CROATIAN MEDICAL JOURNAL
CMJ
42(3):328-335,2001

FORENSIC SCIENCES

Population Variation of Human Mitochondrial DNA Hypervariable Regions I and II in 105 Croatian Individuals Demonstrated by Immobilized Sequence-specific Oligonucleotide Probe Analysis

Matthew N. Gabriel[1,2], Cassandra D. Calloway[1], Rebecca L. Reynolds[1], Šimun Anđelinović[3], Dragan Primorac[4]

[1]Roche Molecular Systems, Department of Human Genetics, Alameda, CA, USA; [2]San Mateo County Sheriff's Office, Forensic Laboratory, San Mateo, CA, USA; [3]Department of Pathology and Forensic Medicine; and [4]Laboratory for Clinical and Forensic Genetics, Split University Hospital and School of Medicine, Split, Croatia

Aim. To detect sequence variation in 105 Croatian individuals by the use of duplex polymerase chain reaction amplification of full-length hypervariable region I and II (HVI/HVII) products and subsequent hybridization to a linear array of 27 immobilized sequence-specific oligonucleotide (SSO) probes, which targets six regions within HVI and HVII, and two additional sites, 189 and 16093.

Methods. Chelex-extracted bloodstains were used for amplification of HV regions. In all cases, a single robust amplification was sufficient for immobilized SSO probe typing and subsequent direct sequence analysis for both HVI and HVII. This method, suitable for a range of forensic samples (including shaft portions of single hairs), was also applied to the analysis of 18 skeletal elements recovered from a mass grave. Using a panel of immobilized SSO probes, we have developed a rapid screening approach to mitochondrial DNA (mtDNA) haplotyping before direct sequence analysis.

Results. We established a reference sequence database of mtDNA haplotypes for 105 randomly selected Croatian individuals. Fifty different mitotypes were observed (33 unique). The most frequent mitotypes occurred 18 times or ~17.1% [111111 189 (A) 16093 (T)] and 11 times or ~10.5% [131111 189 (A) 16093 (T)]; all other mitotypes occurred 5% or less. The corresponding genetic diversity value for this database was ~0.952. The usefulness of establishing an mtDNA reference database with immobilized SSO probe testing has been demonstrated by determining the strength of a match comparison obtained for one skeletal element and a corresponding maternal reference from 18 specimens recovered from a mass grave.

Conclusion. The sequence variation detected by the panel of immobilized SSO probes is sufficiently diverse to be used for identification of human skeletal remains from mass graves. The immobilized SSO typing strip targets polymorphic regions within HVI and HVII and is a useful identification tool for mass grave and mass disaster analysis, as well as for criminal casework testing.

Key words: *DNA, mitochondrial; forensic medicine; genetic markers; haplotypes; oligonucleotide probes; polymerase chain reaction; short tandem repeat*

Significant efforts are currently underway to identify missing individuals discovered in mass graves situated throughout Croatia and Southern Bosnia (1). For those cases submitted for DNA typing to the Split University Laboratory for Clinical and Forensic Genetics, approximately 85% have been successfully typed with forensic nuclear DNA markers (PM-DQA1 and short tandem repeat [STR] typing methodologies) (2). Environmental factors, including soil acidity and composition, heat, and humidity, can contribute to the fragmentation of DNA molecules, thereby preventing two-copy nuclear markers to persist over time. Further, polymerase chain reaction (PCR) inhibition and low template amounts can lead to uninterpretable typing results. For ~15% of cases that are unsuccessful with nuclear DNA analysis, we have implemented mitochondrial DNA (mtDNA) analysis of skeletal remains as an alternative identification method.

Methods for the detection of sequence variation within the human mtDNA control region have been established in the molecular anthropology and forensic DNA communities (3-7). PCR-based techniques commonly used to detect variation include direct sequence analysis of amplification products and screening with a panel of sequence-specific oligonucleotide (SSO) probes (5,6,8-14). These studies have established a high degree of polymorphism primarily

within two regions of the human mtDNA control region known as hypervariable region I (HVI) and hypervariable region II (HVII). The remainder of the non-coding region contains additional polymorphic stretches defined as variable regions I and II (VRI/VRII: 16366-72 and 341-577). Due to the high number of mtDNA genomes per cell (~500- 1,000 copies), forensic specimens such as shed hairs from crime scenes and significantly aged bone samples often contain enough intact DNA for analysis (15-22). For this reason, mtDNA testing is often successful in cases when amplification of two-copy nuclear markers fails. Further, the maternal inheritance of mtDNA molecules allows for a greatly expanded reference sample population for human identification, which is particularly helpful for mass grave analysis when distant maternal reference samples are required for match comparisons.

The most common forensic casework strategy for determining human mtDNA control region haplotypes utilizes direct sequence analysis of PCR amplification products (23). Though this method provides the greatest amount of sequence information, cases that consist of even small numbers of samples require a great deal of time and effort. Laboratory manipulation and data analysis requirements, coupled with reporting criteria for mtDNA haplotypes, make mtDNA analysis of a modest number of specimens a time-consuming task. Furthermore, many forensic casework analyses require confirmation of mtDNA types for only a subset of the total number of evidentiary samples submitted. Forensic human identification cases which involve reassociation of remains or mass grave analysis can benefit from a screening tool that enables accurate sorting or grouping of remains. In addition, the rapid screening of numerous crime scene samples to target those specimens that include or exclude a suspect(s) reference sample can aid the criminal investigative process dramatically. Recognizing the benefit of an immobilized SSO probe analysis for mtDNA haplotype screening in various forensic settings, we have developed a rapid linear array assay to target regions within HVI and HVII.

