*John E. B. Stewart,[1] Ph.D.; Constance L. Fisher,[1] Ph.D.; Patricia J. Aagaard,[1] B.S.; Mark R. Wilson,[1] M.S.; Alice R. Isenberg,[1] Ph.D.; Deborah Polanskey,[1] B.S.; Eric Pokorak,[1] B.S.; Joseph A. DiZinno,[1] D.D.S.; and Bruce Budowle,[2] Ph.D.*

# Length Variation in HV2 of the Human Mitochondrial DNA Control Region

**REFERENCE:** Stewart JEB, Fisher CL, Aagaard PJ, Wilson MR, Isenberg AR, Polanskey D, Pokorak E, DiZinno JA, Budowle B. Length variation in HV2 of the human mitochondrial DNA control region. J Forensic Sci 2001;46(4):862–870.

**ABSTRACT:** Hair samples were typed from three individuals who exhibited length heteroplasmy in the homopolymeric cytosine stretches (C-stretch) in hypervariable region 2 (HV2). The study demonstrated that for different hairs within an individual, the HV2 C-stretch region can vary with respect to the number of cytosines and/or proportion of C-stretch length variants. Length heteroplasmy may occur regardless of the prominent length variant present in this region. Differences in the number of cytosines at the C-stretch region, or a variation in the relative amounts of heteroplasmic length variants, cannot be used to support an interpretation of exclusion.

**KEYWORDS:** forensic science, mitochondrial DNA, length heteroplasmy, C-stretch

Mitochondrial DNA (mtDNA) typing has been demonstrated to be a useful technique for characterizing biological evidence (1–3). Hypervariable regions one and two (HV1 and HV2) are the most polymorphic regions of mtDNA. The FBI and other laboratories typically sequence position sites 16024 to 16365, and 73 to 340 of HV1 and HV2, respectively, to characterize evidentiary samples (1,2,4). In the majority of cases, mtDNA is homoplasmic at operational levels. However, in some cases, heteroplasmy (2–7) may be observed. Heteroplasmy is the condition where more than a single mtDNA sequence is present in an individual. Generally, heteroplasmic mtDNA sequences differ at a single site within an individual. Heteroplasmy at two or more sites is expected to occur at much lower levels. Heteroplasmy may be observed in several ways including but not limited to: 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 3) individuals may be heteroplasmic in one tissue sample and homoplasmic in another tissue sample.

The possibility of heteroplasmy must be considered when evaluating and interpreting mtDNA types derived from evidentiary and reference samples. All three possible heteroplasmic patterns described above were observed for a point substitution at site 16355 in hair samples from maternally related family members (5). Thus, the relative amounts of heteroplasmic mtDNA can vary between tissues. Therefore, within the limits defined by the detection method, minor differences in mtDNA sequences between two samples may not be exclusionary. The studies by Wilson, et al. (5) and Bendall, et al. (6) demonstrated variation in the degree of heteroplasmic expression in hair which ranged from homoplasmy to various degrees of heteroplasmy.

Length heteroplasmy is observed when two or more length variants are found within an individual. Length heteroplasmy is most often manifested as multiple lengths of homopolymeric tracts at a particular site. Within and between cells or tissues, homopolymeric cytosine stretches (C-stretches) of different lengths can occur (7). Bendall and Sykes (7) observed length heteroplasmy in HV1 between sites 16184 and 16193 which resulted in C-stretch lengths ranging from 8 to 14 cytosine residues. Greenberg, et al. (8) reported that HV2 insertions at sites 309.1 (8 cytosines preceding a thymine) and 315.1 (6 cytosines after a thymine) can occur. Hauswirth and Clayton (9) observed different HV2 C-stretch lengths in clones of the KB cell line. The mechanism responsible for generating HV1 length heteroplasmy has been proposed to be replication slippage after a T to C transition at position 16189 has occurred (7). HV2 length heteroplasmy could be generated by the same mechanism (7) or by the insertion of cytosines in the 303–310 area.

To investigate the variation of length heteroplasmy in hairs, samples were typed from three individuals who exhibited length heteroplasmy in the C-stretch region in HV2. The data will be useful for interpreting results in human identity testing.

## Methods

DNA was extracted from individual hairs, pooled hair roots, and pooled hair shafts. Pooled hairs consisted of five hairs that were cut to approximately 1 cm in length and washed together in the same manner as individual hairs (1,10). The pooled hairs were divided into two samples: hair shafts with roots and hair shafts only (individually typed hairs were not included in the pooled hair samples). DNA was extracted, amplified, and sequenced from hairs, blood, and saliva as previously described (1,5,10) with the following modifications: the reagents, plastic tubes, Microcon-100 units, PCR tubes, and micro tissue glass grinders were exposed to UV light for a minimum of 15 min. UV light performance was monitored using a Uvicide® card (Vangard International, Inc., Neptune, NJ) according to manufacturer's recommendations (i.e., 3.6 Joules/$cm^2$/min). The Kontes glass micro tissue grinders were cleaned by rinsing with $ddH_2O$, followed by scrubbing with a sterile cotton-tipped applicator dipped in a warm detergent solution of 5% (w/v)

[1] DNA Analysis Unit II, FBI Laboratory, 935 Pennsylvania Avenue NW, Washington, D.C.

