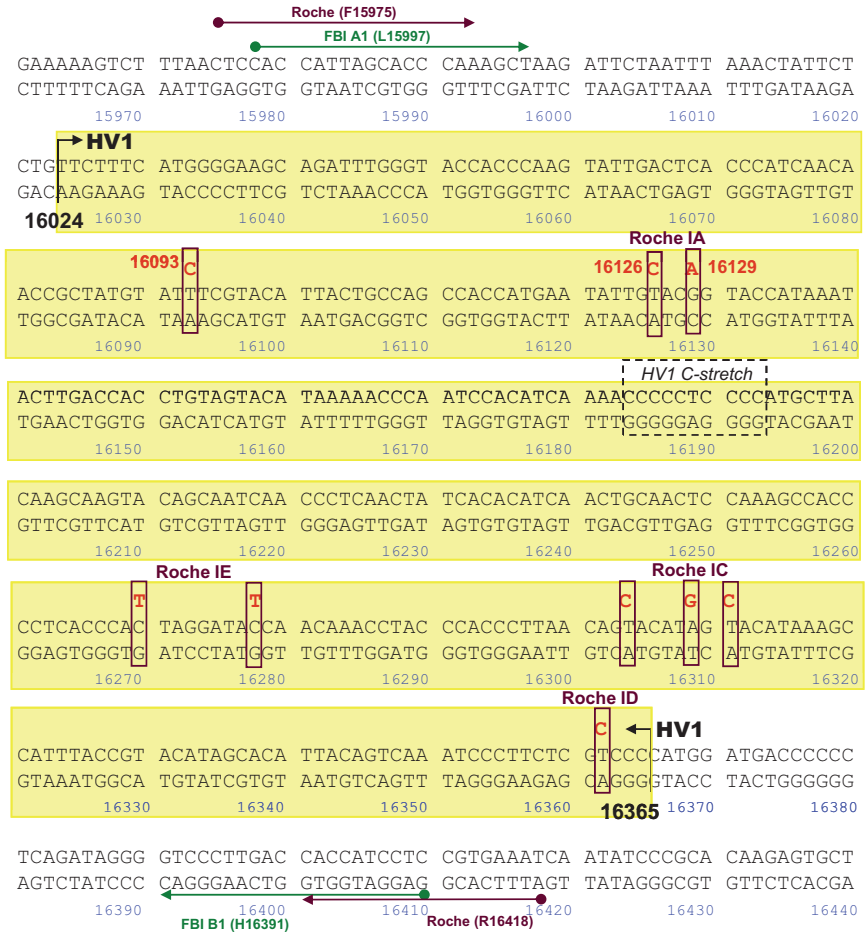

FIGURE 14.4a Annotation of the revised Cambridge Reference Sequence for the HV1 portion of the mtDNA control region with primer positions and common sequence polymorphisms examined in screening assays (see Figure 14.9).

dye, ddATP (adenine) is labeled with a green dye, and ddGTP (guanine) is labeled with a yellow dye although it is typically displayed in black for easier visualization. These are similar dyes to those used for STR detection as described in Chapter 6. Fluorescent dye labels have simplified DNA sequencing as have the widespread use of automated detection systems and capillary electrophoresis. The Human Genome Project was completed with these sequencing technologies.

The performance of DNA sequencing chemistries has progressed over the past decade from use of a simple Taq polymerase, which often had high backgrounds and poor incorporation rates for many nucleotide combinations, to the well-balanced Big Dye chemistries used today. Signal-to-noise ratios have improved with brighter dyes (Lee et al. 1997), which in turn now permit results to be obtained from less material. As little as 1 ng of mtDNA PCR product can now be used for each DNA sequencing reaction (Stewart et al. 2003).

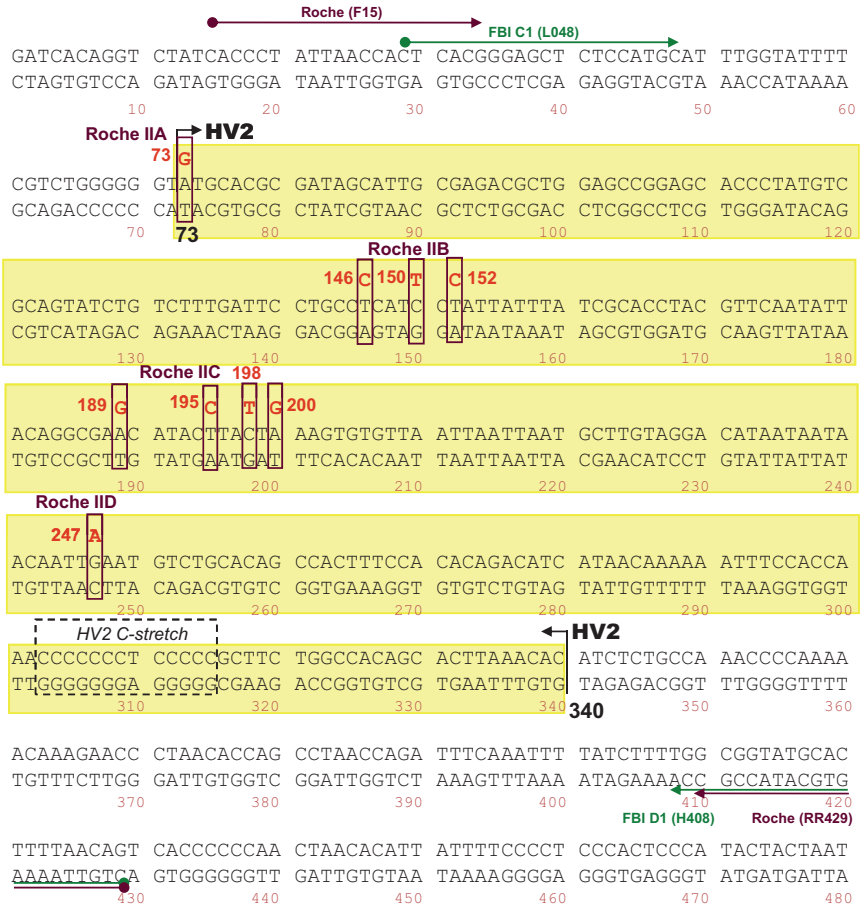

FIGURE 14.4b Annotation of the revised Cambridge Reference Sequence for the HV2 portion of the mtDNA control region with primer positions and common sequence polymorphisms examined in screening assays (see Figure 14.9).

DNA sequencing of mtDNA is usually performed with the following steps: (1) PCR amplification of the entire control region or a portion of it with various primer sets as will be explained below; (2) removal of remaining dNTPs and primers from PCR through spin filtration using a Microcon 100 filter or enzymatic digestion with shrimp alkaline phosphatase and exonuclease I (Dugan et al. 2002); (3) determination of PCR product quantity (Wilson et al. 1995a, 1995b); (4) performance of DNA sequencing reaction to incorporate fluorescent ddNTPs as described above with each reaction containing a different primer to dictate which strand is sequenced; (5) removal of unincorporated fluorescent dye terminators from the completed sequencing reaction usually through spin column filtration; (6) dilution of purified sequencing reaction products in formamide, (7) separation through a capillary electrophoresis instrument (see Chapter 6); and (8) sequence analysis of each reaction performed and interpretation of compiled sequence information as will be described below.

DNA sequencing has been reliably performed on a variety of platforms including the ABI 310, ABI 377, and ABI 3100 (Stewart et al. 2003). The multi-capillary ABI 3500 and 3730 instruments also work well. The primary difference between STR analysis and mtDNA sequencing on these multi-color fluorescence detection instruments is that a separation medium capable of single base resolution is necessary for DNA sequence analysis while it is not always needed for STR typing. Thus, the separation medium POP-6 is commonly used for DNA sequencing while POP-4, a less viscous and lower-resolution polymer, is used for STR typing.

Next-generation DNA sequencing has also been used for mtDNA sequencing (Mikkelsen et al. 2009). Currently the equipment, reagents, and data evaluation software required for this approach are quite expensive and the process temperamental and time consuming. However, the technology for next-generation DNA sequencing is rapidly developing and may eventually supplement if not displace Sanger methodologies. See Chapter 17 for further information.

Another approach to capturing mtDNA structural information is to measure base composition with mass spectometry (Hall et al. 2005, Oberacher et al. 2007, Hall et al. 2009). Although this approach cannot provide positional information on where the mass difference is located within a fragment, the indication that there is a sequence difference can be helpful in sample screening.

Primers Used for Control Region Amplification and Sequencing

PCR primers commonly used by the FBI Laboratory for mtDNA sequencing are shown as arrows in Figure 14.4 (Wilson et al. 1995b). Their primer nomenclature uses the strand corresponding to the primer (L for light and H for heavy) and the 3′ nucleotide position. Thus, primer A1 is designated as L15997 and corresponds to the light strand of the Anderson reference sequence and ends at position 15997. Note that this nomenclature system does not indicate the 5′-end of the primer and therefore can make it more difficult to determine the overall PCR product size.

Another approach to mtDNA primer nomenclature is used by the Armed Forces DNA Identification Laboratory (AFDIL). The primer positions for their primer sets (PS) I-IV are indicated in Figure 14.5. Strand designation in this case is by forward (F) and reverse (R) rather than light (L) and heavy (H). Also the 5′ nucleotide position is noted rather than the 3′ nucleotide as done by the FBI Laboratory. The AFDIL approach permits an easier determination of the overall PCR product size defined by a primer pair. It is worth noting that two of the primers in the FBI and AFDIL sets are identical even though their names are different: FBI B1 (H16391) is the same primer as AFDIL R16410 used in PSII.

Small Amplicons to Improve Amplification Success

As noted in Chapter 10 when encountering highly degraded DNA samples where the molecules have been fragmented to small sizes, the use of smaller-sized PCR products improves recovery of information from the original DNA template. This is also the case with mitochondrial DNA, and “mini-mito primer sets” have been developed to amplify smaller portions of HV1 and HV2 (Gabriel et al. 2001a, Edson et al. 2004).

The bottom portion of Figure 14.5 shows the relative position and PCR product sizes for eight mini-products ranging in size from 126 bp to 170 bp. Additional mini-mito primer

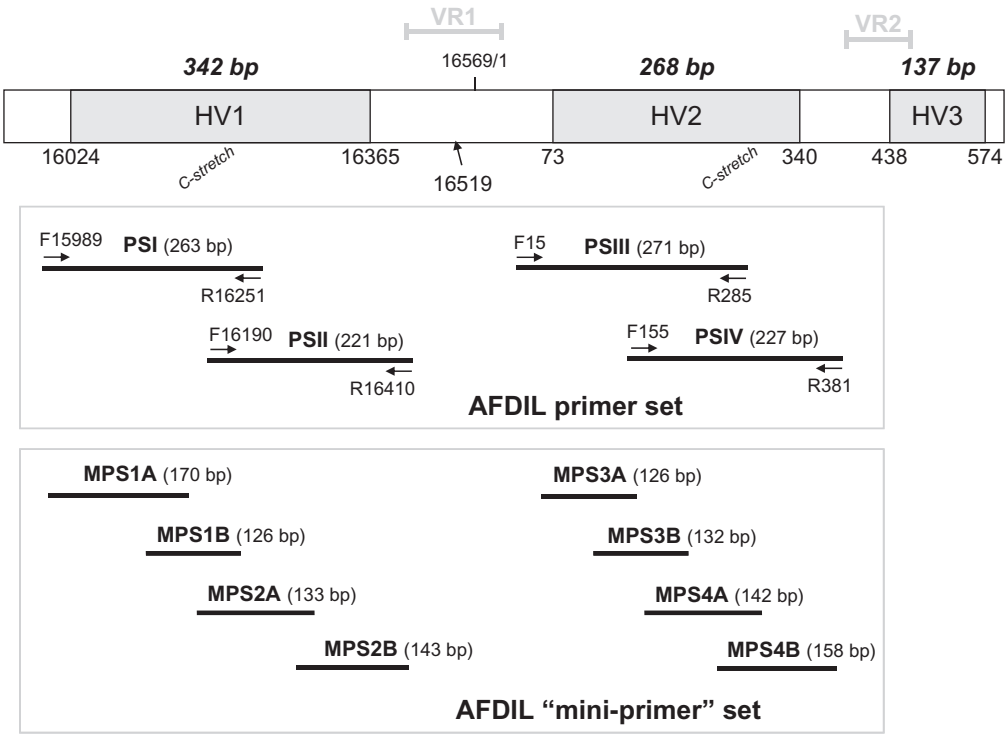

FIGURE 14.5 The three hypervariable (HV) regions of the mtDNA control region. HV1 spans nucleotide positions 16024–16365 (342 bp), HV2 spans positions 73–340 (268 bp), and HV3, which is rarely examined in forensic testing, spans positions 438–574 (137 bp). The general positions for variable regions VR1 and VR2 are noted although these are rarely used. PCR primer sets (PS) commonly used by the Armed Forces DNA Identification Laboratory (AFDIL) are illustrated. Primer nomenclature designates the 5′-nucleotide for each primer. PCR product sizes for each set of primers are noted in parentheses. The bottom section shows "mini-primer" PCR product sizes that are used with highly degraded DNA samples to enable greater recovery of sequence information (see Gabriel et al. 2001a).

sets have been developed to aid work with highly degraded DNA specimens by a group at Innsbruck, Austria (Eichmann & Parson 2008, Berger & Parson 2009) and a team in South Korea (Lee et al. 2008).

The use of mini-amplicons that overlap one another is sometimes referred to as an "ancient DNA" approach and is capable of recovering abundant DNA in a sample that might otherwise fail to produce results with a standard protocol (Gabriel et al. 2001a, Melton and Nelson 2001). This approach has been used to successfully recover information from Neanderthal bones that are thousands of years old (Krings et al. 1997, 1999).

Data Review and Editing

DNA sequencing is performed in both the forward and reverse directions so that the complementary strands can be compared to one another for quality control purposes. If it is not

possible to get a sequence from both strands, for example following a C-stretch (see below), then the same strand can be sequenced twice in separate reactions. The goal is to have at least double coverage of every nucleotide being assessed either through sequencing the top and bottom strand or sequencing the same strand twice.

The sequencing process does not always lead to beautiful data that is unambiguous for each base. Some regions, such as the C-stretches, are challenging to decipher and may not even be included in the final interpretation (Stewart et al. 2001). Further, as is discussed later in this chapter, the large copy number and relatively high mutation rate of mtDNA can lead to intra-individual sequence variability.

Sequencing chemistries and instruments have improved in recent years leading to more even peaks, better sensitivity, and less noise. However, experienced analysts must still manually review and potentially edit the software-provided base calls for each nucleotide. At present there is no publicly available software that can robustly evaluate mtDNA sequence data in a reliable and automated fashion without manual intervention.

The sequence editing process is aided by alignments from the multiple sequences generated over a region for the same sample. Computer programs such as Sequencer (GeneCodes, Ann Arbor, MI) align the forward and reverse sequencing reactions and allow the sequencing electropherograms for each reaction to be evaluated side-by-side. For casework samples that utilize smaller PCR products, the overlap between products (see Figure 14.5) permits a further measure of quality assurance in the final compiled sequence. In addition, two forensic analysts must independently examine, interpret, and edit sequence matching results as a final quality assurance measure (Isenberg 2004).

Challenges with Sequencing Beyond Polymeric C-Stretches

Note that a dotted box is found around a stretch of cytosine nucleotides in both the HV1 (Figure 14.4a) and HV2 (Figure 14.4b) regions. These regions are commonly referred to as "C-stretches." On the revised Cambridge Reference Sequence, the HV1 C-stretch spans nucleotides 16184 to 16193 with a T at position 16189. In some samples, position 16189 is a C giving rise to a stretch of 10 or more cytosines in a row (Figure 14.6). The HV2 C-stretch region spans positions 303 to 315 on the reference sequence with a T at position 310 (Figure 14.4b). This T can become a C in some samples leading to a homopolymeric C-stretch.