Unlike most mtDNA strategies that amplify HVI and HVII separately, this method involves co-amplification of full-length HVI and HVII products (444 bp and 415 bp, respectively) in a single reaction. We are currently developing two separate duplex reactions that target ~270-290 bp amplicons for HV region amplification from specimens that contain small amounts of intact mtDNA molecules due to limited specimen size and/or a high degree of DNA fragmentation. Both amplification strategies are sensitive to subcellular levels of mtDNA copies, and a single robust reaction is sufficient for SSO linear array typing and subsequent sequence analysis for both HV regions. These amplification strategies, then, conserve valuable sample extract. Hybridization of the PCR products to the immobilized SSO probe panel is performed without any sample preparation, allowing preliminary match comparisons to be made in a matter of hours. Samples that provide an SSO probe match are then subjected to sequence analysis, thereby keeping the number of subsequent confirmatory sequencing reactions to a minimum.

Immobilized SSO probe typing has also proven to be a more sensitive method of sequence detection compared with direct sequence analysis. The SSO system described here is capable of detecting the minor component of a known ratio of mixed samples at levels of 5-10%. Direct sequencing detects minor components at roughly ~15-25% (16) and is both position and chemistry dependent. Immobilized SSO probe analysis can then aid in making accurate determinations regarding mixtures, such as contamination (DNA contributions from more than one individual) and mtDNA heteroplasmy (the presence of more than one mtDNA sequence within an individual). Since contamination and heteroplasmy may appear as a mixture of sequences, often additional information is needed to report the 'true' mtDNA haplotype. Heteroplasmy appears to occur more frequently at the highly polymorphic positions within HVI and HVII (24) and, since the array of immobilized SSO probes targets the most polymorphic regions, detection of secondary sequences within an individual can be obtained routinely with this typing method. Likewise, mixtures between two unrelated specimens are likely to contain polymorphisms that are readily detected by the current panel of immobilized SSO probes. Thus, mtDNA typing with linear arrays of immobilized SSO probes provides a method for mixture determination in addition to direct sequencing.

This amplification and typing strategy has also been applied to the analysis of 18 skeletal elements recovered from a Croatian mass grave (report in preparation), and is suitable for a range of forensic samples, including shaft portions of single hairs (16,25).

We report sequence variation detection from 105 Croatian individuals using duplex PCR amplification of full-length HVI/HVII products and subsequent hybridization to a linear array of 27 immobilized SSO probes. For this study, chelex-extracted bloodstains were used for the amplification of HV regions, and in all cases, a single robust amplification was sufficient for SSO typing and subsequent direct sequence analysis for both HVI and HVII. The sequence variation detected by the panel of immobilized SSO probes we report here is sufficiently diverse to be used for human identification of skeletal remains from mass graves located throughout Croatia and Southern Bosnia. This paper describes the application of this typing approach to mass grave analysis and outlines many of the interpretation issues specific to immobilized SSO probe analysis of forensic specimens, as well as general mtDNA characteristics that affect the interpretation and reporting of mtDNA haplotypes.

Material and Methods

DNA Extraction from Reference Bloodstains

Approximately 0.5x0.5 cm of a bloodstain material was extracted from gauze swatches, using the standard Chelex 100© protocols (26). The final volume of each extract was ~200 μL and 1.5 μL of each extract was used for duplex PCR amplification.

HVI/HVII Duplex Polymerase Chain Reaction Amplification

PCR reactions were performed in 60 μL volumes using 1X AmpliTaq DNA Polymerase Buffer (12 mmol/L Tris-HCl, pH 8.3, 60 mmol/L KCl) with 2.4 mmol/L $MgCl_2$ final concentration (Applied Biosystems, Foster City, CA, USA), 200 μmol/L each dinucleotide triphosphate (dNTP), 0.2 mmol/L each primer, and 0.25 units/μL AmpliTaq Gold DNA Polymerase (Applied Biosystems). All reactions were performed in an Applied Biosystems 9600 thermal cycler with a 12-minute activation step at 92°C. Samples were subjected to 34-38 cycles at 92°C for 30 s, 60°C for 30 s, and 72°C for 30 s, followed by a final 72°C extension step for 10 min. Sequences for the 5′ biotinylated HVI and HVII primers were as follows: HVI (16024-16365): F15975 5′-CTCCACCATTAGCACCCAA-3′ and R16418 5′-ATTTCACGG AGGATGGTG-3′; HVII (73-340): F15 5′-CACCCTATTAACCACT CACG-3′ and R429 5′-CTGTTAAAAGTGCATACCGCCA-3′. A separate duplex amplification was performed to target variable region I and II as follows: VRI (16366-72): F16286 5′-CCTACCCAC CCTTAACAGTACA-3′ and R145 5′-GGCAGGAATCAAAGACA GA-3′; VRII (341-577): F314 5′-CCGCTTCTGGCCACAGCAC TTA-3′ and R641 5′-TGGGGTGATGTGAGCCCGTCTA-3′.

PCR products were analyzed on 1.5% agarose gels in 1 × TBE (89 mmol/L Tris base, 89 mmol/L boric acid, 2 mmol/L ethylenediaminetetraacetic acid [EDTA], pH 8.3) stained with 0.5 μg/mL ethidium bromide. Products were used directly for SSO typing and then purified with Millipore Ultrafree-MC centrifugal filter devices (Millipore Co., Bedford, MA, USA) to eliminate excess primer before sequence analysis. For Millipore device filtration, 30 μL of PCR product was added to 350 μL of TE (10 mmol/L Tris-HCl, 1mmol/L EDTA, pH 8.0) and centrifuged for 4 min at 12,000 rpm. Then, 20 μL of TE was added to the membrane and the recovery was transferred to a clean microcentrifuge tube for cycle sequencing.