[2] Senior scientist, FBI Laboratory, 935 Pennsylvania Avenue NW, Washington, D.C.

Received 23 Sept. 1999; and in revised form 20 March 2000, 24 May 2000, 8 Sept. 2000; accepted 11 Sept. 2000.

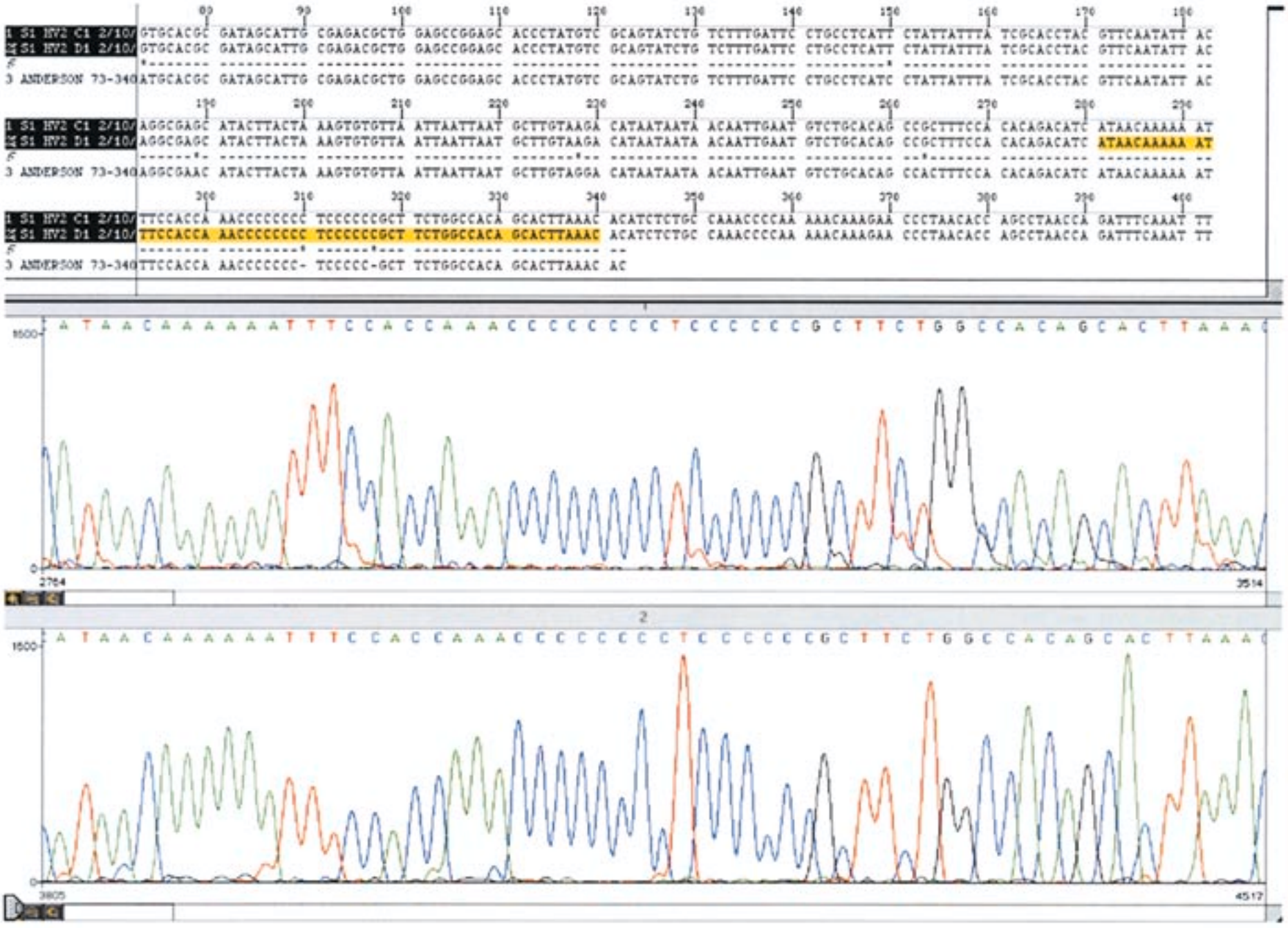


FIG. 1*a*—*Sequence and electropherogram from individual X exhibiting a major component of eight cytosines from sites 303 to 310, with a minor component of nine cytosines. The top panel shows eight cytosines and a characteristic out of phase sequence downstream of the C stretch, suggesting the presence of nine cytosines. The bottom panel, which is the sequence of the reverse strand, shows nine cytosines; the ninth C can be seen at position 302 under the A.*