Unfortunately, this homopolymeric stretch of cytosines creates problems for polymerases as they synthesize a complementary strand to the mtDNA template present in the reaction. Length heteroplasmy in HV1 between positions 16184 and 16193 can result in C-stretch lengths ranging from 8 to 14 cytosine residues (Bendall & Sykes 1995). Length heteroplasmy likely results from replication slippage after a T to C transition has occurred at position 16189. The mixture of length variants may already be present in the original DNA or generated in the sequencing reaction itself. Regardless of the source of the length variants, the impact of a 16189 T-to-C transition on sequencing results downstream of the C-stretch region can be seen in Figure 14.6b.

A similar situation occurs with the HV2 C-stretch region when insertions of cytosines occur in the 303 to 310 area or a transition of T-to-C occurs at position 310 (Stewart et al. 2001). The presence of intra-individual variation in the number of cytosines observed when multiple hairs were tested from the same individual has led to the decision to not call an

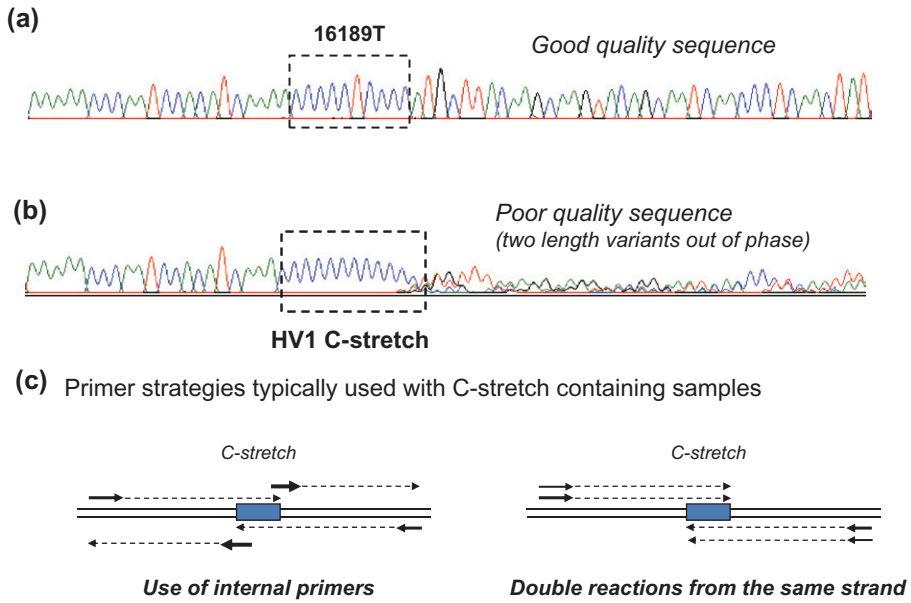

FIGURE 14.6 Comparison of a sample with (a) 16189T (no HV1 C-stretch) to (b) one with the C-stretch. Notice how the sequence quality quickly drops after the string of cytosine residues due to the presence of two or more length variants that creates a situation where the extension products are out of phase or register with one another. Different primer combinations are typically used on samples containing a C-stretch as illustrated in (c) to recover sequence information from both strands or to provide a double read of the same strand.

exclusion based solely on differences in the HV2 C-stretch region (Stewart et al. 2001). The issue of heteroplasmy and intra-individual variation will be discussed in more detail later in this chapter.

The ability to rapidly screen for the C-stretch prior to sequencing is advantageous and can be performed by noting the presence of extra heteroduplex peaks in quality control analyses of HV1 PCR products (Butler et al. 1998a). In the event that the C-stretch is present in a sample, different sequencing primers may be used to obtain reliable mtDNA sequence information downstream of the homopolymeric stretches (Rasmussen et al. 2002). For example, the FBI A4/B4 primer set (L16209 and H16164) shown in Figure 14.4a can be used on individuals possessing the HV1 C-stretch in order to recover sequence information from both sides of the homopolymeric stretch of cytosines (Wilson 1997). Alternatively the same strand may be examined twice in separate sequencing reactions to provide double coverage of all nucleotides (Figure 14.6c).

Use of Positive and Negative Controls

As noted by Melton and Nelson (2001) the two primary goals in mtDNA testing are (1) to protect the integrity of the evidence by preventing contamination at any stage in the testing and (2) to collect the maximum amount of available mtDNA data inherent to any sample.

Control samples that are processed in parallel with evidentiary samples through each step of the process serve to monitor performance and assess one's success with the two goals noted above.

Contamination assessment is performed with reagent blanks and negative controls. Reagent blanks monitor contamination from extraction to final sequence analysis while negative controls monitor contamination from amplification to final sequence analysis (SWGDAM 2003). All of the procedures performed on a sample are also performed on the reagent or extraction blank with the exception of adding DNA. Negative controls or amplification blanks are introduced at the PCR amplification step and use the same reagents as the sample with sterile water in place of the DNA template. If the reagent blank and/or the negative control associated with a particular amplification results in a sequence that is the same as that of the sample, all data for the sample must be rejected (Isenberg 2004). The analysis must then be repeated beginning with the re-amplification of the sample in question.

Reagent blank contamination is sometimes observed in spite of great efforts to keep the laboratory environment clean. Since mtDNA analysis is a very sensitive technique, the presence of low-level contamination is not uncommon (Isenberg 2004). For example, Mitotyping Technologies reported that reagent blanks resulted in amplification products in 29 of 1218 (2.4%) of PCR reactions performed in casework over a two-year period of time (Melton & Nelson 2001). These contaminants did not match a staff member's type or the type of the recently handled sample. This suggests sporadic contamination of disposable tips or PCR tubes during manufacture or packaging. This type of contamination is not uncommon when working with low-copy-number DNA as noted in Chapter 11.

If contamination is observed with either the reagent blank or the negative control, results from the unknown sample being run in parallel do not always have to be disregarded. Research with artificial sample mixtures has demonstrated that a threshold of background contamination can be set for still obtaining reliable sequence data. For example, the FBI Laboratory has established a 10:1 rule where any contamination seen in a reagent blank or negative control during post-PCR analysis must be less than one-tenth the amount of the sample being processed (Wilson et al. 1995a, 1995b). This sample-to-contamination ratio determination is possible due to the PCR product quantification analysis performed in their procedure (Butler et al. 1994). A more recent study demonstrated that the 10:1 rule is conservative and reliable (Stewart et al. 2003).

A positive control is a sample of known mtDNA sequence that serves to demonstrate that amplification and sequencing reaction components are working properly. This positive control is typically an extracted DNA sample that is processed through the steps of amplification, sequencing, and data analysis. For example, the FBI Laboratory uses the HL-60 cell line as a positive control (Levin et al. 2003).

Interlaboratory Studies

Interlaboratory studies in which laboratories perform testing on the same sample are valuable for demonstrating that a technique is reliable (see Chapter 7). As of 2010, no manufacturers supply commercially available kits for the entire process of mtDNA sequencing, such as are available for STR typing. Thus, a number of different methods exist for mtDNA testing without a single universal protocol.

A number of interlaboratory studies involving mtDNA sequencing have been conducted and have demonstrated that the same results can be successfully obtained in multiple laboratories using different protocols (Carracedo et al. 1998, Alonso et al. 2002, Prieto et al. 2003, Parsons et al. 2004, Tully et al. 2004).

Certified Reference Materials for mtDNA Sequence Analysis

Certified reference materials along with positive controls serve to demonstrate that mtDNA sequence analysis is being performed appropriately (Szibor et al. 2003a). The U.S. National Institute of Standards and Technology (NIST) has developed two Standard Reference Materials (SRMs) to aid in confirming sequencing results with mtDNA (Levin et al. 1999, Levin et al. 2003). Information is available for the entire mtGenome on the cell line HL-60 (SRM 2392-I) and on three samples (SRM 2392). The certificates for these reference materials were updated in 2007 and 2009, respectively.

The FBI Revised Quality Assurance Standards require that U.S. laboratories run a NIST SRM or material traceable to a NIST SRM at least once a year or whenever a protocol is changed to help verify that DNA sequencing and interpretation are being performed accurately (see Appendix 4, Standard 9.5.5).

INTERPRETATION OF mtDNA RESULTS

Following completion of mtDNA sequence analysis, as outlined in Figure 14.3, results from the edited and reviewed sequences for a question (Q) and a known (K) sample are compared as illustrated in Figure 14.7 for a portion of HV1. All 610 nucleotides (positions 16024–16365 and 73–340) are normally evaluated between samples being compared.

(a) mtDNA Sequences Aligned with rCRS (positions 16071-16140)

```
                16090       16100       16110       16120       16130       16140
rCRS  ACCGCTATGT ATTTCGTACA TTACTGCCAG CCACCATGAA TATTGTACGG TACCATAAAT
Q     ACCGCTATGT ATCTCGTACA TTACTGCCAG CCACCATGAA TATTGTACAG TACCATAAAT
K     ACCGCTATGT ATCTCGTACA TTACTGCCAG CCACCATGAA TATTGTACAG TACCATAAAT
```

(b) Reporting Format with Differences from rCRS

Sample Q	Sample K
16093C	16093C
16129A	16129A

FIGURE 14.7 (a) Comparison of sequence alignments for hypothetical Q and K samples with (b) conversion to the revised Cambridge Reference Sequence (rCRS) differences for reporting purposes.

Based on the Q-K comparison, mtDNA sequence results can generally be grouped into three categories: exclusion, inconclusive, or failure to exclude. The SWGDAM guidelines for mtDNA interpretation makes the following recommendations (SWGDAM 2003):

- *Exclusion* – if there are two or more nucleotide differences between the questioned and known samples, the samples can be excluded as originating from the same person or maternal lineage.
- *Inconclusive* – if there is one nucleotide difference between the questioned and known samples, the result will be inconclusive.
- *Cannot Exclude (Failure to Exclude)* – if the sequences from questioned and known samples under comparison have a common base at each position or a common length variant in the HV2 C-stretch, the samples cannot be excluded as originating from the same person or maternal lineage.

A common base is defined as a shared base in the case of ambiguity (e.g., heteroplasmy) in the sequence (Isenberg 2004). For example, if one sequence possesses heteroplasmy at a site and another does not (see Figure 14.8), then they cannot be excluded from one another. A length variant alone especially in the HV2 homopolymeric C-stretch cannot be used to support an interpretation of exclusion (Stewart et al. 2001, SWGDAM 2003). Several examples are provided in Table 14.4 with their respective interpretations based on the SWGDAM guidelines.

The reason that a single base difference is classified in terms of an "inconclusive result" is that mutations have been observed between mother and children (Parsons et al. 1997). For example, if a maternal relative is used for a reference sample, the possibility of a single base

TABLE 14.4 Example mtDNA Sequences and Interpretations for Known (K) and Question (Q) Sample Pairs (Adapted from Isenberg 2004).

Sequence Results	Observations	Interpretation
Q TATTGTACGG **K** TATTGTACGG	Sequences are fully concordant with common bases at every position	Cannot Exclude
Q TATTGCACAG **K** TATTGTACGG	Sequences differ at two positions	Exclusion
Q TATTNTACGG **K** TATTGTACGG	A single unspecified base in one of the sequences; common base at every position	Cannot Exclude
Q TATTNTACGG **K** TATTGTACNG	Ambiguous bases in both sequences at different positions; common base at every position	Cannot Exclude
Q TATTGTACA/GG **K** TATTGTAC G G	Heteroplasmic mixture at a position in one sample that is not present in the other; common base at every position (G in both Q and K)	Cannot Exclude
Q TATTGTACA/GG **K** TATTGTACA/GG	Heteroplasmic mixture at the same site in both sequences; common base at every position	Cannot Exclude
Q TATTGCACGG **K** TATTGTACGG	Sequences identical at every position except one; no indication of heteroplasmy	Inconclusive

difference may exist between two samples that are in fact maternally related. Often additional samples, usually more reference samples, are run if an inconclusive result is obtained in an attempt to clarify the interpretation (Wilson et al. 1997). Hairs from an individual might be pooled in an attempt to detect heteroplasmy (Isenberg 2004).

More recently, Parson and Bandelt (2007) have offered some extended guidelines for mtDNA analysis and interpretation.

Reporting Statistics

When "failure to exclude" is the interpretation for reference and evidence samples, then a statistical estimate of the significance of a match is needed. Mitochondrial DNA is inherited in its entirety from our mother without recombination (discussed later in the chapter). Therefore individual nucleotide positions are inherited in a block and must be treated as a single locus haplotype, the same as with Y-chromosome information discussed in Chapter 13. The product rule applied to independently segregating STR loci found on separate chromosomes cannot be used with mtDNA polymorphisms.

As was previously noted with Y-STR interpretation, the current practice of conveying the rarity of a mtDNA type among unrelated individuals involves counting the number of times a particular haplotype (sequence) is seen in a database (Wilson et al. 1993, Budowle et al. 1999). This approach is commonly referred to as the "counting method" and depends entirely on the number of samples present in the database that is searched. Thus, the larger the number of unrelated individuals in the database, the better the statistics will be for a random-match frequency estimate.

The true population frequencies for around 60% of mtDNA sequences are not presently known because they occur only a single time in a database (Isenberg 2004). Based on available population information, confidence intervals can be used to estimate the upper and lower bounds of a frequency calculation (Holland & Parsons 1999, Tully et al. 2001). An example is worked in D.N.A. Box 14.1.

One of the challenges with rare mtDNA haplotypes is how to express the weight of evidence when a particular type has not been seen before in the database. Charles Brenner has developed an approach for handling this situation (Brenner 2010). Although other methods, such as likelihood ratios, may be used for estimating the weight of evidence, it is important to keep in mind that mtDNA can never have the power of discrimination that an autosomal STR marker can since its inheritance is uniparental.

Reporting Differences to the Revised Cambridge Reference Sequence

For reporting purposes, sequences are listed in a minimum data format as differences relative to the rCRS. When differences are observed, the nucleotide position is cited followed by the base present at that site. For example in Figure 14.7a, differences are observed at positions 16093 and 16129 and are noted in Figure 14.7b in their minimum data format at 16093C and 16129A. In this format, all other nucleotides are assumed to be identical to the revised Cambridge Reference Sequence. Bases that cannot be unambiguously determined are usually coded N. At confirmed positions of ambiguity (e.g., sequence heteroplasmy), the

D.N.A. BOX 14.1

CALCULATION OF mtDNA PROFILE FREQUENCY ESTIMATES USING THE COUNTING METHOD

In cases where an mtDNA profile is observed a particular number of times (X) in a database containing N profiles, its frequency (p) can be calculated as follows:

$$p = X/N$$

A 95% upper bound confidence interval can be placed on the profile's frequency using:

$$p + 1.96\sqrt{\frac{(p)(1-p)}{N}}$$

In cases where the profile has not been observed in a database, the 95% upper bound on the confidence interval is

$$1 - \alpha^{1/N} = 1 - (0.05)^{1/N}$$

where $\alpha = 0.05$ is the confidence coefficient and N is the number of individuals in the database.

For example, the mtDNA type 16129A, 263G, 309d, 315.1C occurs twice in 1148 African-American profiles, twice in 1655 Caucasian profiles, and not at all in 686 Hispanic profiles when searched against the mtDNA Population Database (Monson et al. 2002). Using the equations above, calculations for the rarity of this profile in the respective sample sets are as follows:

For African-Americans: $p = 2/1148 + 1.96\ [(2/1148)(1 - (2/1148))/1148]^{1/2} = 0.0017 + 0.002 = 0.004 = 0.40\%$

For Caucasians: $p = 2/1655 + 1.96\ [(2/1655)(1 - (2/1655))/1655]^{1/2} = 0.0012 + 0.0017 = 0.0029 = 0.29\%$

For Hispanics: $1 - (0.05)^{1/686} = 1 - 0.9956 = 0.0044 = 0.44\%$

These calculations demonstrate that the statistical weight can be similar whether or not a match is found to a few previously observed samples in a database.