Immobilized Sequence-specific Oligonucleotide Probe Strip Typing

PCR product yield was compared to a low DNA mass ladder (Life Technologies, Gaithersburg, MD, USA) and roughly 20-150 ng was used for immobilized SSO probe typing. The typing protocol follows similar conditions as the AmpliType PM/PM+DQA1 protocol (Applied Biosystems) with these modifications (16,25):

1) PCR product was prepared for hybridization by mixing 1:1 with Amplicor™ denaturant solution (1.6% NaOH) (Roche Diagnostics, Branchburg, Germany) and incubated at room temperature for 5 minutes.

2) Washes were performed with 3 mL per strip and the solution consisted of 2 x sodium phosphate buffer (SSPE) and 0.5% sodium dodecyl sulphate (SDS).

3) 8 μL of streptavidin-horseradish peroxidase conjugate was added per strip.

4) Color development was performed with a 3 mL volume of color development buffer (16,25).

Mitotype Designations

There are 4 types of probe signals that can occur at each polymorphic region. The most commonly observed strip typing result is a single visible probe within each of the regions. These probe signals are of similar intensity across the strip. Some DNA sequences that are not complementary to any of the probes for a given region result in the absence of probe signal. A "0" designation is given to the region with no detectable signal. The "0" types do not necessarily have the same sequence, because there are a number of variant sequences that will prevent hybridization to the probes for a given region. Further, a "W" designation is given to signals that appear weaker than the probe intensities in other regions. Weak signals are due to the presence of additional variants that partially destabilize the template-probe interaction. Lastly, multiple probe signals observed within a single polymorphic region suggest either the presence of mixed sequences due to a DNA contaminant (multiple DNA contributors) or to mtDNA heteroplasmy (presence of two detectable mtDNA sequences within an individual).

The linear arrays used to type these samples contain probes for the detection of polymorphisms at the following regions and positions: HVI A – 16126, 16129; HVI C – 16304, 16309, and 16311; HVI position 16093; HVII A – 73; HVII B – 146, 150, and 152; HVII C – 189, 195, 198, 200; HVII D – 247; HVII position 189. The mitotype designations [e.g., 131111 189 (A) 16093 (T)] refer to the particular polymorphism present within each region and at positions 189 and 16093. The specific sequence variant detected by each probe is given in Table 1.

Cycle Sequencing

We used 5-20 ng of the purified product, as determined by quantitation against the low DNA mass ladder described above, for cycle sequencing with BigDye™ Dye Terminator chemistry (Applied Biosystems). Cycle sequencing was performed in 20 μL volumes with 8.0 μL of BigDye™ Terminator mix and 1.6 μmol/L of the appropriate forward or reverse primer used for amplifica-

Table 1. Hypervariable regions (HVI) and (HVII) sequence variation detection[a]

Probe	Hypervariable region I												
			16126			16129							
A1	T	G	T	A	C	G	G	T					
A2	–	–	C	–	–	–	–	–					
A3	–	–	–	–	–	A	–	–					
			16304					16309	16311				
C1	A	G	T	A	C	A	T	A	G	T	A	C	
C2	–	–	C	–	–	–	–	–	–	–	–	–	
C3	–	–	–	–	–	–	–	–	–	C	–	–	
C4	–	–	–	–	–	–	–	G	–	–	–	–	
			16093										
16093T	A	T	T	T	C								
16093C	–	–	C	–	–								

Probe	Hypervariable region II														
			73												
A1	G	T	A	T	G										
A2	–	–	G	–	–										
			146				150	152							
B1	C	C	T	C	A	T	C	C	T	A	T				
B2	–	–	C	–	–	–	–	–	–	–	–				
B3	–	–	–	–	–	–	–	–	C	–	–				
B4	–	–	C	–	–	–	–	–	C	–	–				
B5	–	–	–	–	–	–	T	–	–	–	–				
B6	–	–	–	–	–	–	T	–	C	–	–				
B7	–	–	C	–	–	–	T	–	C	–	–				
			189						195		198		200		
C1	G	A	A	C	A	T	A	C	T	T	A	C	T	A	A	A
C2	–	–	–	–	–	–	–	–	C	–	–	–	–	–	–	–
C4	–	–	–	–	–	–	–	–	C	–	–	T	–	–	–	–
C5	–	–	G	–	–	–	–	–	–	–	–	–	–	G	–	–
			247												
D1	T	T	G	A	A										
D2	–	–	A	–	–										
			189												
189A	G	A	A	C	A										
189G	–	–	G	–	–										

[a]A list of the HVI and HVII positions targeted with the panel of immobilized SSO probes along with the sequence detected for the probes within each region.