Alconox® Terg-a-zyme™. Then, the grinders were rinsed with $ddH_2O$. Subsequently, 200 μL of 1 N $H_2SO_4$ were added to the mortar, the matching pestle was inserted, and grinding was briefly simulated. The mortars and pestles were allowed to soak for a minimum of 20 min in the 1 N $H_2SO_4$. The grinders were rinsed with $ddH_2O$ and the excess water was removed by centrifugation. The cleaned grinders were exposed to UV light for a minimum of 15 min. Grinders were used once and discarded. Hairs were washed and ground separately except where noted.

Cycle sequencing was performed using the ABI PRISM™ dRhodamine Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer, Foster City, CA). Samples were subjected to electrophoresis in a PE/ABI PRISM 377 DNA Sequencer. FMC Long Ranger™ gels (FMC BioProducts, Rockland, ME) were used with the ABI DNA sequencers.

## Results

Seven individual hairs, a blood sample, and a saliva sample were sequenced for the HV2 region for individual X. Four different HV2 C-stretch patterns were seen. Three individual hairs, the saliva sample, and the blood sample exhibited a length predominately of eight cytosine residues (8Cs) ($C_8TC_6$, Fig. 1*a*); a minor 9C variant can be seen conclusively in the bottom panel of Fig. 1*a.* One hair showed 7C's with a minor component of 8C's ($C_7TC_6$ and $C_8TC_6$, Fig. 1*b*). One hair exhibited 8C's with a minor component of 7C's ($C_8TC_6$ and $C_7TC_6$, Fig. 1*c*). In Fig. 1*c* there is potential evidence of a trace 9C variant; alternately the presence of the "apparent 9C" may be the result of overlapping frame shifts due to the presence of the 7C and 8C length variants. Two hairs exhibited 8Cs with a trace level of 7Cs and 9Cs ($C_8TC_6$, $C_7TC_6$, and $C_9TC_6$, Fig. 1*d*).

Fourteen individual hairs were typed from individual Y. Four different HV2 C-stretch patterns were observed. There were ten individual hairs that exhibited a length of 9C's with a minor component of 8C's and 10C's ($C_9TC_6$, $C_8TC_6$, and $C_{10}TC_6$, Fig. 2*a*). Four individual hairs had eight and nine C components ($C_8TC_6$ and $C_9TC_6$, Fig. 2*b*). One hair exhibited 9C's with a minor component of 10C's and 11C's ($C_9TC_6$, $C_{10}TC_6$, and $C_{11}TC_6$, Fig. 2*c*). The sample of five pooled hairs with roots and the sample of five pooled hair shafts exhibited a predominate length of 9C's, with a lower component of 8C's and 10C's ($C_9TC_6$, $C_8TC_6$, and $C_{10}TC_6$).

Four individual hairs and a blood sample were sequenced for individual Z. Four different HV2 C-stretch patterns were observed. The blood sample and one hair showed a length of 8C's with a minor component of 9C's ($C_8TC_6$ and $C_9TC_6$). One hair showed an almost equal mixture of 8C's and 9C's and minor component of 10C's ($C_8TC_6$, $C_9TC_6$, and $C_{10}TC_6$, Fig. 3*a*). Another hair exhibited a mixture of 7C's and 8C's ($C_7TC_6$ and $C_8TC_6$, Fig. 3*b*). One hair showed only 7C's ($C_7TC_6$, Fig. 3*c*). The results are listed in Table 1.

## Discussion

This study demonstrates that different hairs from an individual can vary in the HV2 C-stretch region with respect to the number of C's and in the relative proportion of C-stretch length. These data also demonstrate that length heteroplasmy may occur regardless of the prominent C-stretch length present. A major component is defined as the most abundant type observed in a sample. Length heteroplasmy was observed in an individual whose major sequence length in at least one hair consisted of as few as seven C's. Thus, variation can occur in the hairs from individuals whose major type is seven or more C's. Generally, length heteroplasmy becomes more prevalent when the C-stretch length reaches eight or more C's. In addition, some hairs from heteroplasmic individuals may show homoplasmy. The presence of hairs with and without length