As shown in D.N.A. Box 13.2, the Clopper-Pearson method (Clopper & Pearson 1934) may also be used to provide a conservative estimate for the upper 95% confidence interval.

Sources:

Clopper, C.J., & Pearson, E.S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika, *26, 404–413.*

Evett, I.W., & Weir, B.S. (1998). Interpreting DNA Evidence. *Sunderland, MA: Sinauer Associates, Inc., p. 142.*

Monson, K.L., et al. (2002). The mtDNA population database: an integrated software and database resource. Forensic Science Communications, *4(2). Available at <http://www2.fbi.gov/hq/lab/fsc/backissu/april2002/miller1.htm>.*

Tully, G., et al. (2001). Considerations by the European DNA profiling (EDNAP) group on the working practices, nomenclature and interpretation of mitochondrial DNA profiles. Forensic Science International, *124, 83–91.*

International Union of Pure and Applied Chemistry (IUPAC) codes should be used, such as A/G = R and C/T = Y (SWGDAM 2003).

Insertions in a DNA sequence relative to the rCRS are described by noting the site immediately 5′ to the insertion as compared to the rCRS followed by a point and a "1" (for the first insertion), a "2" (if there is a second insertion), and so on, and then by the nucleotide

that is inserted (Isenberg 2004). For example, 315.1C is a common observation where six Cs are observed following the T at position 310 in the rCRS. The rCRS contains five Cs in positions 311 through 315 (Andrews et al. 1999). Therefore, the notation 315.1C describes the presence of an extra C as an insertion (".1C") prior to position 316.

Deletions are noted by a dash ("−") or a "D," "d" or "del" following the nucleotide position where the deletion was observed relative to the rCRS (e.g., 309D, 309-, or 309del). Some insertion and deletion combinations can lead to multiple possibilities for reporting a result in terms of differences from the reference sequence. Therefore, recommendations have been made for consistent treatment of length variants as will be described in the next section.

Nomenclature Issues

Ambiguities with respect to mtDNA nomenclature can result in different analysts describing the same sample differently although they agree on the nucleotide sequence. Likewise population databases could have multiple entries for the same mtDNA haplotype preventing an accurate estimate for the frequency of a particular type. Thus, standardization in designation of mtDNA sequences is important to generate comparable data that can easily be shared among laboratories.

Length variants present a challenge when alignments are made between a sample of interest and the Cambridge Reference Sequence. Treatments of insertions and deletions (gaps) can vary between laboratories causing some laboratories to code the same sequence differently. Mark Wilson and colleagues at the FBI Laboratory have made a number of recommendations to enable consistent treatment of length variants (Wilson et al. 2002a, 2002b). Three primary recommendations were made: (1) characterize profiles using the least number of differences from the reference sequence; (2) if there is more than one way to maintain the same number of differences with respect to the reference sequence, differences should be prioritized in the following manner: (a) insertions/deletions (indels), (b) transitions, and (c) transversions; (3) insertions and deletions should be placed 3′ with respect to the light strand. Insertions and deletions should be combined in situations where the same number of differences to the reference sequence is maintained. These recommendations are hierarchical; that is recommendation (1) should take precedence over recommendations (2) and (3). A total of 41 specific examples are provided to demonstrate the need for consistent treatment of length variants in mtDNA sequence analysis and reporting (Wilson et al. 2002a, 2002b).

Some groups prefer to use a phylogenetic approach to expressing the nomenclature of a mtDNA sequence (Brandstätter et al. 2004, 2007). In the future, string searches that utilize the full mtDNA sequence (Irwin et al. 2007, Röck et al. 2011) will remove the ambiguity and potential mismatches that can occur when reducing sequences to differences from a reference sequence using either hierarchical rules or phylogenetic approaches.

ISSUES IMPACTING INTERPRETATION

In this section, several issues that can arise when considering mtDNA evidence particularly in courts of law are further elaborated upon (see Walker 2003). A National Institute of Justice-funded study also found that there can be confusion and misperceptions by jurors in terms of the strength of the evidence when mtDNA data is presented in court (Dann et al. 2004).

Heteroplasmy

Heteroplasmy is the presence of more than one mtDNA type in an individual (Melton 2004). Two or more mtDNA populations may occur between cells in an individual, within a single cell, or within a single mitochondrion. It is now thought that all individuals are heteroplasmic at some level—many below the limits of detection in DNA sequence analysis (Comas et al. 1995, Bendall et al. 1996, Steighner et al. 1999, Tully et al. 2000). It is highly unlikely that millions of mtDNA molecules scattered throughout an individual's cells are completely identical given that regions of the mtGenome have been reported to evolve at 6 to 17 times the rate of single-copy nuclear genes (see Brown et al. 1979, Wallace et al. 1987, Tully 1999). Consider that whereas only a single copy of each nuclear chromosome is present in an egg, there are approximately 100,000 copies of the mtDNA genome present (Chen et al. 1995). Thus, for the transmission of a mtDNA mutation to become detectable it must spread to an appreciable frequency among a cell's mtDNA molecules.

Heteroplasmy may be observed in several ways: (1) individuals may have more than one mtDNA type in a single tissue; (2) individuals may exhibit one mtDNA type in one tissue and a different type in another tissue; and/or (3) individuals may be heteroplasmic in one tissue sample and homoplasmic in another tissue sample (Carracedo et al. 2000). In fact, heteroplasmy has been reported inside a single mitochondrion isolated with optical tweezers (Pfugradt et al. 2010, Reiner et al. 2010). Given that heteroplasmy happens, interpretation guidelines must take into account how to handle differences between known and questioned samples.

Both sequence and length heteroplasmy have been reported in the literature (Bendall & Sykes 1995, Bendall et al. 1996, Melton 2004). Length heteroplasmies often occur around the homopolymeric C-stretches in HV1 at positions 16184 to 16193 and HV2 at positions 303 to 310 (Stewart et al. 2001, Parson & Bandelt 2007) (see Figure 14.4). Sequence heteroplasmy is typically detected by the presence of two nucleotides at a single site, which show up as overlapping peaks in a sequence electropherogram (Figure 14.8).

Heteroplasmy at two sites in the same individual, a condition known as "triplasmy," has been reported (Tully et al. 2000), but occurs at lower frequencies than single-site heteroplasmy. Since it is rare to find more than one heteroplasmic position in the 610 nucleotides sequenced for HV1 and HV2, a report of as many as six heteroplasmic sites in an individual mtDNA sequence (Grzybowski 2000) raised suspicions about the sequencing strategy used. The Grybowski study has been criticized as possibly containing contamination due to the excessive number of amplification cycles used (Budowle et al. 2002a, Brandstätter & Parson 2003). A reanalysis of the same samples used in the original Grybowski study with a direct rather than a nested PCR approach resulted in a reduction in the reported number of samples with heteroplasmic positions (Grzybowski et al. 2003).

One of the major challenges of heteroplasmic samples is that the ratio of bases may not stay the same across different tissues, such as blood and hair or between multiple hairs (Sullivan et al. 1997, Wilson et al. 1997, Sekiguchi et al. 2003). Some mtDNA protocols now recommend sequencing multiple hairs from an individual in order to confirm heteroplasmy.

Hotspots for heteroplasmy include the following positions in HV1: 16093, 16129, 16153, 16189, 16192, 16293, 16309, and 16337 (Stoneking 2000, Tully et al. 2000, Brandstätter & Parson 2003) and 72, 152, 189, 207, and 279 in HV2 (Calloway et al. 2000, Melton & Nelson

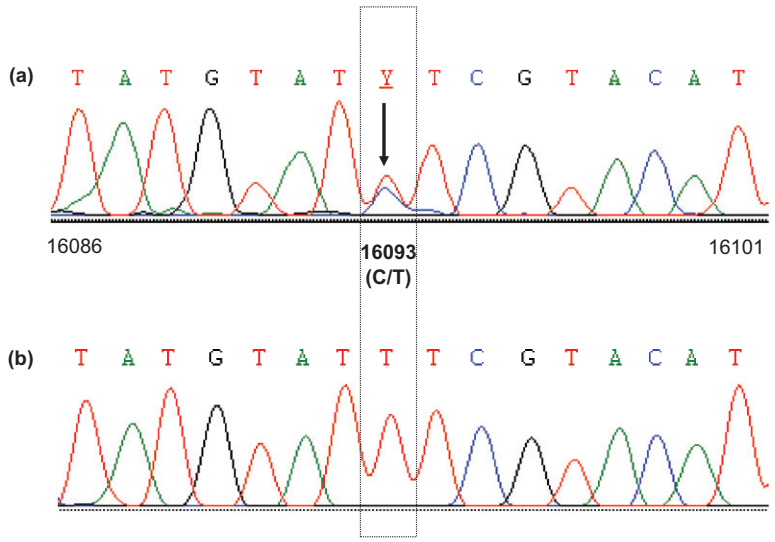

FIGURE 14.8 (a) Sequence heteroplasmy at position 16093 possessing both C and T nucleotides compared to (b) the same region (positions 16086–16101) on a different sample containing only a T at position 16093.

2001). One study found that the frequency of heteroplasmy can differ across tissue types, with muscle tissue being the highest, and was statistically significant across different age groups suggesting that heteroplasmy increases with age (Calloway et al. 2000). Heteroplasmy has also been reported to remain stable over time in the same individuals and thus be inherited rather than age related (Lagerström-Fermér et al. 2001). While heteroplasmy can sometimes complicate the interpretation of mtDNA results, the presence of heteroplasmy at identical sites can improve the probability of a match, such as seen in the Romanov study (Ivanov et al. 1996).

Sample Mixtures

A major advantage of mtDNA in terms of sequencing is that it is haploid and therefore only a single type exists (barring detectable heteroplasmy) for analysis. However, mixed samples from more than one biological source are commonly encountered in forensic settings. Generally speaking attempts are not made to decipher samples containing a mixture due to the complexity of the sequencing signals that could arise. Peak height ratios for two different bases cannot be used for reliable quantification of the two components because incorporation rates are not always even. Thus, the ratio of an A:G mixed base might be 50:50 at a particular position but when the complementary strand is sequenced a 70:30 or 80:20 ratio for the T:C bases might be observed because the polymerase incorporates the fluorescently-labeled ddTTP and ddCTP with different efficiencies than the A and G dideoxynucleotides.

If three or more sites within the 610 bases evaluated across HV1 and HV2 are found to possess multiple nucleotides at a position (i.e., sequence heteroplasmy), then the sample can usually be considered a mixture—either by contamination or from the original source material. Presently mixture interpretation is not attempted in forensic laboratories performing routine casework.

Some researchers are pursuing efforts to resolve mtDNA mixtures through cloning and sequencing the resulting HV1/HV2 regions from individual colonies (Bever et al. 2003, Walker et al. 2004). Theoretically, each individual colony produced during the process of cloning corresponds to the control region from a single individual or a single component of heteroplasmy. Interpretation of mixtures is being attempted with statistical analysis from multiple clones. A number of pitfalls exist with this approach including the possibility of overestimating the number of contributors due to the occurrence of heteroplasmic mitochondria. The number of contributors will be underestimated if individuals are closely related and members of the same mtDNA haplogroup (Walker et al. 2004). Denaturing HPLC has also been used for separating mtDNA amplicon mixtures (Danielson et al. 2007) as has a mismatch primer-induced restriction site analysis method (Szibor et al. 2003b).

Nuclear Pseudogenes

Segments of the mtGenome are present in the human nuclear genome (Collura & Stewart 1995, Zischler et al. 1995, Wallace et al. 1997). These "molecular fossils" or pseudogenes are rare events caused by migration and integration of a portion of the mtGenome into nuclear DNA and are sometimes referred to as "*numts*" (Lopez et al. 1994). Zischler et al. (1995) reported that human chromosome 11 carries a portion of the mtDNA control region that reflects an ancient genetic transposition from the mitochondrion to the nuclear genome. This element differs from typical modern mtDNA sequences by approximately 7.5% and has not created problems with regular forensic casework (Morgan et al. 1998).

Molecular fossils potentially complicate mtDNA human identity testing if they rather than the intended mtDNA target are amplified when a high number of PCR cycles are used to try to tease out mtDNA sequence information from a particularly difficult sample (Morgan et al. 1998). Under unique circumstances, nuclear pseudogenes could contaminate the true mtDNA sequence. Such was likely the case with the high degree of heteroplasmy reported on some hair samples amplified with a nested PCR approach involving a cumulative number of 60 cycles (Grzybowski 2000, Budowle et al. 2002a, Brandstätter & Parson 2003). However, with primer sets commonly used in forensic mtDNA testing and a direct PCR with fewer than 40 cycles, nuclear DNA sequences that are similar to mtDNA rarely cause a problem because their initial copy number is so much lower than that of mtDNA.

Possibilities of Recombination or Paternal Leakage

Several years ago three papers were published suggesting the possibility of recombination in mtDNA or inheritance from the paternal rather than the maternal line (Hagelberg et al. 1999, Eyre-Walker et al. 1999, Awadalla et al. 1999). Paternal inheritance of mtDNA has been reported in mice (Gyllensten et al. 1991). The Hagelberg and Eyre-Walker papers

created quite a stir in the mtDNA forensic and population genetic circles (Macaulay et al. 1999, Parsons & Irwin 2000, Kivisild & Villems 2000, Jorde & Bamshad 2000, Kumar et al. 2000). Hagelberg and co-workers later retracted their paper due to problems with the data (Hagelberg et al. 2000). Since there really appears to be no direct evidence to support either recombination within or between mtGenomes, this issue has been laid to rest for most scientists in the field (see Ingman et al. 2000, Elson et al. 2001, Wiuf 2001, Herrnstadt et al. 2002).

However, there has been a single report published of the transmission of a paternal human mtDNA type in skeletal muscle (Schwartz & Vissing 2002). This paternal haplotype was not found in any other tissues though. Several additional studies with individuals having a similar muscle disease failed to find any evidence of paternal transmission of mtDNA (Johns 2003, Filosto et al. 2003, Taylor et al. 2003). With tens of thousands of mtDNA samples confirming the maternal inheritance pattern established over three decades ago (Giles et al. 1980), it is safe to conclude that the central dogma of maternal inheritance for mtDNA is sound.

Size of mtDNA Population Database and the Quality of Information

There are now population databases with thousands of mtDNA profiles in them. The availability of population data for the HV1/HV2 regions that are sequenced in forensic mtDNA analysis will be discussed in more detail later in the chapter.

Most Common Types

One of the biggest weaknesses of mtDNA analysis is that some haplotypes are rather common in various population groups. For example, in the FBI mtDNA Population Database of 1655 Caucasians there are 131 individuals who match at 263G, 315.1C and 264 additional profiles that have only a single difference. Thus, 395 out of 1655 (23.9%) of the Caucasian database would not be able to be excluded if a sample was observed with this common mtDNA type!

However, additional sequence information from polymorphic sites around the entire mtGenome have been characterized to help better resolve many of these most common types (Parsons & Coble 2001, Coble et al. 2004). Using this information, assays have been developed to help subdivide several of the most common Caucasian, African American, and Hispanic mtDNA types (Parsons 2006).

LABORATORIES PERFORMING mtDNA TESTING

The first efforts in mtDNA sequence analysis with a forensic applications focus were performed by the Forensic Science Service in England (Sullivan et al. 1991, Hopgood et al. 1992, Sullivan et al. 1992) although the FBI Laboratory had thought about its use in the late 1980s (Budowle et al. 1990). Today there are a number of laboratories internationally that perform mtDNA testing. One of the most widely respected is Walther Parson's lab at the University of Innsbruck in Austria. The EMPOP database described later in this chapter was

created and is maintained by this group (Parson & Dür 2007). Within the United States, the Armed Forces DNA Identification Laboratory and the FBI Laboratory have led the efforts in mtDNA analysis but in slightly different arenas.