tion. Cycle sequencing was performed as follows: 96°C for 15 s, 50°C for 5 s, and 60°C for 4 min for 25 cycles. One positive sequence control was performed for each group of samples with pGEM plasmid and M13 primer provided in the sequencing kit. Isopropanol precipitation of the product was performed by the addition of 80 μL of 75% isopropanol, mixing, and incubation at room temperature for 45 min. Centrifugation was carried out at 3,500 rpm for 45 min. The tray of tubes was then inverted and the isopropanol residue was removed by centrifugation at 2200 rpm for 2 minutes. Samples were resuspended in 20 μL of Template Suppression Reagent (TSR, Applied Biosystems) or 6 μL of 50 mg/mL blue dextran in 25 mmol/L EDTA with deionized formamide (1:5) loading buffer for ABI PRISM™ 310 capillary electrophoresis (Applied Biosystems) injection or polyacrylamide gel loading, respectively, and denatured by heating at 92°C for 3 min. For capillary electrophoresis, 5-30 s injections were performed with a 47 cm capillary and POP6™ polymer (Applied Biosystems). For samples run on the ABI PRISM™ 377 DNA Sequencer (Applied Biosystems), 1.5 μL of the 6 μL resuspension volume was electrophoresed on a 4% polyacrylamide gel (19:1) prepared from Long Ranger Singel packs (FMC Bioproducts, Rockland, ME, USA). Data analysis was performed with Sequencher™ software (Gene Codes, Ann Arbor, MI, USA) after fluorescent data collection.

Table 2. Mitotype frequencies observed in 105 Croatian individuals[a]

	HVI		HVII				Position			
Mitotype	A	C	A	B	C	D	189	16093	No.	freq.
1	0	1	2	3	1	1	A	T	2	0.0190
2	1	0	2	4	0	1	AW	T	1	0.0095
3	1	1	1	0	1	1	A	T	1	0.0095
4	1	1	1	1	1	1	A	T	18	0.1714
5	1	1	1	1	1W	1	G	T	1	0.0095
6	1	1	1	1	2	1	A	T	2	0.0190
7	1	1	1	2	1	1	A	T	5	0.0476
8	1	1	1	2	2	1	A	T	1	0.0095
9	1	1	1	3	1	1	A	T	1	0.0095
10	1	1	2	0	2	1	A	T	1	0.0095
11	1	1	2	1	0	1	G	T	2	0.0190
12	1	1	2	1	1	1	A	T	4	0.0381
13	1	1	2	1W	1	1	A	T	1	0.0095
14	1	1	2	3	0	1	0	T	1	0.0095
15	1	1	2	3	1	1	A	T	4	0.0381
16	1	1	2	5	1	1	A	T	2	0.0190
17	1	2	1	1	1	1	A	T	6	0.0571
18	1	2	1	1	1	1	A	C	1	0.0095
19	1	2	1	1	2	1	A	T	1	0.0095
20	1	2	2	1	1	1	A	T	1	0.0095
21	1	2	2	1	1	1	A	T/C	1	0.0095
22	1	2	2	5	1	1	A	T	1	0.0095
23	1	3	1	1	1	1	A	T	11	0.1048
24	1	3	1	1	1	1	A	C	1	0.0095
25	1	3	1	1	2	1	A	T	1	0.0095
26	1	3	2	1	0	1	G	T	1	0.0095
27	1	3	2	1	1	1	A	T	1	0.0095
28	1	3	2	4	2	1	A	T	1	0.0095
29	1	3	2	5	2	1	A	T	1	0.0095
30	1	4	2	3	1	1	A	T	1	0.0095
31	2	1	1	1	1	1	A	T	1	0.0095
32	2	1	2	1	0	1	0	T	3	0.0286
33	2	1	2	1	1	1	A	T	4	0.0381
34	2	1	2	1	1	1	A	C	1	0.0095
35	2	1	2	1	1	1	0	T	1	0.0095
36	2	1	2	1	2	1	A	T	1	0.0095
37	2	1	2	3/4W	0	1	0	T	1	0.0095
38	2	1	2	4	1	1	A	T	1	0.0095
39	2	1	2	5	1	1	0	T	1	0.0095
40	2	2	2	1	1	1	A	T	2	0.0190
41	2	2	2	3	1	1	A	T	1	0.0095
42	2	3	2	1	1	1	0	C	1	0.0095
43	3	0	1	1	1/2W	1	A	C	1	0.0095
44	3	1	1	1	1	1	A	T	2	0.0190
45	3	1	1	1	1	1	A	C	1	0.0095
46	3	1	2	1	0	1	A	T	1	0.0095
47	3	1	2	1	1	1	A	T	2	0.0190
48	3	1	2	3	0	1	G	T	1	0.0095
49	3	2	2	5	1	1	A	T	1	0.0095
50	3	3	2	1	1	1	A	T	2	0.0190
									105	0.9988

[a]A complete listing of the 50 mitotypes obtained for this reference database along with the corresponding frequencies for each type. A mitotype designation refers to the particular polymorphism present within each HVI and HVII region and at positions 189 and 16093 (e.g., 131111 189(A) 16093 (T) for mitotype 23).