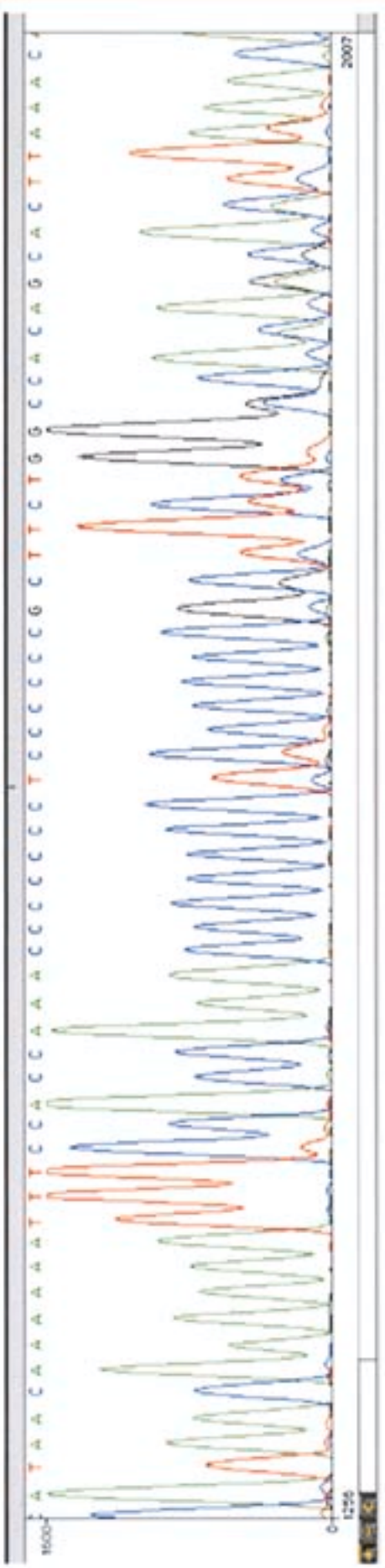

FIG. 1*b*—*Electropherogram from individual X exhibiting major component of seven cytosines and a minor component of eight cytosines.*

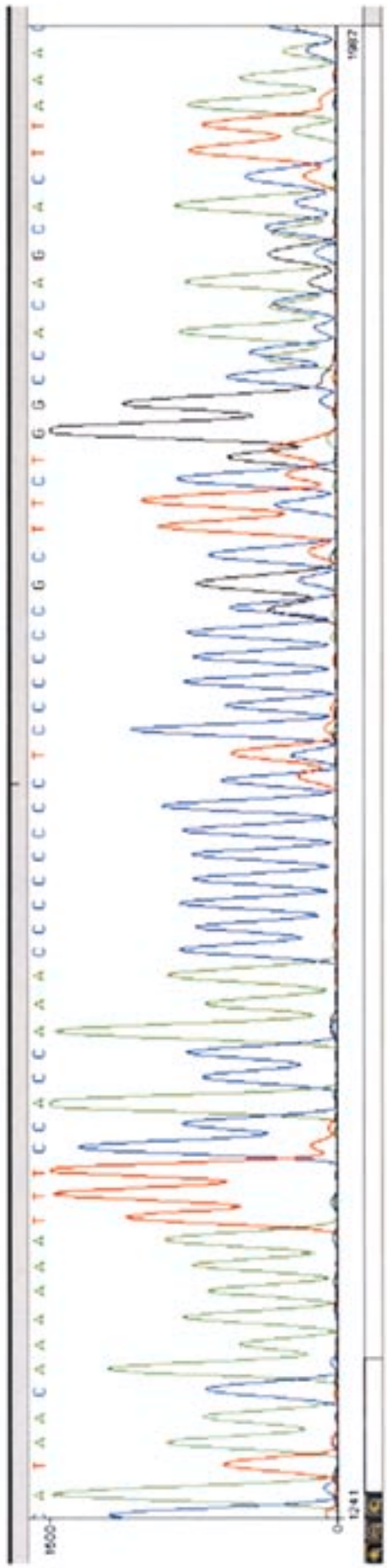

FIG. 1c—*Electropherogram from individual X with major component of eight cytosines, a minor amount of seven cytosines, and a trace amount of nine cytosines.*