The Armed Forces DNA Identification Laboratory (AFDIL) is located in Rockville, Maryland and is charged with identifying the remains of military personnel (Holland et al. 1993). Bones recovered from Vietnam, Korea, and World War II operations have been successfully analyzed with mtDNA (Holland et al. 1995, Holland & Parsons 1999). AFDIL also aids mass disaster victim identification programs including those necessitated by U.S. airline crashes (see Chapter 9) and has helped solve historical puzzles such as identifying remains from the Tomb of the Unknown Soldier (Holland & Parsons 1999) and the Romanov family (Ivanov et al. 1996, Coble et al. 2009).

The FBI Laboratory focuses on the use of forensic evidence including mtDNA in criminal investigations. Until they were further subdivided in 2009, two DNA units existed within the FBI Laboratory: DNA Unit I, which focused exclusively on nuclear DNA, and DNA Unit II, which performed mtDNA analysis and aided missing persons investigations.

The FBI Laboratory first explored the feasibility of using mtDNA in human identity applications in the late 1980s (Budowle et al. 1990) and aggressively began researching analysis methods in 1992. The FBI Laboratory DNA Unit II, now called the Mitochondrial DNA Unit, has conducted mitochondrial DNA casework since June 1996. Their first case involving court testimony came in August 1996 with the State of Tennessee versus Paul William Ware, which involved mtDNA analysis of a single pubic hair found in the throat of a young victim that matched the defendant who was subsequently convicted (Marchi & Pasacreta 1997). Much of the mtDNA evidence processed by the FBI involves shed hairs to aid criminal and counterterrorism investigations.

In 2005, four regional FBI-funded mtDNA laboratories became operational to conduct mtDNA casework as an extension of the FBI's own operations. The four original regional mtDNA labs were the Arizona Department of Public Safety (Phoenix, Arizona), the Connecticut State Police (Meriden, Connecticut), the Minnesota Bureau of Criminal Apprehension (St. Paul, Minnesota), and the New Jersey State Police (Trenton, New Jersey). As of early 2011, Arizona, Minnesota, and New Jersey are the regional FBI mtDNA labs. Each of these satellite laboratories was originally designed to be able to analyze approximately 120 cases per year. Taken collectively, the goal of creating these regional mtDNA laboratories was to double the FBI's capacity to provide mtDNA analysis for the criminal justice system.

Several private laboratories in the United States have validated mtDNA procedures and offer mtDNA testing for a fee. These laboratories include Mitotyping Technologies, LLC (State College, Pennsylvania), Bode Technology Group (Lorton, Virginia), Orchid Cellmark (Dallas, Texas), and Laboratory Corporation of America (Research Triangle Park, North Carolina). These laboratories typically charge around $2000 per sample for mtDNA testing in order to sequence the 610 nucleotides in HV1 and HV2. The University of North Texas Center for Human Identification (Ft. Worth, Texas) is funded by the National Institute of Justice to perform mtDNA sequence analysis in aiding missing persons work (see Chapter 9).

Mitotyping Technologies reported processing 105 cases between February 1999 and February 2001 (Melton & Nelson 2001). These cases involved 199 questioned items of which 130 were hairs. A total of 137 known reference samples were also processed including 111 that were in the form of blood. Only 17 of their 199 questioned samples failed to yield any

mtDNA amplification products. Length heteroplasmy was observed 15 times in the HV1 C-stretch region and 77 times in the HV2 C-stretch region with 17 samples having both HV1 and HV2 length heteroplasmy. Sequence site heteroplasmy was reported 19 times mostly at positions 16093 but also at nucleotide positions 16166, 16286, 72, 152, 189, 207, and 279. In 57 out of 105 cases (54.3%), the known reference sample could not be excluded as donor of a biological sample.

Mitotyping has also published an evaluation of their success with 691 casework hair samples (Melton et al. 2005) and 116 casework skeletal remains (Nelson & Melton 2007).

SCREENING ASSAYS FOR mtDNA TYPING

Due to the effort both in terms of time and labor required to obtain full sequence information from mtDNA sequencing, screening approaches and rapid low-resolution typing assays can and have been used to eliminate the need for full analysis of samples that can be easily excluded from one another. Many times physical screening methods can put samples into context without having to indiscriminately perform mtDNA sequencing on all samples. For example, microscopic examinations of hair can help eliminate as many questioned hairs as possible leaving the mtDNA laboratory to concentrate their efforts on only key hairs (Houck & Budowle 2002). Likewise anthropological evaluations of bones or teeth can be important first screens prior to making the effort to analyze the mtDNA sequence (see Edson et al. 2004).

With the expense and effort required to obtain full mtDNA sequences across HV1 and HV2, the ability to rapidly screen out samples that do not match can be advantageous to overworked, understaffed, and poorly funded crime laboratories. Several assays have been developed and even validated for use in screening forensic casework (Table 14.5).

SSO Probes and Linear Array Typing Assays

One of the most widely used screening assays for assessing mtDNA variation used to date are the sequence-specific oligonucleotide (SSO) probes originally designed by Mark Stoneking and colleagues in 1991. Rather than sequencing the entire HV1 and HV2 regions, the most polymorphic sites are examined through hybridization of PCR products to oligonucleotide probes designed to anneal to different variants. The original paper describes 23 probes across 9 regions that permit evaluation of variation at 14 different nucleotide positions (Stoneking et al. 1991). The sites that are probed include 16126, 16129, 16217, 16223, 16304, 16311, 16362, 73, 146, 152, 195, 199, 247, and 309.1. A number of population studies have been conducted with these SSO probes including an examination of 2282 individuals from North America (Melton et al. 2001).

The original SSO probe assay required that the PCR products be attached through UV cross-linking to a nylon membrane. Then each radioactively labeled probe was individually hybridized at different temperatures and finally exposed to autoradiographic film for several hours (Stoneking et al. 1991). Roche Molecular Systems (Alameda, CA) has converted the SSO probe assay into a more convenient format involving colorimetric detection (e.g., Gabriel et al. 2001b). In a "reverse dot blot" format, the SSO probes are attached to the nylon membrane in a linear array of spatially resolved lines of probes. Biotin-labeled PCR

TABLE 14.5 Methods for Screening mtDNA Variation (See Butler & Levin 1998; Budowle et al. 2004).

Technique	Description	Reference
Sequence-specific oligonucleotide (SSO) dot blot assay	23 SSO probes testing 14 sites within nine regions from HV1 and HV2; 274 mtDNA types observed among 525 individuals from five ethnic groups.	Stoneking et al. (1991); Melton et al. (2001)
Mini-sequencing	Single base primer extension with fluorescent ddNTPs and poly(T)-tailed primers to yield different electrophoretic mobilities; 10 substitution and two length polymorphisms measured in the control region; 65 haplotypes observed from 152 British Caucasian samples.	Tully et al. (1996); Morley et al. (1999)
Single-strand conformational (SSCP)	Differences in DNA secondary structure are detected on a native polyacrylamide gel; 25 mtDNA types observed polymorphism among 45 Spanish individuals tested.	Alonso et al. (1996)
Low-stringency single- specific-primer PCR (LSSP-PCR)	Following regular PCR, a single primer and a low annealing temperature are used to generate a "signature" pattern; for 30 unrelated individuals, all signature patterns were different across the control region (1024 bp).	Barreto et al. (1996)
PCR-restriction fragment length polymorphism (PCR-RFLP)	A 199 bp region of HV1 is digested with RsaI; 19 unrelated mother-child pairs were examined with an 8% probability of a random match.	Pushnova et al. (1994); Butler et al. (1998a)
Denaturing gradient gel electrophoresis (DGGE)	Two DNA samples are mixed and run on a denaturing gradient gel; heteroduplexes, which travel more slowly through the gel, may be separated from the homoduplexes; samples that differ at a single location have been resolved.	Steighner et al. (1999); Tully et al. (2000)
Affymetrix high-density DNA chip hybridization array	135,000 probes complementary to the entire mtGenome are contained on a microchip for parallel processing through hybridization.	Chee et al. (1996)
Pyrosequencing	Sequencing by synthesis over ~50 nucleotides per reaction through an enzyme cascade that produces visible light; a total of 4 HV1, 4 HV2, and 11 coding region reactions were run.	Andreasson et al. (2002)
SNaPshot (mini-sequencing)	Allele-specific primer extension with 11 coding region SNPs combined into a single multiplex amplification and detection assay.	Vallone et al. (2004)
Denaturing HPLC	HV1 and HV2 PCR products for a known and an unknown sample source are generated and then mixed together; samples that differ from one another by at least one nucleotide will form a heteroduplex on the HPLC.	LaBerge et al. (2003)
Luminex 100 liquid bead array	30 SNPs within HV1 and HV2 are examined by allele-specific hybridization with SSO probes attached to different colored beads that are separated using flow cytometry.	Budowle et al. (2004)
LINEAR ARRAYs	Reverse dot blot hybridization with lines instead of dots using 18 SNPs in the same general probe regions as Stoneking et al. (1991).	Gabriel et al. (2003)

products are washed over nylon membrane strips containing immobilized SSO probes in the linear array and hybridized under uniform conditions. A streptavidin-horseradish peroxidase enzyme conjugate coupled with 3,3′,5,5′-tetramethyl-benzidine creates a light-blue precipitate using the same chemistry described for HLA-DQα reverse dot blot SSO probes (Saiki et al. 1989).

Figure 14.9 illustrates the probe layout for the LINEAR ARRAY Mitochondrial DNA HVI/HVII Region-Sequence Typing Kit now available from Roche Applied Sciences (Indianapolis, IN). The final linear array format examines 18 SNPs with 33 SSO probes present on 31 different lines. The Roche SSO probe sites are shown in Figure 14.4.

Two hypothetical results are illustrated in Figure 14.9 for non-matching K and Q samples. The K sample reported type of 1-1-1-1-1-1-1-1-1-1-1 is equivalent to the Cambridge Reference Sequence (see Figure 14.4). The Q sample possesses a different pattern and therefore can be excluded from the K sample. Notice that probe IE within HVI did not produce a signal from any of the three possible probes. This result is referred to as a "blank" and occurs due to additional polymorphisms that are present in close proximity to the polymorphic sites designed for detection in the assay. These additional polymorphisms disrupt hybridization of the PCR product and therefore no signal is seen for any of the probes in HVIE. Likewise weak (w) signals such as the "w1" type are caused by mismatches between the PCR products and the SSO probes attached on the nylon strip.

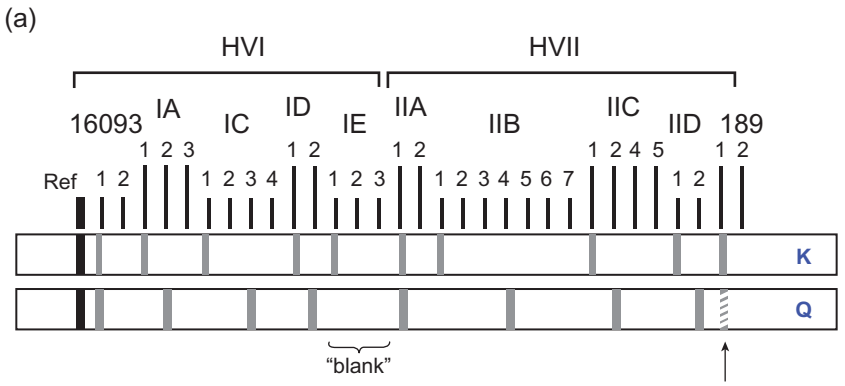

FIGURE 14.9 (a) Results schematically displayed of a known (K) reference and a question (Q) sample that do not match one another using the Roche LINEAR ARRAY mtDNA HVI/HVII Region-Sequence Typing Strips. (b) Types are reported as a string of numbers representing the LINEAR ARRAY probe results. Failure of the PCR product to bind to a probe region (e.g., HVIE in sample Q) is referred to as a "blank," is reported as a zero in the string of numbers, and is due to polymorphisms in the sample near the probe site that disrupt hybridization. Weak signals such as indicated by the arrow for 189 in sample Q are also due to a closely spaced polymorphism that disrupts full hybridization of the PCR product to the sequence-specific probe present on the LINEAR ARRAY.

Results from screening assays, such as the LINEAR ARRAY system described above, can be considered presumptive tests. They are useful in eliminating samples that can be excluded from one another. However, full HV1/HV2 sequencing would normally be performed to confirm any matches to see if differences outside of the SSO probe regions exist.

POPULATION DATABASES

Population databases play an important role in estimating the expected frequency of mtDNA haplotypes that are observed in casework when a suspect's mtDNA sequence matches that of an evidentiary sample. A great deal of effort has been expended to gather information from thousands of maternally unrelated individuals in various population groups around the world. Having high-quality information in the database is also important in order to make a reliable estimate of the frequency for a random match.

mtDNA typing results on samples from unknown sources are most useful if they are evaluated in comparison to a known sample or a database. Databases of more than 1000 unrelated individuals now exist and have been compiled from multiple population groups (Handt et al. 1998, Budowle et al. 1999, Attimonelli et al. 2000, Wittig et al. 2000, Röhl et al. 2001, Monson et al. 2002). The size of the database is important because without recombination between mtDNA molecules, an mtDNA sequence is treated as a single locus (i.e., haplotype instead of genotype).

The largest compiled database described to date contains HV1 and HV2 sequences from 14,138 individuals (Röhl et al. 2001). This information was collated from 103 mtDNA publications prior to January 2000, 13 data sets published in 2000 and 2001, and two unpublished data sets. Authors of the original publications were contacted in an effort to confirm and correct sequence errors, eliminate duplications, and harmonize nomenclatures, but not every query was answered. Of the 116 publications, 90 required some kind of change to correct errors or adjust nomenclature illustrating the challenge of compiling accurate mtDNA sequence databases. The authors conclude that their annotated database probably still contains errors and that while it can be used for qualitative identification of relevant reference populations for a given mtDNA type, the determination of a "legally defensible" frequency estimate of an mtDNA type within a population should be performed with higher-quality data yet to be produced (Röhl et al. 2001).

FBI mtDNA Database

The FBI has compiled the mtDNA Population Database also known as CODIS^{mt} (Monson et al. 2002) for the purpose of being able to determine a legally defensible frequency estimate. The CODIS^{mt} database has a forensic and a published literature component to it (Miller & Budowle 2001) in order to separate data obtained from laboratories following validated forensic protocols and academic research laboratories where data quality is not reviewed as carefully prior to publication.

The forensic database contains 4839 mtDNA profiles from 14 different populations (Table 14.6). These samples have been sequenced and the electropherograms carefully reviewed across positions 16024 to 16365 for HV1 and positions 73 to 340 for HV2.

TABLE 14.6 Summary of High Quality Forensic Profiles Present in the FBI Laboratory's mtDNA Population Database When It Was Released to the Public in April 2002 (Monson et al. 2002).

Population Name	Number of Profiles	Data Analysis on Group
African-American	1148	Budowle et al. (1999)
Apache	180	Budowle et al. (2002b)
Caucasian	1655	Budowle et al. (1999), Allard et al. (2002)
China/Taiwan	356	Allard et al. (2004)
Egypt	48	
Guam	87	Allard et al. (2004)
Hispanic	686	Budowle et al. (1999)
India	19	
Japan	163	Budowle et al. (1999)
Korea	182	Allard et al. (2004)
Navajo	146	Budowle et al. (2002b)
Pakistan	8	
Sierra Leone	109	Budowle et al. (1999)
Thailand	52	Allard et al. (2004)
Total	**4839**	

An additional 6106 published profiles have been compiled from the literature with annotated population information (Miller et al. 1996, Miller & Budowle 2001). For classification of mtDNA profiles, a standard 14-character nucleotide sequence identifier was assigned to each profile where the first three characters represent the country of origin, the second three characters the group or ethnic affiliation, and the final six characters are sequential acquisition numbers (Miller & Budowle 2001, Monson et al. 2002).