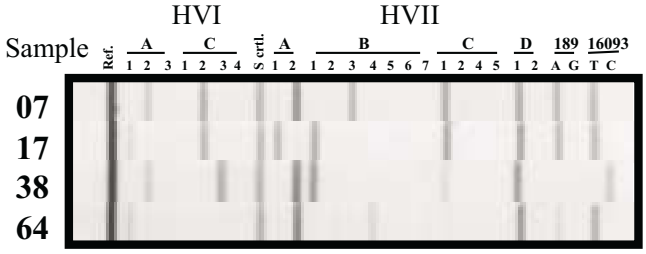
Results

Immobilized SSO probe typing of 105 Croatian individuals yielded a total of 49 different mitotypes, 32 of which were observed only once. Table 2 provides a complete listing of the mitotypes obtained with immobilized SSO probe analysis. The most frequent mitotypes occurred 18 times or ~17.1% [111111 189 (A) 16093 (T)] and 11 times or ~10.5% [131111 189 (A) 16093 (T)]. All other mitotypes were observed at frequencies of 5% or less. Regions designated with a "W" gave weak signals relative to other probe regions for a given sample, and those designated with "0" gave no detectable signal. For these samples, direct sequence analysis was performed to determine the cause of weak or undetectable signals. Decreased signal intensities or absent signals were concluded to be due to sequence polymorphisms within probe-binding regions that acted to destabilize the template-probe interaction. Figure 1 shows several samples that gave either weak signal intensity or absence of signal by SSO typing. The most common occurrence of absent probe signal was in the HVII C region. A "C0" mitotype was assigned 11 times (10.5%) for this region as a result of a cluster of polymorphisms. A transition at position 185 (G A) alone, or in combination with a 189 transition (A G), was sufficient to prevent signal detection. Likewise, polymorphisms at 189 (A G), 194 (C T), and 195 (T C) prevented template binding. Similarly, for the HVII B region, a variant at position 153 (A G) caused decreased template binding and a corresponding weak signal.

Multiple probe signals within a single probe region were observed in three samples and are noted in Table 2. As discussed above, multiple probe signals detected within a given region suggest a mixture of sequences. Each of the potential mixtures was confirmed by bi-directional sequence analysis and typing of a second amplification. Sample CR09 showed the presence of a "t/C" mixture (t/C = minor "t" with predominant "C"), where the first base refers to the Anderson reference sequence (27) and the second base refers to the variant sequence at position 16093. CR02 showed "T/c" at position 195 (corresponding to positive probe signals HVII C1/C2) and CR103 a roughly equal "T/C" mixture at 146 (corresponding to probe signals HVII B3/B4). SSO typing and sequence analysis did not reveal any additional mixed positions across the HV regions. Typing results and sequence analysis confirming the mixture of sequences ob-

served at position 16093 are shown in Figure 2. Based on this analysis, as well as information from our previous studies and the current literature characterizing heteroplasmy (16,25,28), we concluded that the blood samples of these three individuals are most likely heteroplasmic.

Direct sequence analysis was performed for the two most frequent mitotypes to determine the presence of additional polymorphic sites that might offer increased discrimination to the immobilized SSO probe analysis. Table 3 provides a complete listing of sequence information for these 29 samples. For the mitotype observed 18 out of 105 times (111111AT), there was additional HVI and HVII variation undetected by the current panel of SSO probes. A total of 16 additional polymorphic sites were revealed with direct sequencing (differences in the length of cytosine residues were excluded because of the high frequency of length heteroplasmy in the HVI and HVII cytosine stretches). One set of five samples (CR01, 59, 60, 87, and 101) could not be distinguished by direct sequence analysis of the HVI and HVII regions and had a common polymorphism at 16298 (C T). A second set of two samples (CR58 and 98) contained variants at 16169 (C T), 93 (A G), 207 (G A), and 338 (C T). The final set of two samples (CR92 and



HVI		HVII					
A	C	A	B	C	D	189	16093
2	2	2	3	1	1	A	T
1	2	1	1	1	1	A	T
2	3	2	1	1W	1	0	C
1	0	2	4	0	1	AW	T

Figure 1. Range of SSO types obtained from Croatian reference database. Figure shows the typing results of four Croatian samples (CR07, 17, 38, 64) with the corresponding mitotypes. Number designations are given to signals obtained in each of the typing regions. "W" is assigned for signals that are weaker relative to other probe regions and due to additional sequence variants that destabilize template-probe hybridization. "0" designates absent signals within a region due to the presence of variants within probe binding regions, which prevent hybridization. Samples 38 and 64 are examples of weak or absent signals, and 07 and 17 are shown for comparison.

Table 3. Sequence variation observed in the two most frequent mitotypes[a]

	HVI										VRI							HVII											VRII		
	16169	16183	16189	16239	16298	16311	16354	16355	16356	16362	16366	16482	16519	42.1	55	56.1	72	93	114	153	204	207	239	263	309.1	309.2	315.1	338	508	522	523
And.	C	T	T	C	T	T	C	C	T	T	C	A	T	T	T	T	T	A	G	A	T	G	T	A	T	T	T	C	A	C	A
111111AT																															
CR01	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	G	C	.	C	.	G	.	.
CR10	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	.	C	.	.	.	.
CR28	.	.	.	.	.	.	.	.	.	.	T	.	.	.	C	C	.	.	.	.	.	.	.	G	.	.	C	.	.	.	.
CR52	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.	.	.	.	.	.	.	.	G	.	.	C	.	.	.	.
CR58	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	A	.	G	C	.	C	T	.	.	.
CR59	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	G	C	.	C	.	.	.	.
CR60	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	G	C	.	C	.	.	.	.
CR65	.	.	.	.	.	.	.	.	.	C	.	G	.	.	.	.	.	.	.	.	.	.	C	G	.	.	C	.	.	.	.
CR76	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	C	.	.	D[b]	D
CR78	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	G	.	.	C	.	.	.	.
CR81	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	G	C	.	.	G	C	C	C	.	.	.	.
CR87	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	G	C	.	C	.	.	.	.
CR90	.	.	.	T	.	.	.	.	C	.	.	.	C	.	.	.	.	.	.	.	.	.	.	G	.	.	C	.	.	.	.
CR91	.	C	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	G	C	C	C	.	.	D	D
CR92	.	.	.	.	.	.	.	T	.	.	.	.	C	C	.	.	.	.	.	.	.	.	.	G	.	.	C	.	.	.	.
CR93	.	.	.	.	.	.	.	T	.	.	.	.	C	C	.	.	.	.	.	.	.	.	.	G	.	.	C	.	.	.	.
CR98	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	A	.	G	.	.	C	T	.	.	.
CR101	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	G	C	.	C	.	.	.	.
131111AT																															
CR03	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	.	C	.	.	D	D
CR04	.	.	.	.	.	C	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	G	.	.	C	.	.	.	.
CR27	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	.	C	.	.	.	.
CR31	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	.	C	.	.	.	.
CR32	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	.	C	.	.	D	D
CR57	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	.	C	.	.	.	.
CR61	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	C	C	.	.	.	.
CR85	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	C	.	.	.	.
CR95	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	.	C	.	.	D	D
CR96	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	C	.	.	.	.
CR100	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	C	.	.	.	.