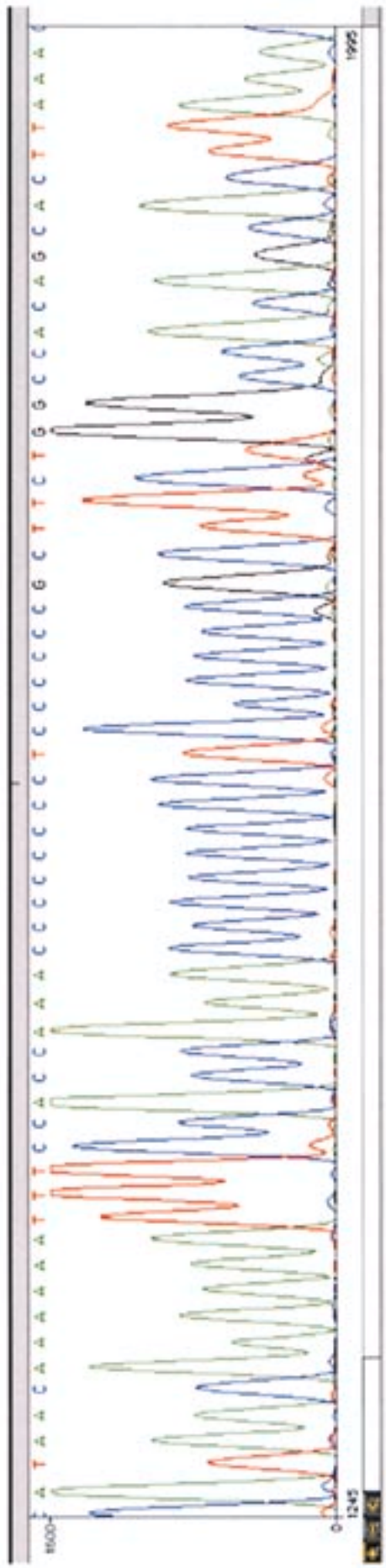

FIG. 1d—*Electropherogram from individual X exhibiting major component of eight cytosines with trace levels of 7 and 9 cytosines.*

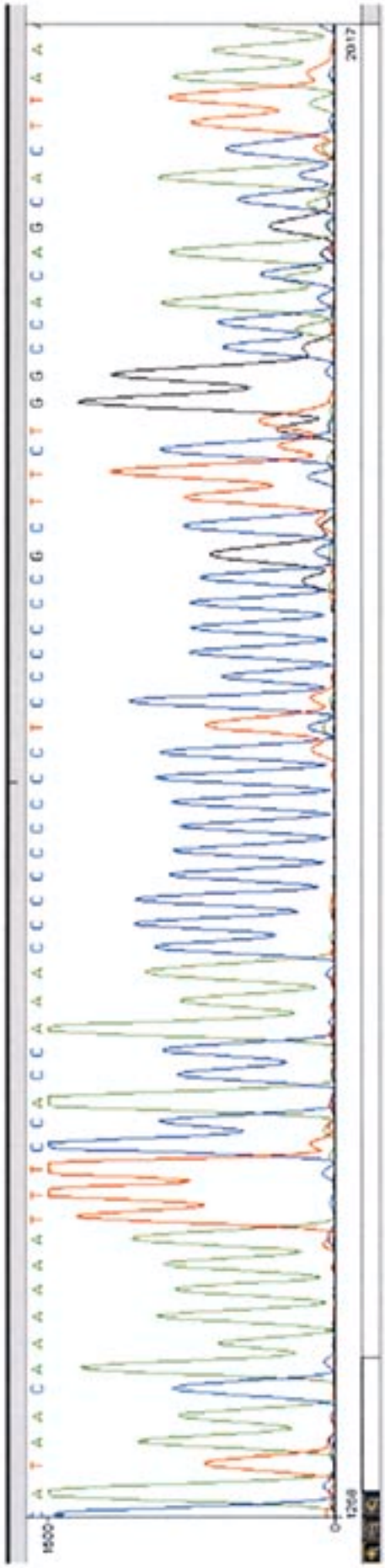

FIG. 2a—*Electropherogram from individual Y with major component of nine cytosines and a minor amount of eight and ten cytosines.*

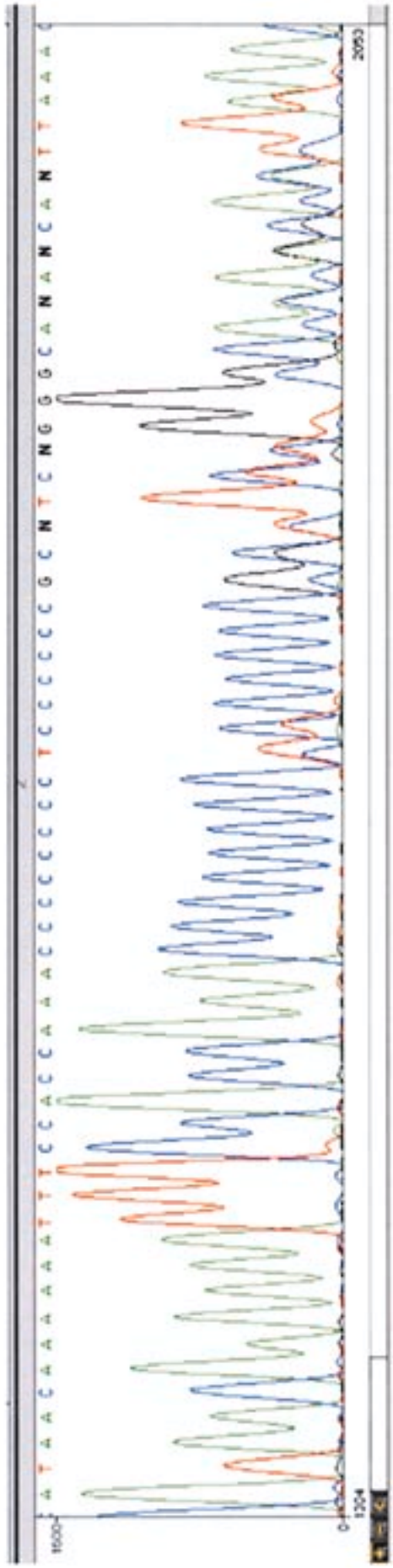

FIG. 2b—*Electropherogram from individual Y with major components of eight cytosines and of nine cytosines.*