Both of these databases were publicly released in April 2002 in a Microsoft Access format and can be downloaded from the FBI website along with the "MitoSearch" analysis tool (Monson et al. 2002). MitoSearch can examine the population data sets listed in Table 14.6 for specific mtDNA sequences, which are entered based on differences from the Cambridge Reference Sequence. The software returns the number of times that the specified profile appears in each population group. For example, the mtDNA type 16129A, 263G, 309del, 315.1C occurs twice in 1148 African-American profiles, twice in 1655 Caucasian profiles, and not at all in 686 Hispanic profiles.

EMPOP

The European forensic mtDNA sequencing community has been actively engaged for a number of years in developing new high-quality population databases for forensic and

human identity testing applications. A European DNA Profiling Group mitochondrial DNA population database project (EMPOP) has gathered thousands of mtDNA sequences and constructed a high-quality mtDNA database that can be accessed at http://www.empop.org (Parson & Dür 2007). As of December 2010, there were 12,247 "forensic" (high-quality) data that could be searched online. A majority of the current samples are classified as West Eurasian (Caucasian).

mtDNAmanager

A Korean group from Yonsei University (Seoul, Korea) has created an online mtDNA searchable population database called mtDNAmanager (Lee et al. 2008). As of December 2010, this database contains 9294 mtDNA control region sequences grouped into five subsets: African (1496), West Eurasian (3673), East Asian (2326), Oceanian (114), and Admixed (1685). Many of these sequences are shared among the FBI and EMPOP databases and thus are not from "unique" haplotypes from unrelated individuals. mtDNAmanager can be accessed at http://mtmanager.yonsei.ac.kr/.

Issues with Sequence Quality

Concerns with mtDNA database sequence quality and the impact that it might have on accurately estimating frequency estimates for random matches have been raised by Peter Forster and Hans Bandelt (Röhl et al. 2001, Bandelt et al. 2001, Bandelt et al. 2002, Forster 2003, Salas et al. 2005). Using a statistical analysis clustering approach called phylogenetics, the similarities and differences between multiple and closely related DNA sequences (i.e., from the same region) can be compared systematically (see Wilson & Allard 2004). Sequence alignments are created and compared to identify samples that are extremely different. Extreme or unusual differences may be an indication that the sample was contaminated or the sequence data was incorrectly recorded. For example, a laboratory may put HV1 data for a sample with another sample's HV2 sequence and thereby create an artificial recombinant or accidental composite sequence. Thus phylogenetic analyses can play a role in verifying sequence quality (Bandelt et al. 2001, Wilson & Allard 2004).

Errors that creep into mitochondrial DNA population databases can be segregated into four different classes (Parson et al. 2004): (1) mistakes in the course of transcription of the results (i.e., clerical errors); (2) sample mix-up (e.g., putting data from HV1 on one sample together with data from HV2 on another sample); (3) contamination; and (4) use of different nomenclatures.

From a pilot collaborative study of 21 laboratories, 14 non-concordant haplotypes (16 individual errors) were observed out of a total of 150 submitted samples/haplotypes representing the examination of approximately 150,000 nucleotides (Parson et al. 2004). Measures are being put into place for complete electronic transfer of data and base calling to avoid the primary problem of clerical errors when transferring information from raw sequence data to final report. In the future, mtDNA databases may require retention of raw data for population samples in order to more easily verify authenticity of results should an inquiry into the origin of sequence results be needed at a later date (Parson et al. 2004). Search strategies using the complete query sequence will also likely be implemented (Röck et al. 2011).

Whole Mitochondrial Genome Sequencing

The first description of a methodology for sequencing the entire mtGenome was by Deborah Nickerson's group at the University of Washington (Rieder et al. 1998). They used 24 pairs of primers to amplify PCR products ranging in size from 765 bp to 1162 bp. These primer pairs provide on average almost 200 bases of overlap between the various PCR products spanning the mtGenome. Ingman et al. (2000) used the Nickerson laboratory sequencing strategy to launch the era of mitochondrial population genomics when they sequenced 53 mtGenomes from diverse world population groups. Max Ingman maintains an mtGenome polymorphism database at http://www.genpat.uu.se/mtDB/.

In the past few years, a number of other methodologies have appeared in the literature for sequencing entire mtGenomes (Table 14.7). Regardless of the sequencing strategy used, the biggest challenge in conducting this work remains efforts to reduce and eliminate errors in sequence review (see Herrnstadt et al. 2003). Fortunately, the reference sequence (rCRS) was updated prior to the explosion of mtGenome information that began with Ingman et al. (2000).

Resolving "Most Common Types"

One of the major challenges of mtDNA typing lies in the fact that many sequences fall into common groupings termed "most common types." For example, a review of the HV1/HV2 type distribution in 1655 Caucasians of U.S. and European descent (Monson et al. 2002)

TABLE 14.7 Summary of Published mtGenome DNA Sequencing Efforts from December 2000 to February 2004 Representing Almost 1000 Complete mtGenomes. As of December 2010, over 8000 mtGenomes Are Available on GenBank (See http://mitomap.org/bin/view.pl/MITOMAP/MitoSeqs). See Also http://www.phylotree.org/mtDNA_seqs.htm.

Population	Number Sequenced	Reference (GenBank Accessions)	Approach Taken
Samples of diverse worldwide origin	53	Ingman et al. (2000) AF346963–AF347015	24 PCR reactions, 48 sequencing reactions
Samples of diverse worldwide origin	33	Maca-Meyer et al. (2001) AF381981–AF382013	32 PCR reactions, 64 sequencing reactions
African, Asian, European origin	560 (coding region only)	Herrnstadt et al. (2002) Sequences available at www.mitokor.com	68 PCR reactions, 136 sequencing reactions
East Asian lineages	48	Kong et al. (2003) AY255133–AY255180	15 PCR reactions, 47 sequencing reactions
Australian and New Guinean Aborigines and Polynesians	52	Ingman & Gyllensten (2003) AY289051–AY289102	24 PCR reactions, 48 sequencing reactions
Most common Caucasian types	241	Coble et al. (2004) AY495090–AY495330	12 PCR reactions, 95 sequencing reactions

found that the most common mtDNA type, which matches the rCRS, occurred 7.1% of the time (Coble et al. 2004). Furthermore, it was observed that only 18 mtDNA types account for 20.8% of the total Caucasian data set (Coble et al. 2004). The presence of these most common types suggests that one out of every five times a mtDNA sequence analysis is performed on a Caucasian individual, the result would be expected to match numerous other individuals in a population database. While the same analysis revealed that approximately 50% of the 1655 individuals present in the European Caucasian population are "unique in the databasem," having a sample that falls into one of these most common types can present a disappointing statistic after all of the hard work taken to generate the full mtDNA HV1/HV2 sequence.

There has been an extensive search for distinguishing single nucleotide polymorphisms in samples possessing the most common Caucasian types (Parsons & Coble 2001). A total of 241 complete mtGenomes were sequenced from the 18 common European Caucasian HV1/HV2 types mentioned above (Coble 2004, Coble et al. 2004). The samples typed come from mtDNA haplogroups H, J, T, V, and K (see the next section for more discussion on haplogroups).

Examination of whole mtGenome sequence information expanded the 18 most common Caucasian HV1/HV2 types to 209 resolvable haplotypes (Coble et al. 2004). This almost 12-fold improvement in resolving power for these common HV1/HV2 types required about 27 times the amount of DNA sequencing—from 610 bases for just HV1/HV2 alone to ≈16,569 for the entire mtGenome. Obviously, this approach is not a cost effective one. Furthermore, even with the expansion in sequence information, 32 of the 241 individuals matched one or more individuals across the entire mtGenome.

From their extensive sequencing information, Coble et al. (2004) selected a battery of SNP markers to aid in resolving the most common Caucasian mtDNA HV1/HV2 types without the costly and time-consuming venture of having to sequence the entire mtGenome. A total of 59 informative SNPs were placed into eight multiplex panels (Coble et al. 2004). The first panel provides maximum resolution of the most common Caucasian HV1/HV2 mtDNA type (i.e., that matching rCRS) and examines the following nucleotides spanning the mtGenome: 477, 3010, 4580, 4793, 5004, 7028, 7202, 10211, 12858, 14470, and 16519. Vallone et al. (2004) combined these 11 SNP sites into a multiplex allele-specific primer extension or "SNaPshot" assay (see Chapter 12) that can reliably type a sample that contains only a few hundred copies of mtDNA.

Defining mtDNA Haplogroups

Over the course of typing mtDNA samples from various populations, researchers have observed that individuals often cluster into haplogroups that can be defined by particular polymorphic nucleotides (see Wallace et al. 1999, Ruiz-Pesini et al. 2004). These haplogroups were originally defined in the late 1980s and 1990s by grouping samples possessing the same or similar patterns when subjected to a series of restriction enzymes that were used to separate various mtDNA types from diverse populations around the world (Table 14.8). Mitochondrial DNA haplogroups have now been correlated to HV1/HV2 polymorphisms as well as the entire mtGenome variation. Haplogroups A, B, C, D, E, F, G, and M

TABLE 14.8 Major Mitochondrial Haplogroups and the Specific Polymorphisms in the Coding Region or Control Region That Define Them (See Finnila et al. 2001, Herrnstadt et al. 2002, Brandstatter et al. 2003, Kong et al. 2003, Allard et al. 2004, Quintans et al. 2004). Note That Not All Haplogroups, which Have Been Defined in the Literature, Are Listed Here. For Updated Haplogroup Information, See van Oven & Kayser (2009) and http://mitomap.org/bin/view.pl/MITOMAP/HaplogroupMarkers.

Haplogroup (Population)	Coding Region Polymorphisms	Control Region Polymorphisms (*not including 263G, 315.1C)
A (Asian)	663G	16233T, 16290T, 16319A, 235G
B (Asian)	9bp deletion, 16159C	16217C, 16189C
C (Asian)	13263G	16233T, 16298C, 16327T
D (Asian)	2092T, 5178A, 8414T	16362C
H (Caucasian)	7028C, 14766C	73A and lack of CRS differences*
H1 (Caucasian)	3010A	73A and lack of CRS differences*
H2 (Caucasian)	1438A, 4769A	73A and lack of CRS differences*
H3 (Caucasian)	6776C	73A and lack of CRS differences*
H4 (Caucasian)	3992T	73A and lack of CRS differences*
H5 (Caucasian)	4336C	73A and lack of CRS differences*
H6 (Caucasian)	3915A	73A and lack of CRS differences*
H7 (Caucasian)	4793G	73A and lack of CRS differences*
I (Caucasian)	1719A, 8251A, 10238C	16223T, 199C, 204C, 250C
J (Caucasian)	4216C, 12612G, 13708A	16069T, 16126C, 295T
J1 (Caucasian)	3010A	462T
J2 (Caucasian)	7476T, 15257A	195C
K (Caucasian)	12372A, 14798C	16224C, 16311C
L1 (African)	2758A, 3594T, 10810C	16187T, 16189C, 16223T, 16278T, 16311C
L2 (African)	3594T	16223T, 16278T
L3 (African)	3594C	16223T
M (Asian)	10400T, 10873C	16223T, 16298C
T (Caucasian)	709A, 1888A, 4917G, 10463C, 13368A, 14905A, 15607G, 15928A, 8697A	16126C, 16294T
U5 (Caucasian)	3197C	16270T
V (Caucasian)	4580A, 15904T	16298C, 72C
W (Caucasian)	709A, 1243C, 8251A, 8697G, 8994A	16223T, 189G, 195C, 204C, 207A
X (Caucasian)	1719A, 6221C, 8251G, 14470C	16189C, 16223T, 16278T, 195C

are typically associated with Asians while most Native Americans fall into haplogroups A, B, C, and D. Haplogroups L0, L1, L2, and L3 are African, and haplogroups H, I, J, K, T, U, V, W, and X are typically associated with European populations (Wallace et al. 1999).

Along the same lines as the multiplex SNP detection assay described above for resolving samples containing the most common HV1/HV2 types, Brandstätter et al. (2003) described a multiplex SNP system for categorizing European Caucasian haplogroups. This approach involves the analysis of 16 coding region SNPs to aid assignment of individual samples into one of the nine major European Caucasian mtDNA haplogroups listed above. For example, the presence of a cytosine at position 7028 indicates that the sample can be grouped into haplogroup H as opposed to the other groups whose individuals possess a thymine at 7028.

Another SNP typing assay was recently reported to examine 17 coding region SNPs in a single multiplexed detection assay (Quintans et al. 2004). A SNaPshot reaction is used to probe the following mtDNA nucleotide positions: 3010, 3915, 3992, 4216, 4336, 4529, 4580, 4769, 4793, 6776, 7028, 10398, 10400, 10873, 12308, 12705, and 14766. This assay was capable of breaking 266 samples into 20 different mtDNA haplogroup designations and aided in resolving some of the most common type (i.e., 263G, 315.1C) haplogroup H samples from one another.

Forensic population databases have been analyzed in terms of haplogroup information to aid in quality control of samples contained within a population group (Allard et al. 2002, Budowle et al. 2003, Allard et al. 2004).

Genetic Genealogy with mtDNA

Scientists have been using DNA for several decades to try to understand human migration patterns (Relethford 2001, Relethford 2003). Samples have been gathered from a number of individuals around the world often from isolated populations such as the Australian aborigines. The uniparental inheritance of mtDNA and Y-chromosome markers (see Chapter 13) makes it easier to trace ancestral lineages through multiple generations since the shuffling effects of recombination that promotes the diversity of autosomal DNA profiles are not present in haploid systems. The ability to successfully obtain mtDNA results from ancient bones is also useful, as has been demonstrated with the recovery of HV1 and HV2 sequences from Neanderthal remains that are thousands of years old (Krings et al. 1997).

While the same DNA markers are being used in these types of studies as in forensic DNA typing, the sample groups are often analyzed differently since direct comparisons cannot usually be made. Rather the DNA information obtained is extrapolated over many generations between the various populations tested. There is not a one-to-one unique match being made between a "suspect" and "evidence." Instead scientists are often guessing at what genetic signatures existed in the past based on various assumptions—with a bit of "storytelling" mixed in (see Goldstein & Chikhi 2002). However, large amounts of data are being collected in an attempt to better understand our heritage and travels as a human species (e.g., Helgason et al. 2003). Forensic DNA testing, disease diagnostics and anthropological and genealogical research efforts will all continue to benefit from the growth and developments in mitochondrial DNA analysis.

READING LIST AND INTERNET RESOURCES

General Information

Bandelt, H.-J., Richards, M., & Macaulay, V. (Eds.), (2006). *Human mitochondrial DNA and the evolution of homo sapiens*. Berlin-Heidelberg: Springer-Verlag Press.

Butler, J. M., & Levin, B. C. (1998). Forensic applications of mitochondrial DNA. *Trends in Biotechnology*, *16*, 158–162.

Butler, J. M., & Coble, M. D. (2006). *Y-Chromosome and mitochondrial DNA workshop*. Available at <http://www.cstl.nist.gov/biotech/strbase/YmtDNAworkshop.htm>.

Fisher, C. L. (2000). Mitochondrial DNA: Today and tomorrow. In *Proceedings of the 11th international symposium on human identification*. Available at <http://www.promega.com/geneticidproc/ussymp11proc/content/fisher.pdf>.