[a]Complete sequence information for hypervariable (HV) and variable regions (VR) in the two most frequent mitotypes [111111 189 (A) 16093 (T) and 131111 189 (A) 16093 (T) – 17.1% and 10.5%, respectively] observed in this population. Sequence designations follow the numbering scheme for the Anderson reference sequence (5).
[b]D indicates deletion.

93) that could not be distinguished by sequence analysis contained a variant at 16355 (C T). In general, variation was widespread across these regions; however, several polymorphisms did occur multiple times and may serve as candidate sites for additional probes in this population. The most common site of variation undetected by the panel of SSO probes was 16298, which occurred 5 times among the most frequent types. There were 7 additional positions that occurred twice.

The second most frequent mitotype (131111AT), which occurred 11 out of 105 times, did not contain any additional HVI or HVII variation aside from the 16311 (C T) variant, accounting for the HVI C3 signal, and 263 (A G) and 315.1 (C), which were observed for all 29 of the most frequent mitotypes. Primers were then designed to target variable region I and II (VRI/VRII) sequences. All 29 samples were sequenced with these primers in an attempt to distinguish shared haplotypes. Results from VR sequencing are shown in Table 3. Sample CR01 from the "111111AT" group contained an A G transition at position 508 in VRII, which was not shared by CR59, 60, 87, and 101. Three of the 11 "131111AT" mitotypes (CR03, 32, and 95) contained VRII deletions at position 522 and 523, whereas CR04 contained a T C transition at 16519 in VRI. A total of 10 VR positions uncovered additional variable sites, with the most common polymorphisms occurring at 16519 (T C), 522 and 523 (deletion), and 72 (T C) for 7, 5, and 5 times, respectively.

After HV and VR sequencing, three sets of samples were still indistinguishable: 92 and 93; 58 and 98; and 59, 60, 87, and 101 (see Table 3 for sequences). Despite efforts to collect samples at random, the possibility that the individuals within these sets are maternal relatives must be addressed. Numerous population studies of sequence variation in mitochondrial DNA reveal that common haplotypes are found within populations. Therefore, to answer the question of relatedness for the shared haplotypes, we will need to analyze nuclear DNA markers. This set of 105 individuals has been genotyped on the basis of the 13 STR loci used by the forensic DNA community and the data generated in this study are currently being analyzed.