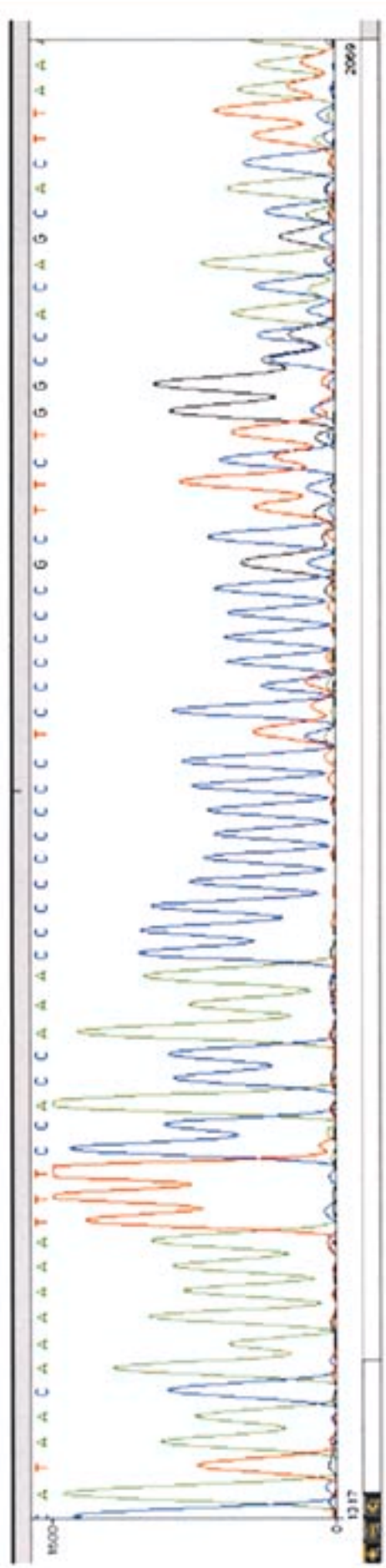

FIG. 2c—*Electropherogram from individual Y showing a major component of nine cytosines and a minor component of ten and eleven cytosines. Trace levels of other length variants are present.*

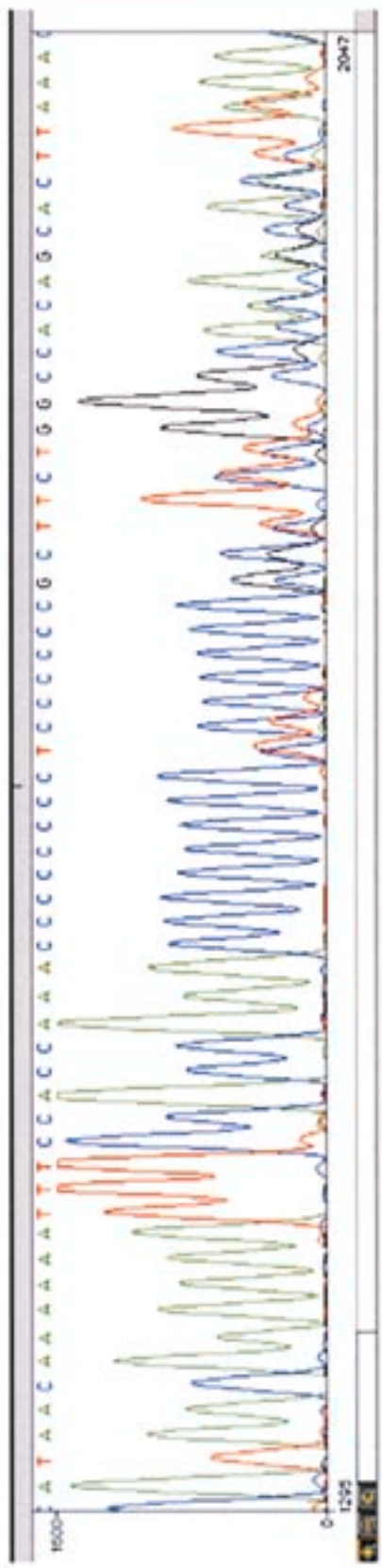

FIG. 3a—*Electropherogram from individual Z predominately with an almost equal mixture of eight and nine cytosines and a minor component of ten cytosines.*