Fourney, R. M. (1998). Mitochondrial DNA and forensic analysis: A primer for law enforcement. *Canadian Society of Forensic Sciences Journal*, *31*, 45–53.

Legros, F., et al. (2004). Organization and dynamics of human mitochondrial DNA. *Journal of Cell Science*, *117*, 2653–2662.

Wong, L. J., & Boles, R. G. (2005). Mitochondrial DNA analysis in clinical laboratory diagnostics. *Clinica Chimica Acta*, *354*, 1–20.

Characteristics of mtDNA

Anderson, S., et al. (1981). Sequence and organization of the human mitochondrial genome. *Nature*, *290*, 457–465.

Andrews, R. M., et al. (1999). Reanalysis and revision of the Cambridge Reference Sequence for human mitochondrial DNA. *Nature Genetics*, *23*, 147.

Bodenteich, A., et al. (1992). Dinucleotide repeat in the human mitochondrial D-loop. *Human Molecular Genetics*, *1*, 140.

Brown, W. M., et al. (1979). Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Science of the United States of America*, *76*, 1967–1971.

Chen, X. J., & Butow, R. A. (2005). The organization and inheritance of the mitochondrial genome. *Nature Reviews Genetics*, *6*, 815–825.

Foran, D. R. (2006). Relative degradation of nuclear and mitochondrial DNA: An experimental approach. *Journal of Forensic Sciences*, *51*, 766–770.

Robin, E. D., & Wong, R. (1988). Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells. *Journal of Cellular Physiology*, *136*, 507–513.

Satoh, M., & Kuroiwa, T. (1991). Organization of multiple nucleoids and DNA molecules in mitochondria of a human cell. *Experimental Cell Research*, *196*, 137–140.

Scheffler, I. E. (1999). *Mitochondria*. New York: Wiley-Liss.

Szibor, R., et al. (1997). Mitochondrial D-loop 3′ (CA)n repeat polymorphism: Optimization of analysis and population data. *Electrophoresis*, *18*, 2857–2860.

Applications

Coble, M. D., et al. (2009). Mystery solved: The identification of the two missing Romanov children using DNA analysis. *PLoS ONE*, *4*, e4838.

Gill, P., et al. (1994). Identification of the remains of the Romano family by DNA analysis. *Nature Genetics*, *6*, 130–135.

Gill, P., et al. (1995). Establishing the identity of Anna Anderson Manahan. *Nature Genetics*, *9*, 9–10. (Erratum in: *Nature Genetics*, *9*, 218)

Ivanov, P. L., et al. (1996). Mitochondrial DNA sequence heteroplasmy in the Grand Duke of Russia Georgij Romanov establishes the authenticity of the remains of Tsar Nicholas II. *Nature Genetics*, *12*, 417–420.

Just, R. S., et al. (2011). Titanic's unknown child: the critical role of the mitochondrial DNA coding region in a re-identification effort. *Forensic Science International: Genetics*, *5*, 231–235.

Krings, M., et al. (1997). Neandertal DNA sequences and the origin of modern humans. *Cell*, *90*, 19–30.

Stone, A. C., et al. (2001). Mitochondrial DNA analysis of the presumptive remains of Jesse James. *Journal of Forensic Sciences*, *46*, 173–176.

Wallace, D. C., et al. (1999). Mitochondrial DNA variation in human evolution and disease. *Gene*, *238*, 211–230.

Measuring mtDNA Variation

Ingman, M., et al. (2000). Mitochondrial genome variation and the origin of modern humans. *Nature*, *408*, 708–713.
Richards, M., & Macaulay, V. (2001). The mitochondrial gene tree comes of age. *American Journal of Human Genetics*, *68*, 1315–1320.
Ruiz-Pesini, E., et al. (2004). Effects of purifying and adaptive selection on regional variation in human mtDNA. *Science*, *303*, 223–226.
Torroni, A., et al. (1996). Classification of European mtDNAs from an analysis of three European populations. *Genetics*, *144*, 1835–1850.

Mitochondrial DNA Sequencing in Forensic Casework

Forensic Casework

Budowle, B., et al. (1990). Mitochondrial DNA: a possible genetic material suitable for forensic analysis. In H. C. Lee & R. E. Gaensslen (Eds.), *Advances in Forensic Sciences* (pp. 76–97). Chicago: Year Book Medical Publishers.
Holland, M. M., et al. (1993). Mitochondrial DNA sequence analysis of human skeletal remains: Identification of remains from the Vietnam War. *Journal of Forensic Sciences*, *38*, 542–553.
Jarman, P. G., et al. (2009). Mitochondrial DNA validation in a state laboratory. *Journal of Forensic Sciences*, *54*, 95–102.
Marchi, E., & Pasacreta, R. J. (1997). Capillary electrophoresis in court: The landmark decision of the People of Tennessee versus Ware. *Journal of Capillary Electrophoresis*, *4*, 145–156.
Melton, T., & Nelson, K. (2001). Forensic mitochondrial DNA analysis: Two years of commercial casework experience in the United States. *Croatian Medical Journal*, *42*, 298–303.
Mosquera-Miguel, A., et al. (2009). Testing the performance of mtSNP minisequencing in forensic samples. *Forensic Science International: Genetics*, *3*, 261–264.
Reynolds, A. M., et al. (2005). Forensic casework analysis using the HVI/HVII mtDNA linear array assay. *Journal of Forensic Sciences*, *50*, 548–554.
Sullivan, K. M., et al. (1992). Identification of human remains by amplification and automated sequencing of mitochondrial DNA. *International Journal of Legal Medicine*, *105*, 83–86.

Steps Involved

Budowle, B., et al. (2003). Forensics and mitochondrial DNA: Applications, debates, and foundations. *Annual Reviews in Genomics and Human Genetics*, *4*, 119–141.
Chong, M. D., et al. (2005). Optimization of a duplex amplification and sequencing strategy for the HVI/HVII regions of human mitochondrial DNA for forensic casework. *Forensic Science International*, *154*, 137–148.
Hopgood, R., et al. (1992). Strategies for automated sequencing of human mitochondrial DNA directly from PCR products. *Biotechniques*, *13*, 82–92.
Isenberg, A. R. (2004). Forensic mitochondrial DNA analysis. In R. Saferstein (Ed.), *Forensic science handbook* (Vol. II, pp. 297–327). Upper Saddle River, New Jersey: Pearson Prentice Hall.
Isenberg, A. R., & Moore, J. M. (1999). Mitochondrial DNA analysis at the FBI Laboratory. *Forensic Science Communications*, *1*(2). Available at <http://www2.fbi.gov/hq/lab/fsc/backissu/july1999/dnalist.htm>.
Mabuchi, T., et al. (2007). Typing the 1.1 kb control region of human mitochondrial DNA in Japanese individuals. *Journal of Forensic Sciences*, *52*, 355–363.
Sanger, F., et al. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, *74*, 5463–5467.
Sullivan, K. M., et al. (1991). Automated amplification and sequencing of human mitochondrial DNA. *Electrophoresis*, *12*, 17–21.
Wilson, M. R., et al. (1995). Validation of mitochondrial DNA sequencing for forensic casework analysis. *International Journal of Legal Medicine*, *108*, 68–74.

Bones & Hair

Allen, M., et al. (1998). Mitochondrial DNA sequencing of shed hairs and saliva on robbery caps: sensitivity and matching probabilities. *Journal of Forensic Sciences*, *43*, 453–464.

Edson, S. M., et al. (2004). Naming the dead—confronting the realities of rapid identification of degraded skeletal remains. *Forensic Science Review*, *16*, 63–90.
Giambernardi, T. A., et al. (1998). Bovine serum albumin reverses inhibition of RT-PCR by melanin. *Biotechniques*, *25*, 564–566.
Hellmann, A., et al. (2001). STR typing of human telogen hairs—a new approach. *International Journal of Legal Medicine*, *114*, 269–273.
Houck, M. M., & Budowle, B. (2002). Correlation of microscopic and mitochondrial DNA hair comparisons. *Journal of Forensic Sciences*, *47*, 964–967.
Jehaes, E., et al. (1998). Evaluation of a decontamination protocol for hair shafts before mtDNA sequencing. *Forensic Science International*, *94*, 65–71.
Lee, H. Y., et al. (2006). Differential distribution of human mitochondrial DNA in somatic tissues and hairs. *Annals of Human Genetics*, *70*, 59–65.
Linch, C. A., et al. (2001). Human hair histogenesis for the mitochondrial DNA forensic scientist. *Journal of Forensic Sciences*, *46*, 844–853.
Loreille, O., et al. (2007). High efficiency DNA extraction from bone by total demineralization. *Forensic Science International: Genetics*, *1*, 191–195.
Melton, T., et al. (2005). Forensic mitochondrial DNA analysis of 691 casework hairs. *Journal of Forensic Sciences*, *50*, 73–80.
Nelson, K., & Melton, T. (2007). Forensic mitochondrial DNA analysis of 116 casework skeletal samples. *Journal of Forensic Sciences*, *52*, 557–561.
Pfeiffer, H., et al. (1999). Mitochondrial DNA in human hair shafts—existence of intra-individual differences? *International Journal of Legal Medicine*, *112*, 172–175.
Pfeiffer, H., et al. (1999). Mitochondrial DNA typing from human axillary, pubic and head hair shafts – success rates and sequence comparisons. *International Journal of Legal Medicine*, *112*, 287–290.
Wilson, M. R., et al. (1995). Extraction, PCR amplification and sequencing of mitochondrial DNA from human hair shafts. *Biotechniques*, *18*, 662–669.
Yoshii, T., et al. (1992). Presence of a PCR-inhibitor in hairs. *Nihon Hoigaku Zasshi*, *46*, 313–316.

mtDNA Quantitation

Alonso, A., et al. (2004). Real-time PCR designs to estimate nuclear and mitochondrial DNA copy number in forensic and ancient DNA studies. *Forensic Science International*, *139*, 141–149.
Andréasson, H., et al. (2002). Real-time DNA quantification of nuclear and mitochondrial DNA in forensic analysis. *BioTechniques*, *33*, 402–411.
Andréasson, H., et al. (2006). Nuclear and mitochondrial DNA quantification of various forensic materials. *Forensic Science International*, *164*, 56–64.
Meissner, C., et al. (2000). Quantification of mitochondrial DNA in human blood cells using an automated detection system. *Forensic Science International*, *113*, 109–112.
Niederstätter, H., et al. (2007). A modular real-time PCR concept for determining the quantity and quality of human nuclear and mitochondrial DNA. *Forensic Science International: Genetics*, *1*, 29–34.
von Wurmb-Schwark, N., et al. (2002). Quantification of human mitochondrial DNA in a real time PCR. *Forensic Science International*, *126*, 34–39.

PCR Amplification & Sequencing

Bini, C., et al. (2003). Different informativeness of the three hypervariable mitochondrial DNA regions in the population of Bologna (Italy). *Forensic Science International*, *135*, 48–52.
Butler, J. M., et al. (1998). Rapid mitochondrial DNA typing using restriction enzyme digestion of polymerase chain reaction amplicons followed by capillary electrophoresis separation with laser-induced fluorescence detection. *Electrophoresis*, *19*, 119–124.
Butler, J. M., et al. (1994). Quantitation of polymerase chain reaction products by capillary electrophoresis using laser fluorescence. *Journal of Chromatography B Biomedical Applications*, *658*, 271–280.
Comas, D., et al. (1999). Analysis of mtDNA HVRII in several human populations using an immobilised SSO probe hybridization assay. *European Journal of Human Genetics*, *7*, 459–468.

Dugan, K. A., et al. (2002). An improved method for post-PCR purification for mtDNA sequence analysis. *Journal of Forensic Sciences*, *47*, 811–818.
Gabriel, M. N., et al. (2001). Improved mtDNA sequence analysis of forensic remains using a "mini-primer set" amplification strategy. *Journal of Forensic Sciences*, *46*, 247–253.
Gabriel, M. N., et al. (2001). Population variation of human mitochondrial DNA hypervariable regions I and II in 105 Croatian individuals demonstrated by immobilized sequence-specific oligonucleotide probe analysis. *Croatian Medical Journal*, *42*, 328–335.
Gabriel, M. N., et al. (2003). Identification of human remains by immobilized sequence-specific oligonucleotide probe analysis of mtDNA hypervariable regions I and II. *Croatian Medical Journal*, *44*, 293–298.
Lee, L. G., et al. (1997). New energy transfer dyes for DNA sequencing. *Nucleic Acids Research*, *25*, 2816–2822.
Lutz, S., et al. (2000). Is it possible to differentiate mtDNA by means of HVIII in samples that cannot be distinguished by sequencing the HVI and HVII regions? *Forensic Science International*, *113*, 97–101.
Rasmussen, E. M., et al. (2002). Sequencing strategy of mitochondrial HV1 and HV2 DNA with length heteroplasmy. *Forensic Science International*, *129*, 209–213.
Stewart, J. E., et al. (2003). Evaluation of a multicapillary electrophoresis instrument for mitochondrial DNA typing. *Journal of Forensic Sciences*, *48*, 571–580.
Tully, G., et al. (1996). Rapid detection of mitochondrial sequence polymorphisms using multiplex solid-phase fluorescent minisequencing. *Genomics*, *34*, 107–113.
Wilson, M. R. (1997). Update to: Extraction, PCR amplification and sequencing of mitochondrial DNA from human hair shafts. In U. Gyllensten & J. Ellingboe (Eds.), *The PCR technique: DNA sequencing II* (pp. 322–328). Natick, Massachusetts: Eaton Publishing.

Smaller Amplicons

Berger, C., & Parson, W. (2009). Mini-midi-mito: Adapting the amplification and sequencing strategy of mtDNA to the degradation state of crime scene samples. *Forensic Science International: Genetics*, *3*, 149–153.
Eichmann, C., & Parson, W. (2008). "Mitominis" Multiplex PCR analysis of reduced size amplicons for compound sequence analysis of the entire mtDNA control region in highly degraded samples. *International Journal of Legal Medicine*, *122*, 385–388.
Lee, H. Y., et al. (2008). A modified mini-primer set for analyzing mitochondrial DNA control region sequences from highly degraded forensic samples. *Biotechniques*, *44*, 555–558.
von Wurmb-Schwark, N., et al. (2009). A new multiplex-PCR comprising autosomal and y-specific STRs and mitochondrial DNA to analyze highly degraded material. *Forensic Science International: Genetics*, *3*, 96–103.

New Technologies

Hall, T. A., et al. (2005). Base composition analysis of human mitochondrial DNA using electrospray ionization mass spectrometry: A novel tool for the identification and differentiation of humans. *Analytical Biochemistry*, *344*, 53–69.
Hall, T. A., et al. (2009). Base composition profiling of human mitochondrial DNA using polymerase chain reaction and direct automated electrospray ionization mass spectrometry. *Analytical Chemistry*, *81*, 7515–7526.
Mikkelsen, M., et al. (2009). Application of full mitochondrial genome sequencing using 454 GS FLX pyrosequencing. *Forensic Science International: Genetics*, *2*, 518–519.
Oberacher, H., et al. (2006). Profiling 627 mitochondrial nucleotides via the analysis of a 23-plex polymerase chain reaction by liquid chromatography-electrospray ionization time-of-flight mass spectrometry. *Analytical Chemistry*, *78*, 7816–7827.
Oberacher, H., et al. (2007). Liquid chromatography-electrospray ionization mass spectrometry for simultaneous detection of mtDNA length and nucleotide polymorphisms. *International Journal of Legal Medicine*, *121*, 57–67.

Collaborative and Interlab Studies

Alonso, A., et al. (2002). Results of the 1999-2000 collaborative exercise and proficiency testing program on mitochondrial DNA of the GEP-ISFG: an inter-laboratory study of the observed variability in the heteroplasmy level of hair from the same donor. *Forensic Science International*, *125*, 1–7.
Carracedo, A., et al. (1998). Reproducibility of mtDNA analysis between laboratories: A report of the European DNA Profiling group (EDNAP). *Forensic Science International*, *97*, 155–164.