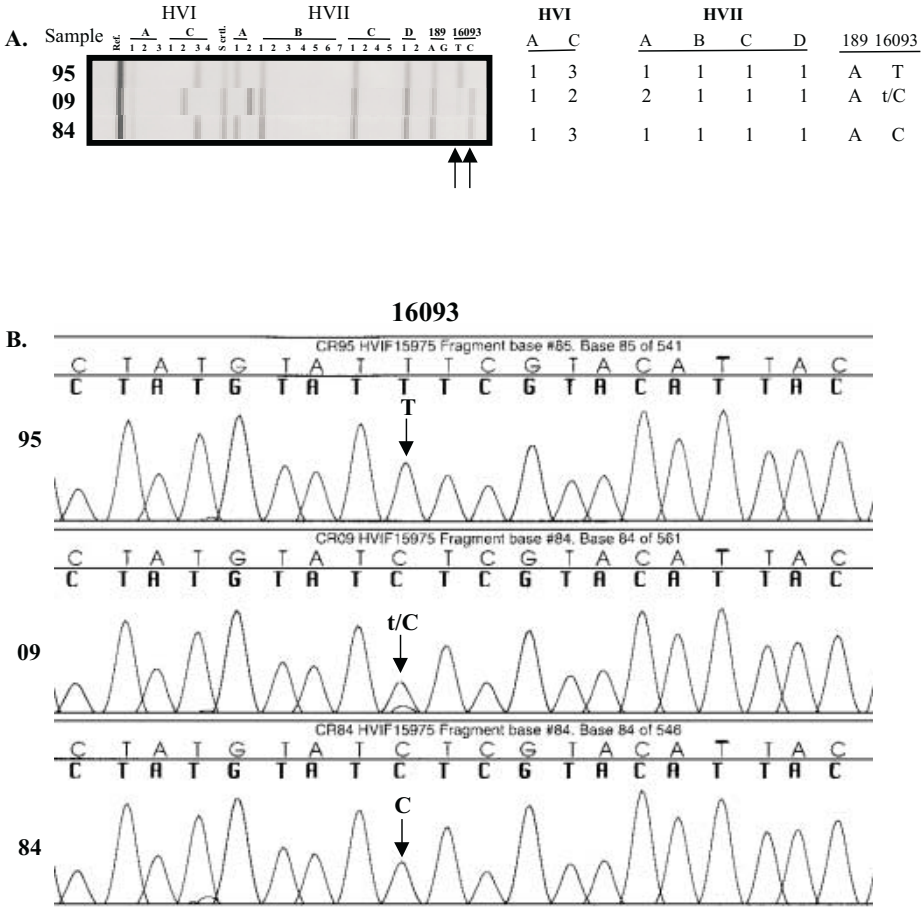


Figure 2. Mixture detection with SSO typing and direct sequencing. A) Mixture detection with the SSO strips at position 16093. Samples CR95, 09, and 84 show the three possible outcomes at this position. A single 16093T signal is observed in CR95, whereas a 16093C signal is seen with CR84. The presence of a mixture of "t/C" can be seen with SSO typing in CR09. B) Sequence information for CR09 confirms the mixture of bases seen with SSO typing. There were no additional positions that suggested mixed sequences for this sample. Independent amplifications and bi-directional sequence analysis were performed to prove the presence of mtDNA heteroplasmy at this position.

Discussion

Immobilized SSO probe typing of mtDNA sequences can offer significant benefit in many forensic identity testing efforts. Reassociation of mass disaster remains, single and mass grave analysis, and screening of large numbers of crime scene samples for inclusion or exclusion of a potential suspect(s) are some of the instances in which SSO typing can aid forensic testing. This strategy is highly beneficial in directing a forensic DNA analysis toward those specimens that possess the greatest evidentiary value and provides a benefit to forensic laboratories that are unable to perform direct sequence analysis (due to time and/or budget constraints) on large numbers of samples. Co-amplification of the HV regions and typing with the immobilized SSO probe strip assay expands the number of samples that can be processed in a given period of time. For this study, 105 samples were amplified and typed in less than two days. This duplex PCR amplification strategy is beneficial for conserving sample extract and is suitable on a range of forensic specimens, including shaft portions of single hairs, blood, and body tissues.

The population variation described here was applied to a separate study, which analyzed 18 skeletal elements recovered from a Croatian mass grave (our unpublished data). Fourteen of 18 samples provided amplification and typing results, and one bone sample extract provided a preliminary match following SSO analysis with one of the maternal references submitted. The corresponding mitotype was unique in this database of 105 randomly chosen Croatian individuals and subsequent sequence analysis confirmed the match between the unknown and maternal reference. Immobilized SSO typing followed by direct sequencing is therefore the recommended casework strategy for sample typing before determining matches.

Mixture interpretation and mtDNA heteroplasmy issues must be addressed for any mtDNA analysis in which more than one sequence is observed. Recent reports of the characterization of heteroplasmy in the non-coding region have suggested that the frequency of heteroplasmy is much higher in the normal population than what was once thought. Reported observations also indicate that the frequency of heteroplasmy varies across tissue types and may occur at some positions more frequently than others (16,25). Results from our study are consistent with previous findings. We observed mtDNA point heteroplasmy in 3 of the 105 blood samples by SSO typing and these results were confirmed by sequencing. One instance was observed in the HVI region and two instances were observed in the HVII region; sample CR09 showed the presence of a "t/C" mixture at position 16093 (Fig. 2), whereas CR02 showed "T/c" at position 195 and CR103 showed a roughly equal "T/C" mixture at 146. These results are consistent with the frequency of heteroplasmy detected in the HVII region by SSO analysis of 689 blood samples previously reported by Reynolds et al (16). Mixtures at both positions 146 and 195 have been reported in other studies and may be potential heteroplasmic hotspots in blood (16). Likewise, heteroplasmy at position 16093 has been reported previously and may be a potential hotspot, but possibly only when "C" is the major component (28). Among our samples, 16093 "C" was observed a total of seven times by SSO typing and a "t/C" mixture was clearly observed by both SSO typing and sequencing in one sample, CR09 shown in Figure 2. The remaining 6 samples might also be heteroplasmic but minor components were not detected by sequence analysis. The sensitivity of the SSO typing method is greater than that of sequencing, which ranges from 15%-25% depending on sequencing chemistry, position, direction, and dominant base. Therefore, if we use a more sensitive method of detection, such as denaturing gradient gel electrophoresis (28) or cloning analysis, heteroplasmy may be revealed.

The current linear array of SSO probes has been expanded to include two additional HVI regions (region D – position 16362, and region E – positions 16270 and 16278) for added discrimination. The presence of numerous polymorphic sites within the mtDNA coding region can offer increased power to mtDNA typing for forensic casework testing and may provide a means to distinguish common haplotypes in the control region. Further, our laboratory has the ability to routinely perform multiplex and megaplex amplification of >50 amplicons in a single reaction tube. Combining such an amplification strategy with the rapid and flexible linear array detection system for targeting additional sites within both the coding and non-coding portions of the mtDNA genome offers both forensic and clinical applications.

Acknowledgments

The authors would like to thank Dr. Henry Erlich, Karen Walker, and Michael Grow for scientific and intellectual contributions and laboratory assistance; Irena Drmić and Boja Režić for sample collection; and Leah Cottam for assistance with editing the manuscript. This project was supported by NIJ grant 96-IJ-CX-0028.

References

1 Huffine E, Crews J, Kennedy B, Bomberger K, Zinbo A. Mass identification of persons missing from the break-up of the former Yugoslavia: structure, function, and role of the International Commission on Missing Persons. Croat Med J 2001;42:271-5.