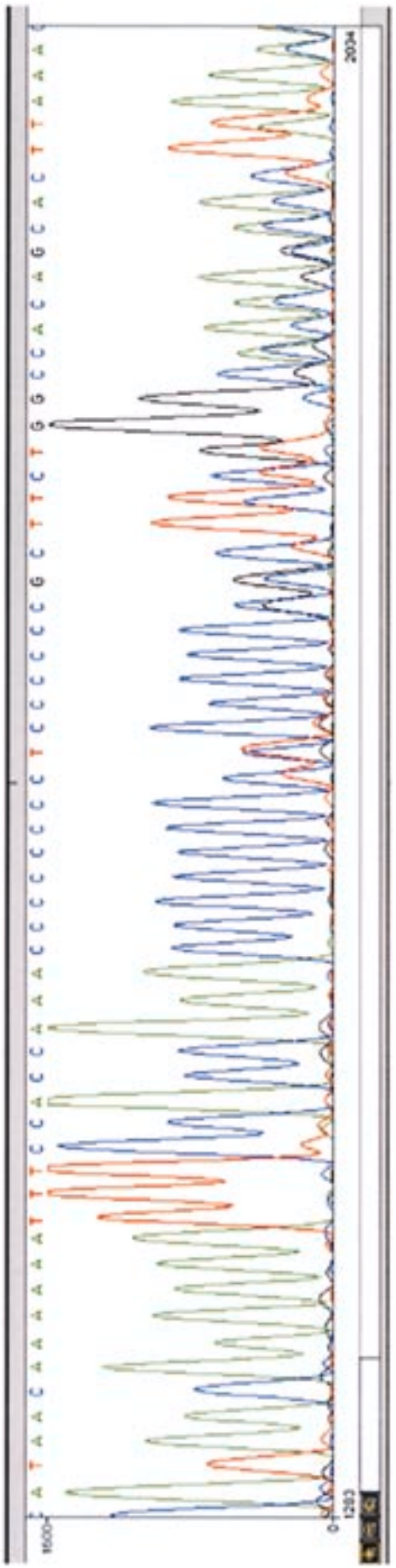

FIG. 3b—*Electropherogram from individual Z with major components of seven and eight cytosines.*