Crespillo, M., et al. (2006). Results of the 2003-2004 GEP-ISFG collaborative study on mitochondrial DNA: Focus on the mtDNA profile of a mixed semen-saliva stain. *Forensic Science International*, *160*, 157–167.

Montesino, M., et al. (2007). Analysis of body fluid mixtures by mtDNA sequencing: An inter-laboratory study of the GEP-ISFG working group. *Forensic Science International*, *168*, 42–56.

Parson, W., et al. (2004). The EDNAP mitochondrial DNA population database (EMPOP) collaborative exercises: organization, results and perspectives. *Forensic Science International*, *139*, 215–226.

Prieto, L., et al. (2003). The 2000-2001 GEP-ISFG Collaborative Exercise on mtDNA: Assessing the cause of unsuccessful mtDNA PCR amplification of hair shaft samples. *Forensic Science International*, *134*, 46–53.

Prieto, L., et al. (2008). 2006 GEP-ISFG collaborative exercise on mtDNA: Reflections about interpretation, artifacts, and DNA mixtures. *Forensic Science International: Genetics*, *2*, 126–133.

Salas, A., et al. (2005). Mitochondrial DNA error prophylaxis: assessing the causes of errors in the GEP'02-03 proficiency testing trial. *Forensic Science International*, *148*, 191–198.

Tully, G., et al. (2004). Results of a collaborative study of the EDNAP group regarding mitochondrial DNA heteroplasmy and segregation in hair shafts. *Forensic Science International*, *140*, 1–11.

Turchi, C., et al. (2008). Italian mitochondrial DNA database: results of a collaborative exercise and proficiency testing. *International Journal of Legal Medicine*, *122*, 199–204.

Reference Materials

Levin, B. C., et al. (1999). A human mitochondrial DNA standard reference material for quality control in forensic identification, medical diagnosis, and mutation detection. *Genomics*, *55*, 135–146.

Levin, B. C., et al. (2001). A review of the DNA standard reference materials developed by the National Institute of Standards and Technology. *Fresenius Journal of Analytical Chemistry*, *370*, 213–219.

Levin, B. C., et al. (2003). Comparison of the complete mtDNA genome sequences of human cell lines—HL-60 and GM10742A—from individuals with pro-myelocytic leukemia and Leber hereditary optic neuropathy, respectively, and the inclusion of HL-60 in the NIST Human Mitochondrial DNA Standard Reference Material – SRM 2392-I. *Mitochondrion*, *2*, 387–400.

Szibor, R., et al. (2003). Cell line DNA typing in forensic genetics—the necessity of reliable standards. *Forensic Science International*, *138*, 37–43.

Interpretation of mtDNA Results

Brenner, C. H. (2010). Fundamental problem of forensic mathematics—the evidential value of a rare haplotype. *Forensic Science International: Genetics*, *4*, 281–291.

Budowle, B., et al. (2010). Automated alignment and nomenclature for consistent treatment of polymorphisms in the human mitochondrial DNA control region. *Journal of Forensic Sciences*, *55*, 1190–1195.

Carracedo, A., et al. (2000). DNA Commission of the International Society for Forensic Genetics: Guidelines for mitochondrial DNA typing. *Forensic Science International*, *110*, 79–85.

Dann, B. M., et al. (2004). Testing the effects of selected jury trial innovations on juror comprehension of contested mtDNA evidence. Final Technical Report, National Institute of Justice, Grant No. 2002-IJ-CX-0026. Available at <http://www.ncjrs.gov/pdffiles1/nij/grants/211000.pdf>.

Holland, M. M., & Parsons, T. J. (1999). Mitochondrial DNA sequence analysis – validation and use for forensic casework. *Forensic Science Review*, *11*, 21–50.

Parson, W., & Bandelt, H. -J. (2007). Extended guidelines for mtDNA typing of population data in forensic science. *Forensic Science International: Genetics*, *1*, 13–19.

Parsons, T. J., et al. (1997). A high observed substitution rate in the human mitochondrial DNA control region. *Nature Genetics*, *15*, 363–368.

Polanskey, D., et al. (2010). Comparison of Mitotyper rules and phylogenetic-based mtDNA nomenclature systems. *Journal of Forensic Sciences*, *55*, 1184–1189.

Stewart, J. E., et al. (2001). Length variation in HV2 of the human mitochondrial DNA control region. *Journal of Forensic Sciences*, *46*, 862–870.

SWGDAM (2003). Guidelines for mitochondrial DNA (mtDNA) nucleotide sequence interpretation. *Forensic Science Communications*, *5*(2) Available at <http://www2.fbi.gov/hq/lab/fsc/backissu/april2003/swgdammitodna.htm>.

Tully, G., et al. (2001). Considerations by the European DNA profiling (EDNAP) group on the working practices, nomenclature and interpretation of mitochondrial DNA profiles. *Forensic Science International*, *124*, 83–91.
Wilson, M. R., et al. (1993). Guidelines for the use of mitochondrial DNA sequencing in forensic science. *Crime Laboratory Digest*, *20*, 68–77.
Wilson, M. R., et al. (2002). Recommendations for consistent treatment of length variants in the human mitochondrial DNA control region. *Forensic Science International*, *129*, 35–42.
Wilson, M. R., et al. (2002). Further discussion of the consistent treatment of length variants in the human mitochondrial DNA control region. *Forensic Science Communications*, *4*(4) Available at <http://www2.fbi.gov/hq/lab/fsc/backissu/oct2002/wilson.htm>.

Heteroplasmy

Bendall, K. E., & Sykes, B. C. (1995). Length heteroplasmy in the first hypervariable segment of the human mtDNA control region. *American Journal of Human Genetics*, *57*, 248–256.
Bendall, K. E., et al. (1996). Heteroplasmic point mutations in the human mtDNA control region. *American Journal of Human Genetics*, *59*, 1276–1287.
Bandelt, H. J., & Parson, W. (2008). Consistent treatment of length variants in the human mtDNA control region: a reappraisal. *International Journal of Legal Medicine*, *122*, 11–21.
Berger, C., et al. (2011). Evaluating sequence-derived mtDNA length heteroplasmy by amplicon size analysis. *Forensic Science International: Genetics*, *5*, 142–145.
Brandstätter, A., & Parson, W. (2003). Mitochondrial DNA heteroplasmy or artefacts—a matter of the amplification strategy? *International Journal of Legal Medicine*, *117*, 180–184.
Brandstätter, A., et al. (2004). Monitoring the inheritance of heteroplasmy by computer-assisted detection of mixed basecalls in the entire human mitochondrial DNA control region. *International Journal of Legal Medicine*, *118*, 47–54.
Budowle, B., et al. (2002). Critique of interpretation of high levels of heteroplasmy in the human mitochondrial DNA hypervariable region I from hair. *Forensic Science International*, *126*, 30–33.
Calloway, C. D., et al. (2000). The frequency of heteroplasmy in the HVII region of mtDNA differs across tissue types and increases with age. *American Journal of Human Genetics*, *66*, 1384–1397.
Chen, X., et al. (1995). Rearranged mitochondrial genomes are present in human oocytes. *American Journal of Human Genetics*, *57*, 239–247.
Comas, D., et al. (1995). Heteroplasmy in the control region of human mitochondrial DNA. *Genome Research*, *5*, 89–90.
Forster, L., et al. (2010). Evaluating length heteroplasmy in the human mitochondrial DNA control region. *International Journal of Legal Medicine*, *124*, 133–142.
Grzybowski, T. (2000). Extremely high levels of human mitochondrial DNA heteroplasmy in single hair roots. *Electrophoresis*, *21*, 548–553.
Grzybowski, T., et al. (2003). High levels of mitochondrial DNA heteroplasmy in single hair roots: reanalysis and revision. *Electrophoresis*, *24*, 1159–1165.
He, Y., et al. (2010). Heteroplasmic mitochondrial DNA mutations in normal and tumour cells. *Nature*, *464*, 610–614.
Irwin, J. A., et al. (2009). Investigation of heteroplasmy in the human mitochondrial DNA control region: A synthesis of observations from more than 5000 global population samples. *Journal of Molecular Evolution*, *68*, 516–527.
Lagerström-Fermér, M., et al. (2001). Heteroplasmy of the human mtDNA control region remains constant during life. *American Journal of Human Genetics*, *68*, 1299–1301.
Melton, T. (2004). Mitochondrial DNA heteroplasmy. *Forensic Science Reviews*, *16*, 1–20.
Paneto, G. G., et al. (2010). Heteroplasmy in hair: study of mitochondrial DNA third hypervariable region in hair and blood samples. *Journal of Forensic Sciences*, *55*, 715–718.
Pflugradt, R., et al. (2011). A novel and effective separation method for single mitochondria analysis. *Mitochondrion*, *11*, 308–314.
Reiner, J. E., et al. (2010). Detection of heteroplasmic mitochondrial DNA in single mitochondria. *PLoS ONE*, *5*, e14359.
Roberts, K. A., & Calloway, C. (2011). Characterization of mitochondrial DNA sequence heteroplasmy in blood tissue and hair as a function of hair morphology. *Journal of Forensic Sciences*, *56*, 46–60.

Santos, C., et al. (2008). Frequency and pattern of heteroplasmy in the control region of human mitochondrial DNA. *Journal of Molecular Evolution*, *67*, 191–200.

Sekiguchi, K., et al. (2003). Inter- and intragenerational transmission of a human mitochondrial DNA heteroplasmy among 13 maternally-related individuals and differences between and within tissues in two family members. *Mitochondrion*, *2*, 401–414.

Sekiguchi, K., et al. (2004). Mitochondrial DNA heteroplasmy among hairs from single individuals. *Journal of Forensic Sciences*, *49*, 986–991.

Stoneking, M. (2000). Hypervariable sites in the mtDNA control region are mutational hotspots. *American Journal of Human Genetics*, *67*, 1029–1032.

Sullivan, K. M., et al. (1997). A single difference in mtDNA control region sequence observed between hair shaft and reference samples from a single donor: *Proceedings of the seventh international symposium on human identification-1996*. Madison, Wisconsin: Promega Corporation. pp. 126–130. Available at <http://www.promega.com/geneticidproc/ussymp7proc/0721.html>.

Steighner, R. J., et al. (1999). Comparative identity and homogeneity testing of the mtDNA HV1 region using denaturing gradient gel electrophoresis. *Journal of Forensic Sciences*, *44*, 1186–1198.

Tully, G. (1999). Mitochondrial DNA: A small but valuable genome. In *First international conference on forensic human identification*. Forensic Science Service.

Tully, L. A., et al. (2000). A sensitive denaturing gradient-gel electrophoresis assay reveals a high frequency of heteroplasmy in hypervariable region 1 of the human mtDNA control region. *American Journal of Human Genetics*, *67*, 432–443.

Wallace, D. C., et al. (1987). Sequence analysis of cDNAs for the human and bovine ATP synthase beta subunit: Mitochondrial DNA genes sustain seventeen times more mutations. *Current Genetics*, *12*, 81–90.

White, H. E., et al. (2005). Accurate detection and quantitation of heteroplasmic mitochondrial point mutations by pyrosequencing. *Genetic Testing*, *9*, 190–199.

Wilson, M. R., et al. (1997). A family exhibiting heteroplasmy in the human mitochondrial DNA control region reveals both somatic mosaicism and pronounced segregation of mitotypes. *Human Genetics*, *100*, 167–171.

Mixtures

Andréasson, H., et al. (2006). Quantification of mtDNA mixtures in forensic evidence material using pyrosequencing. *International Journal of Legal Medicine*, *120*, 383–390.

Bever, R. A., et al. (2003). Resolution of mixtures by cloning of the mitochondrial DNA control region. In *Proceedings of the 14th international symposium on human identification*. Available at <http://www.promega.com/geneticidproc/>.

Danielson, P. B., et al. (2007). Resolving mtDNA mixtures by denaturing high-performance liquid chromatography and linkage phase determination. *Forensic Science International: Genetics*, *1*, 148–153.

Hatsch, D., et al. (2007). A rape case solved by mitochondrial DNA mixture analysis. *Journal of Forensic Sciences*, *52*, 891–894.

Szibor, R., et al. (2003). Identification of the minor component of a mixed stain by using mismatch primer-induced restriction sites in amplified mtDNA. *International Journal of Legal Medicine*, *117*, 160–164.

Walker, J. A., et al. (2004). Resolution of mixed human DNA samples using mitochondrial DNA sequence variants. *Analytical Biochemistry*, *325*, 171–173.

Nuclear DNA Pseudogenes

Collura, R. V., & Stewart, C. B. (1995). Insertions and duplications of mtDNA in the nuclear genomes of Old World monkeys and hominoids. *Nature*, *378*, 485–489.

Lopez, J. V., et al. (1994). Numt, a recent transfer and tandem amplification of mitochondrial DNA in the nuclear genome of the domestic cat. *Journal of Molecular Evolution*, *39*, 171–190.

Morgan, M. A., et al. (1998). Amplification of human nuclear pseudogenes derived from mitochondrial DNA: A problem for mitochondrial DNA identity testing? *Proceedings of the eighth international symposium on human identification*. Madison, Wisconsin: Promega Corporation. Available at <http://www.promega.com/geneticidproc/ussymp8proc/34.html>.

Ramos, A., et al. (2009). Human mitochondrial DNA complete amplification and sequencing: a new validated primer set that prevents nuclear DNA sequences of mitochondrial origin co-amplification. *Electrophoresis*, *30*, 1587–1593.

Wallace, D. C., et al. (1997). Ancient mtDNA sequences in the human nuclear genome: a potential source of errors in identifying pathogenic mutations. *Proceedings of the National Academy of Sciences of the United States of America*, *94*, 14900–14905.

Zischler, H., et al. (1995). A nuclear "fossil" of the mitochondrial D-loop and the origin of modern humans. *Nature*, *378*, 489–492.

Recombination/Paternal Leakage

Awadalla, P., et al. (1999). Linkage disequilibrium and recombination in hominid mitochondrial DNA. *Science*, *286*, 2524–2525.

Elson, J. L., et al. (2001). Analysis of European mtDNAs for recombination. *American Journal of Human Genetics*, *68*, 145–153.

Eyre-Walker, A., et al. (1999). How clonal are human mitochondria? *Proceedings of the Royal Society of London B Biological Sciences*, *266*, 477–483.

Eyre-Walker, A. (2000). Do mitochondria recombine in humans? *Philosophical Transactions of the Royal Society of London B Biological Sciences*, *355*, 1573–1580.

Filosto, M., et al. (2003). Lack of paternal inheritance of muscle mitochondrial DNA in sporadic mitochondrial myopathies. *Annals of Neurology*, *54*, 524–526.

Giles, R. E., et al. (1980). Maternal inheritance of human mitochondrial DNA. *Proceedings of the National Academy of Sciences of the United States of America*, *77*, 6715–6719.

Gyllensten, U., et al. (1991). Paternal inheritance of mitochondrial DNA in mice. *Nature*, *352*, 255–257.

Hagelberg, E., et al. (1999). Evidence for mitochondrial DNA recombination in a human population of island Melanesia. *Proceedings of the Royal Society of London B Biological Sciences*, *266*, 485–492.

Hagelberg, E., et al. (2000). Erratum: Evidence for mitochondrial DNA recombination in a human population of island Melanesia. *Proceedings of the Royal Society of London B Biological Sciences*, *267*, 1595–1596.

Johns, D. R. (2003). Paternal transmission of mitochondrial DNA is (fortunately) rare. *Annals of Neurology*, *54*, 422–424.

Jorde, L. B., & Bamshad, M. (2000). Questioning evidence for recombination in human mitochondrial DNA. *Science*, *288*, 1931.