2 Primorac D, Anđelinović S, Definis-Gojanović M, Drmić I, Režić B, Baden M, et al. Identification of war victims from mass graves in Croatia and Bosnia and Herzegovina through the use of DNA typing and standard forensic methods. J Forensic Sci 1996;41:891-4.

3 Cann R, Stoneking M, Wilson A. Mitochondrial DNA and human evolution. Nature 1987;325:31-6.

4 Harihara S, Saitou N, Hirai M, Gojobori T, Park K, Misawa S, Ellepola S, et al. Mitochondrial DNA polymorphism among five Asian populations. Am J Hum Genet 1988;43:134-43.

5 Monsalve M, Hagelberg E. Mitochondrial DNA polymorphisms in Carib people of Belize. Proc R Soc Lond B Biol Sci 1997;264:1217-24.

6 Torroni A, Schurr T, Yang C, Szathmary E, Williams R, Schanfield M, Troup G, et al. Native American mitochondrial DNA analysis indicates that the Amerind and

the Nadene populations were founded by two independent migrations. Genetics 1992;130:153-62.

7 Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson A. African populations and the evolution of human mitochondrial DNA. Science 1991;253:1503-7.

8 Stoneking M, Hedgecock D, Higuchi R, Vigilant L, Erlich H. Population variation of human mtDNA control region sequences detected by enzymatic amplification and sequence-specific oligonucleotide probes. Am J Hum Genet 1991;48:370-82.

9 Melton T, Peterson R, Redd A, Saha N, Sofro A, Martinson J, et al. Polynesian genetic affinities with Southeast Asian populations as identified by mtDNA analysis. Am J Hum Genet 1995;57:403-14.

10 Melton T, Stoneking M. Extent of heterogeneity in mitochondrial DNA of ethnic Asian populations. J Forensic Sci 1996;41:591-602.

11 Melton T, Wilson M, Batzer M, Stoneking M. Extent of heterogeneity in mitochondrial DNA of European populations. J Forensic Sci 1997;42:437-46.

12 Melton T, Ginther C, Sensabaugh G, Soodyall H, Stoneking M. Extent of heterogeneity in mitochondrial DNA of sub-Saharan African populations. J Forensic Sci 1997;42:582-92.

13 Soodyall H, Jenkins T, Mukherjee A, du Toit E, Roberts D, Stoneking M. The founding mitochondrial DNA lineages of Tristan da Cunha Islanders. Am J Phys Anthropol 1997;104:157-66.

14 Vigilant L, Pennington R, Harpending H, Kocher T, Wilson A. Mitochondrial DNA sequences in single hairs from a southern African population. Proc Natl Acad Sci 1989;86:9350-4.

15 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. J Forensic Sci 1998;43:453-64.

16 Reynolds R, Walker K, Varlaro J, Allen M, Clark E, Alavaren M, et al. Detection of sequence variation in the HVII region of the human mitochondrial genome in 689 individuals using immobilized sequence-specific oligonucleotide probes. J Forensic Sci 2000;45: 1210-31.

17 Gabriel M, Huffine E, Ryan J, Holland M, Parsons T. Improved mtDNA sequence analysis of forensic remains using a "mini-primer set" amplification strategy. J Forensic Sci 2001;46:247-53.

18 Handt O, Richards M, Trommsdorff M, Kilger C, Simanainen J, Georgiev O, et al. Molecular genetic analyses of the Tyrolean ice man. Science 1994;264: 1775-8.

19 Handt O, Krings M, Ward R, Pääbo S. The retrieval of ancient human DNA sequences. Am J Hum Genet 1996;59:368-76.

20 Holland M, Fisher D, Roby R, Ruderman J, Bryson C, Weedn V. Mitochondrial DNA sequence analysis of human remains. Crime Lab Digest 1995;22:3-8.

21 Pääbo S. Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. Proc Natl Acad Sci 1989;86:1939-43.

22 Wilson M, Polanskey D, Butler J, DiZinno J, Replogle J, Budowle B. Extraction, PCR amplification and sequencing of mitochondrial DNA from human hair shafts. Biotech 1995;18:662-9.

23 Holland M, Parsons T. Mitochondrial DNA sequence analysis – validation and use of forensic casework. Forensic Sci Rev 1999;11:21-50.

24 Parsons T, Muniec D, Sullivan K, Woodyatt N, Alliston-Greiner R, Wilson M, et al. A high observed substitution rate in the human mitochondrial DNA control region. Nat Genet 1997;15:363-8.

25 Calloway C, Reynolds R, Herrin G, Anderson W. The frequency of heteroplasmy in the HVII region of mtDNA differs across tissue types and increases with age. Am J Hum Genet 2000;66:1384-97.

26 Walsh P, Metzger D, Higuchi R. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotech 1991;10:506-13.

27 Anderson S, Bankier A, Barrell B, deBruijn M, Coulson A, Drouin J, et al. Sequence and organization of the human mitochondrial genome. Nature 1981;290:457-65.

28 Tully L, Parsons T, Steighner R, Holland M, Marino M, Prenger V. A sensitive denaturing gradient-gel electrophoresis assay reveals a high frequency of heteroplasmy in hypervariable region I of the human mtDNA control region. Am J Hum Genet 2000;67:432-43.

Received: April 2, 2001
Accepted: May 4, 2001

Correspondence to:
Matthew Gabriel
Forensic Laboratory
San Mateo County Sheriff's Office
31 Tower Road
San Mateo, CA 94402-4097, USA
mgabriel@co.sanmateo.ca.us