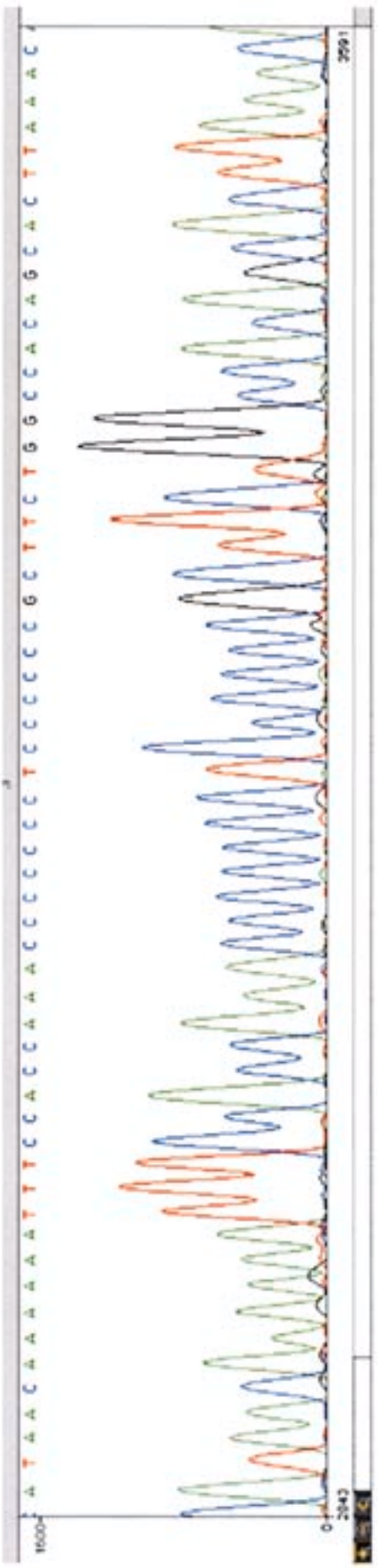

FIG. 3c—*Electropherogram from individual Z exhibiting only seven cytosines.*

TABLE 1—*Results of individual hairs, blood, and saliva sample.*

| Individual | Sample | Cytosines Before Position 310 T | | Type | Figures |
|---|---|---|---|---|---|
| | | Major Component | Minor Component | | |
| X | 3 individual hairs<br>Blood sample<br>Saliva sample | 8 Cs | 9 Cs | $C_8TC_6$<br>$C_9TC_6$ | 1A |
| | 1 hair | 7 Cs | 8 Cs | $C_7TC_6$<br>$C_8TC_6$ | 1B |
| | 1 hair | 8 Cs | 7 Cs<br>possible trace of 9 Cs | $C_8TC_6$<br>$C_7TC_6$ | 1C |
| | 2 individual hairs | 8 Cs | trace amount of | $C_7TC_6$ | 1D |
| | | | 7 Cs and 9 Cs | $C_8TC_6$<br>$C_9TC_6$ | |
| Y | 10 individual hairs<br>Pooled hairs with roots sample and pooled hair shafts sample | 9 Cs | trace amount of<br>8 Cs and 10 Cs | $C_8TC_6$<br>$C_9TC_6$<br>$C_{10}TC_6$ | 2A |
| | 4 individual hairs | 8 and 9 Cs | | $C_8TC_6$<br>$C_9TC_6$ | 2B |
| | 1 individual hair | 9 Cs | 10 Cs and 11 Cs | $C_9TC_6$<br>$C_{10}TC_6$<br>$C_{11}TC_6$ | 2C |
| Z | Blood sample<br>1 hair | 8 Cs | 9 Cs | $C_8TC_6$<br>$C_9TC_6$ | |
| | 1 hair | 8 and 9 Cs | 10 Cs | $C_8TC_6$<br>$C_9TC_6$<br>$C_{10}TC_6$ | 3A |
| | 1 hair | 7 and 8 Cs | | $C_7TC_6$<br>$C_8TC_6$ | 3B |
| | 1 hair | 7 Cs | | $C_7TC_6$ | 3C |

heteroplasmy in an individual is consistent with a bottleneck theory for mtDNA transmission. Thus, an individual whose blood shows 8, 9, or 10 C length variants may possess a hair with only 7C's. Because of the sensitivity of detection, minor length variants may be present but go undetected.

In a forensic setting, the mtDNA sequences from reference hair, blood, and/or saliva samples are compared with the types observed in the evidence samples. None, some, or all of the hairs sampled from an individual may have the exact same appearance as a sequence derived from a reference blood or saliva sample and the hairs may or may not exactly match each other. Based on the results of this study, differences in the number of C's in the C-stretch region or a variation in the relative amounts of heteroplasmic length variants, cannot be used alone to support an interpretation of exclusion. An exclusion based solely on differences in the length of homopolymeric stretches could be an erroneous interpretation. A pooled sample of reference hairs could provide a better representation of the heteroplasmic length profile of an individual and have the ancillary benefit of reducing the labor and cost of typing the hairs individually. Finally, in this study all possible length variants were described to demonstrate the variation that may exist for length heteroplasmy; in casework practice, typically the major length variant component is recorded and differences in C-stretch types are not used for exculpatory purposes.

*Acknowledgments*

We wish to thank Jenifer Smith, Samuel Baechtel, and Alan Guisti for their critical review of this manuscript.

## References

1. Wilson MR, Polanskey D, Butler J, DiZinno JA, Replogle J, Budowle B. Extraction, PCR, amplification and sequencing of mitochondrial DNA from human hair shafts. Biotechniques 1995;16:662–8.
2. Gill P, Ivanov PL, Kimpton C, Piercy R, Benson N, Tully G, et al. Identification of the remains of the Romanov family by DNA analysis. Nat Genet 1994;6:130–5.
3. Ivanov PL, Wadhams MJ, Roby RK, Holland MM, Weedn VW, Parsons TJ. Mitochondrial DNA sequence heteroplasmy in the Grand Duke of Russia Georgij Romanov establishes the authenticity of the remains of Tsar Nicholas II. Nat Genet 12:417–20.
4. Allen M, Engstrom A, Meyers S, Handt O, Saldeen T, von Haeseler A, et al. Mitochondrial DNA sequencing of shed hairs and saliva on robbery caps: sensitivity and matching probabilities. J Forensic Sci 1998; 43:453–64.
5. Wilson MR, Polanskey D, Replogle J, DiZinno JA, and Budowle B. A family exhibiting heteroplasmy in the human mitochondrial DNA control region reveals both somatic mosaicism and pronounced segregation of mitotypes. Hum Genet 1997;100:167–71.
6. Bendall KE, Macauly VA, Sykes BC. Variable levels of a heteroplasmic point mutation in individual hair roots. Am J Hum Genet 1997;61: 1303–8.
7. Bendall KE, Sykes BC. Length heteroplasmy in the first hypervariable

segment of the human mtDNA control region. Am J Hum Genet 1995;57:248–56.
8. Greenberg BD, Newbold JE, Sugino A. Intraspecific nucleotide sequence variability surrounding the origin of replication in human mitochondrial DNA. Gene 1983;21:33–49.
9. Hauswirth WW, Clayton DA. Length heterogeneity of a conserved displacement-loop sequence in human mitochondrial DNA. Nucl Acids Res 1985;13:8093–104.
10. Wilson MR, DiZinno JA, Polanskey D, Replogle J, Budowle B. Validation of mitochondrial DNA sequencing for forensic casework analysis. Int J Legal Med 1995;108:68–74.

Additional information and reprint requests:
John E. B. Stewart
DNA Analysis Unit II
FBI Laboratory
935 Pennsylvania Avenue, N.W.
Washington, D.C. 20535