Kivisild, T., et al. (2000). Questioning evidence for recombination in human mitochondrial DNA. *Science*, *288*, 1931.

Kumar, S., et al. (2000). Questioning evidence for recombination in human mitochondrial DNA. *Science*, *288*, 1931.

Macaulay, V., et al. (1999). Mitochondrial DNA recombination—no need to panic. *Proceedings of the Royal Society of London B Biological Sciences*, *266*, 2037–2039.

Parsons, T. J., & Irwin, J. A. (2000). Questioning evidence for recombination in human mitochondrial DNA. *Science*, *288*, 1931.

Schwartz, M., & Vissing, J. (2002). Paternal inheritance of mitochondrial DNA. *New England Journal of Medicine*, *347*, 576–580.

Taylor, R. W., et al. (2003). Genotypes from patients indicate no paternal mitochondrial DNA contribution. *Annals of Neurology*, *54*, 521–524.

Wiuf, C. (2001). Recombination in human mitochondrial DNA? *Genetics*, *159*, 749–756.

Population Data & Databases

Allard, M. W., et al. (2002). Characterization of the Caucasian haplogroups present in the SWGDAM forensic mtDNA dataset for 1771 human control region sequences. Scientific Working Group on DNA Analysis Methods. Journal of Forensic Sciences, *47*, 1215–1223.

Allard, M. W., et al. (2004). Control region sequences for East Asian individuals in the Scientific Working Group on DNA Analysis Methods forensic mtDNA data set. *Legal Medicine*, *6*, 11–24.

Allard, M. W., et al. (2005). Characterization of human control region sequences of the African American SWGDAM forensic mtDNA data set. *Forensic Science International*, *148*, 169–179.

Allard, M. W., et al. (2006). Evaluation of variation in control region sequences for Hispanic individuals in the SWGDAM mtDNA data set. *Journal of Forensic Sciences*, *51*, 566–573.

Behar, D. M., et al. (2007). The Genographic Project public participation mitochondrial DNA database. *PloS Genetics*, *3*, e104.

Brandon, M. C., et al. (2009). MITOMASTER: A bioinformatics tool for the analysis of mitochondrial DNA sequences. *Human Mutation*, *30*, 1–6. Available at <http://mammag.web.uci.edu/twiki/bin/view/Mitomaster>.

Brandstätter, A., et al. (2004). Mitochondrial DNA control region sequences from Nairobi (Kenya): Inferring phylogenetic parameters for the establishment of a forensic database. *International Journal of Legal Medicine*, *118*, 294–306.

Brandstätter, A., et al. (2007). Generating population data for the EMPOP database – an overview of the mtDNA sequencing and data evaluation processes considering 273 Austrian control region sequences as example. *Forensic Science International*, *166*, 164–175.

Budowle, B., et al. (1999). Mitochondrial DNA regions HVI and HVII population data. *Forensic Science International*, *103*, 23–35.

Budowle, B., et al. (2002). HVI and HVII mitochondrial DNA data in Apaches and Navajos. *International Journal of Legal Medicine*, *116*, 212–215.

EMPOP-Mitochondrial DNA control region database. <http://www.empop.org>.

Ingman, M., & Gyllensten, U. (2006). mtDB: Human Mitochondrial Genome Database, a resource for population genetics and medical sciences. *Nucleic Acids Research*, *34*, D749–D751.

Irwin, J. A., et al. (2007). Development and expansion of high-quality control region databases to improve forensic mtDNA evidence interpretation. *Forensic Science International: Genetics*, *1*, 154–157.

Irwin, J. A., et al. (2011). mtGenome reference population databases and the future of forensic mtDNA analysis. *Forensic Science International: Genetics*, *5*, 222–225.

Lee, H. Y., et al. (2008). mtDNAmanager: a web-based tool for the management and quality analysis of mitochondrial DNA control-region sequences. *BMC Bioinformatics*, *9*, 483. Available at <http://www.biomedcentral.com>.

Lee, Y. S., et al. (2009). MitoVariome: A variome database of human mitochondrial DNA. *BMC Genomics*, *10*, S12.

Miller, K. W., & Budowle, B. (2001). A compendium of human mitochondrial DNA control region: Development of an international standard forensic database. *Croatian Medical Journal*, *42*, 315–327.

MITOMAP: A human mitochondrial genome database. <http://www.mitomap.org/>.

Monson, K. L., et al. (2002). The mtDNA population database: An integrated software and database resource. *Forensic Science Communications*, *4*(2) Available at: <http://www2.fbi.gov/hq/lab/fsc/backissu/april2002/miller1.htm>.

mtDB – Human mitochondrial genome database. <http://www.genpat.uu.se/mtDB/>.

mtDNAmanager. <http://mtmanager.yonsei.ac.kr/>.

Parson, W., et al. (2004). The EDNAP mitochondrial DNA population database (EMPOP) collaborative exercises: Organization, results and perspectives. *Forensic Science International*, *139*, 215–226.

Parson, W., & Dür, A. (2007). EMPOP – a forensic mtDNA database. *Forensic Science International: Genetics*, *1*, 88–92.

Pfeiffer, H., et al. (1999). Expanding the forensic German mitochondrial DNA control region database: Genetic diversity as a function of sample size and microgeography. *International Journal of Legal Medicine*, *112*, 291–298.

Röck, A., et al. (2011). SAM: String-based sequence search algorithm for mitochondrial DNA database queries. *Forensic Science International: Genetics*, *5*, 126–132.

Röhl, A., et al. (2001). An annotated mtDNA database. *International Journal of Legal Medicine*, *115*, 29–39.

Weissensteiner, H., et al. (2010). eCOMPAGT integrates mtDNA: Import, validation and export of mitochondrial DNA profiles for population genetics, tumour dynamics and genotype-phenotype association studies. *BMC Bioinformatics*, *11*, 122. Available at <http://dbis-informatik.uibk.ac.at/ecompagt>.

Quality Concerns with mtDNA Data

Bandelt, H. J., et al. (2001). Detecting errors in mtDNA data by phylogenetic analysis. *International Journal of Legal Medicine*, *115*, 64–69.

Bandelt, H. J., et al. (2002). The fingerprint of phantom mutations in mitochondrial DNA data. *American Journal of Human Genetics*, *71*, 1150–1160.

Bandelt, H. J., et al. (2004). Artificial recombination in forensic mtDNA population databases. *International Journal of Legal Medicine*, *118*, 267–273.

Brandstätter, A., et al. (2005). Phantom mutation hotspots in human mitochondrial DNA. *Electrophoresis*, *26*, 3414–3429.

Budowle, B., et al. (2004). Addressing the use of phylogenetics for identification of sequences in error in the SWGDAM mitochondrial DNA database. *Journal of Forensic Sciences*, *49*, 1256–1261.
Budowle, B., & Polanskey, D. (2005). FBI mtDNA database: a cogent perspective. *Science*, *307*, 845–847.
Forster, P. (2003). To err is human. *Annals of Human Genetics*, *67*, 2–4.
Parson, W. (2007). The art of reading sequence electropherograms. *Annals of Human Genetics*, *71*, 276–278.
Salas, A., et al. (2005). A practical guide to mitochondrial DNA error prevention in clinical, forensic, and population genetics. *Biochemical and Biophysical Research Communications*, *335*, 891–899.
Salas, A., et al. (2007). Phylogeographic investigations: The role of trees in forensic genetics. *Forensic Science International*, *168*, 1–13.

U.S. mtDNA Laboratories

Armed Forces DNA Identification Laboratory. <http://www.afip.org/consultation/AFMES/AFDIL/>.
Bode Technology Group. <http://www.bodetech.com>.
FBI Laboratory DNA Unit II. <http://www.fbi.gov/about-us/lab/mtdna>.
LabCorp. <http://www.labcorp.com>.
Mitotyping. <http://www.mitotyping.com>.
Orchid Cellmark. <http://www.orchidcellmark.com/forensicdna/>.
University of North Texas Center for Human Identification. <http://www.unthumanid.org/>.

Screening Assays

Brandstätter, A., et al. (2003). Rapid screening of mtDNA coding region SNPs for the identification of west European Caucasian haplogroups. *International Journal of Legal Medicine*, *117*, 291–298.
Dario, P., et al. (2009). mtSNP typing before mtDNA sequencing: why do it? *Forensic Science International: Genetics Supplement Series*, *2*, 187–188.
Kline, M. C., et al. (2005). Mitochondrial DNA typing screens with control region and coding region SNPs. *Journal of Forensic Sciences*, *50*, 377–385.
Kristinsson, R., et al. (2009). Comparative analysis of the HV1 and HV2 regions of human mitochondrial DNA by denaturing high-performance liquid chromatography. *Journal of Forensic Sciences*, *54*, 28–36.
Melton, T., et al. (2001). Diversity and heterogeneity in mitochondrial DNA of North American populations. *Journal of Forensic Sciences*, *46*, 46–52.
Nelson, T. M., et al. (2007). Development of a multiplex single base extension assay for mitochondrial DNA haplogroup typing. *Croatian Medical Journal*, *48*, 460–472.
Reynolds, R., et al. (2000). Detection of sequence variation in the HVII region of the human mitochondrial genome in 689 individuals using immobilized sequence-specific oligonucleotide probes. *Journal of Forensic Sciences*, *45*, 1210–1231.
Stoneking, M., et al. (1991). Population variation of human mtDNA control region sequences detected by enzymatic amplification and sequence-specific oligonucleotide probes. *American Journal of Human Genetics*, *48*, 370–382.

Whole mtGenome Sequencing

Coble, M. D. (2004). The identification of single nucleotide polymorphisms in the entire mitochondrial genome to increase the forensic discrimination of common HV1/HV2 types in the Caucasian population. PhD dissertation. Washington, DC: The George Washington University. Available at <http://www.cstl.nist.gov/biotech/strbase/pub_pres/Coble2004dis.pdf>.
Fendt, L., et al. (2009). Sequencing strategy for the whole mitochondrial genome resulting in high quality sequences. *BMC Genomics*, *10*, 139.
Herrnstadt, C., et al. (2002). Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups. *American Journal of Human Genetics*, *70*, 1152–1171.
Herrnstadt, C., et al. (2003). Errors, phantoms and otherwise, in human mtDNA sequences. *American Journal of Human Genetics*, *72*, 1585–1586.
Ingman, M., & Gyllensten, U. (2003). Mitochondrial genome variation and evolutionary history of Australian and New Guinean aborigines. *Genome Research*, *13*, 1600–1606.

Kong, Q. P., et al. (2003). Phylogeny of east Asian mitochondrial DNA lineages inferred from complete sequences. *American Journal of Human Genetics*, *73*, 671–676.

Maca-Meyer, N., et al. (2001). Major genomic mitochondrial lineages delineate early human expansions. *BMC Genetics*, *2*, 13.

Rieder, M. J., et al. (1998). Automating the identification of DNA variations using quality-based fluorescence re-sequencing: analysis of the human mitochondrial genome. *Nucleic Acids Research*, *26*, 967–973.

Most Common Types and Coding Region Assays

Brandstätter, A., et al. (2006). Dissection of mitochondrial superhaplogroup H using coding region SNPs. *Electrophoresis*, *27*, 2541–2550.

Budowle, B., et al. (2005). Forensic analysis of the mitochondrial coding region and association to disease. *International Journal of Legal Medicine*, *119*, 314–315.

Coble, M. D., et al. (2004). Single nucleotide polymorphisms over the entire mtDNA genome that increase the power of forensic testing in Caucasians. *International Journal of Legal Medicine*, *118*, 137–146.

Coble, M. D., et al. (2006). Effective strategies for forensic analysis in the mitochondrial DNA coding region. *International Journal of Legal Medicine*, *120*, 27–32.

Köhnemann, S., et al. (2008). A rapid mtDNA assay of 22 SNPs in one multiplex reaction increases the power of forensic testing in European Caucasians. *International Journal of Legal Medicine*, *122*, 517–523.

Köhnemann, S., et al. (2009). An economical mtDNA SNP assay detecting different mitochondrial haplogroups in identical HVR 1 samples of Caucasian ancestry. *Mitochondrion*, *9*, 370–375.

Nilsson, M., et al. (2008). Evaluation of mitochondrial DNA coding region assays for increased discrimination in forensic analysis. *Forensic Science International: Genetics*, *2*, 1–8.

Parsons, T. J., & Coble, M. D. (2001). Increasing the forensic discrimination of mitochondrial DNA testing through analysis of the entire mitochondrial DNA genome. *Croatian Medical Journal*, *42*, 304–309.

Parsons, T. J. (2006). Mitochondrial DNA genome sequencing and SNP assay development for increased power of discrimination. In *Final report on NIJ grant 2000-IJ-CX-K010*. Available at <http://www.ncjrs.gov/pdffiles1/nij/grants/213502.pdf>.

Quintáns, B., et al. (2004). Typing of mitochondrial DNA coding region SNPs of forensic and anthropological interest using SNaPshot minisequencing. *Forensic Science International*, *140*, 251–257.

Vallone, P. M., et al. (2004). A multiplex allele-specific primer extension assay for forensically informative SNPs distributed throughout the mitochondrial genome. *International Journal of Legal Medicine*, *118*, 147–157.

mtDNA Haplogroups and Phylogenetics

Phylotree. <http://www.phylotree.org/>.

van Oven, M., & Kayser, M. (2009). Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Human Mutation*, *30*, E386–E394.

Wilson, M. R., & Allard, M. W. (2004). Phylogenetics and mitochondrial DNA. *Forensic Science Review*, *16*, 37–62.

Genetic Genealogy and Human Migration Studies

Behar, D. M., et al. (2007). The Genographic Project public participation mitochondrial DNA database. *PLoS Genetics*, *3*, e104.

Bolnick, D. A., et al. (2007). Genetics: The science and business of genetic ancestry testing. *Science*, *318*, 399–400.

Brown, K. (2002). Tangled roots? Genetics meets genealogy. *Science*, *295*, 1634–1635.

DNA Ancestry. <http://dna.ancestry.com>.

FamilyTree DNA. <http://www.familytreedna.com/>.

Genographic project website. <https://genographic.nationalgeographic.com/genographic/>.

Goldstein, D. B., & Chikhi, L. (2002). Human migrations and population structure: What we know and why it matters. *Annual Review of Genomics and Human Genetics*, *3*, 129–152.

Helgason, A., et al. (2003). A populationwide coalescent analysis of Icelandic matrilineal and patrilineal genealogies: Evidence for a faster evolutionary rate of mtDNA lineages than Y chromosomes. *American Journal of Human Genetics*, *72*, 1370–1388.

International Society of Genetic Genealogy. <http://www.isogg.org/>.

Jobling, M. A., Hurles, M. E., & Tyler-Smith, C. (2004). *Human evolutionary genetics: Origins, peoples, and diseases*. New York: Garland Science.

Journal of Genetic Genealogy. <http://www.jogg.info/>.

Olson, S. (2002). *Mapping human history*. New York: Houghton Mifflin.

Relethford, J. H. (2001). *Genetics and the search for modern human origins*. New York: Wiley-Liss.

Relethford, J. H. (2003). *Reflections of our past: How human history is revealed in our genes*. Boulder, Colorado: Westview Press.

Smolenyak, M. S., & Turner, A. (2004). *Trace your roots with DNA: Using genetic tests to explore your family tree*. New York: Rodale.

Sorenson Molecular Genealogy Foundation. <http://www.smgf.org/>.

Stix, G. (2008, July). Traces of a distant past. *Scientific American*, 56–63.

Underhill, P. A., & Kivisild, T. (2007). Use of Y chromosome and mitochondrial DNA population structure in tracing human migrations. *Annual Review of Genetics*, *41*, 539–564.

Wells, S. (2007). *Deep ancestry: Inside the genographic project*. Washington, DC: National Geographic.