Clinical Chemistry 53:2
195–203 (2007)

Mitochondrial DNA Mutation Detection by Electrospray Mass Spectrometry

Yun Jiang,[1] Thomas A. Hall,[1] Steven A. Hofstadler,[1*] and Robert K. Naviaux[2*]

Background: **Mitochondrial DNA (mtDNA) mutations cause a large spectrum of clinically important neurodegenerative, neuromuscular, cardiovascular, and endocrine disorders. We describe the novel application of electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI-FTICR MS) to the rapid and accurate identification of pathogenic mtDNA variants.**

Methods: **In a blinded study, we used ESI-FTICR MS to analyze 24 unrelated samples of total cellular DNA containing 12 mtDNA variants and compared the results with those obtained by conventional PCR-restriction fragment length polymorphism (PCR-RFLP) analysis and gel electrophoresis.**

Results: **From the 24-sample blinded panel, we correctly identified 12 of the samples as bearing an mtDNA variant and found the remaining 12 samples to have no pathogenic variants. The correlation coefficient between the 2 methods for mtDNA variant detection was 1.0; there were no false positives or false negatives in this sample set. In addition, the ESI-FTICR method identified 4 single-nucleotide polymorphisms (SNP) that had previously been missed by standard PCR-RFLP analysis.**

Conclusions: **ESI-FTICR MS is a rapid, sensitive, and accurate method for the identification and quantification of mtDNA mutations and SNPs.**

New analytic schemes based on biological mass spectrometry (MS)[3] are currently being developed for the characterization of various biomolecules, including proteins, nucleic acids, and metabolites. Compared with traditional methods for the characterization of DNA species, the MS platform has several potential advantages, including speed, sensitivity, and feasibility of end-to-end automation *(1–5)*. Recently, MS-based nucleic acid protocols have migrated out of MS labs for application in other disciplines, including human and microbial forensics, drug discovery, infectious-disease diagnostics, and biological warfare agent detection *(6–10)*. Because the masses of the 4 mononucleotides (dAMP, dTMP, dGMP, and dCMP) comprising DNA are known with great accuracy, highly precise measurements of mass can be used to derive a base composition (or a constrained list of base compositions). By taking into account the base complementarity of double-stranded DNA, investigators can further constrain the list of possible base compositions *(11)*. The "softness" of the ionization processes of both matrix-assisted laser desorption/ionization and electrospray ionization (ESI) has enabled the ionization of PCR products for detection by MS. ESI Fourier transform ion cyclotron resonance MS (ESI-FTICR MS) and ESI time-of-flight MS (ESI-TOF MS) have routinely been used in our laboratory for PCR product analysis, a regular component of a biosensor funded by the Defense Advanced Research Projects Agency *(12)*. For example, we recently applied this approach to rapidly genotype 217 isolates of *Acinetobacter* to determine the epidemiology and clonality of outbreaks in soldiers and civilians in Iraq and Kuwait *(13)*. The analytic power of the ESI MS technique is quite general and the technique in principle can be applied to the rapid assay of single-nucleotide polymorphisms (SNPs) and mutation in any gene for which disease-causing variants and informative polymorphisms are known. We describe the specific application of an ESI MS-based method for rapidly genotyping individuals for

[1] Ibis Biosciences, a division of Isis Pharmaceuticals, Carlsbad, CA.
[2] Departments of Medicine and Pediatrics, University of California, San Diego, The Mitochondrial and Metabolic Disease Center, San Diego, CA.
* Address correspondence to these authors at: fax (to R.K.N.) 619-543-7868; e-mail Naviaux@ucsd.edu or (to S.A.H.) 760-603-4653; e-mail SHofstad@isisph.com.

Received June 12, 2006; accepted November 7, 2006.
Previously published online at DOI: 10.1373/clinchem.2006.074823

[3] Nonstandard abbreviations: MS, mass spectroscopy; ESI, electrospray ionization; ESI-FTICR MS, electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry; ESI-TOF MS, electrospray ionization time-of-flight mass spectrometry; SNP, single-nucleotide polymorphism; mtDNA, mitochondrial DNA; PCR-RFLP, PCR-restriction fragment length polymorphism; MELAS, mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes.

mitochondrial DNA (mtDNA) variants associated with mtDNA-based diseases.

Twenty-four blinded patient samples were analyzed using the methods described below. Base composition signatures were derived from PCR products amplified directly from human DNA samples and compared with base compositions expected from nonmutated human mtDNA. ESI-FTICR-MS correctly identified all 12 samples that contained known pathogenic mutations in mtDNA. In addition, the ESI-MS based approach identified unexpected SNPs in 4 samples; these positions were not specifically screened using the traditional restriction fragment length polymorphism method and were thus "invisible" to the standard screening method.

Materials and Methods

REAGENTS, MATERIALS, AND ANALYTIC SAMPLES

Methanol, piperidine, and imidazole were obtained from Sigma-Aldrich and used without additional purification. Microseal™ 96-well skirted V-bottom polypropylene microplates (MJ Research) were used for PCR amplifications and all plate-based purifications. Cellular DNA samples from the whole blood of anonymous donors were purified by standard methods (Puregene; Gentra Systems). DNA concentrations were between 200 and 300 mg/L. Twelve of the 24 samples contained 1 of 7 pathogenic variants at various levels of heteroplasmy: A3243G, T3271C, A8344G, T8356C, T8993G, T8993C, and G11778A. Twelve other samples with no mutation at these sites were used as controls. The genotype of each sample was blinded before MS analysis. Results were compared with those obtained by conventional PCR-restriction fragment length polymorphism (PCR-RFLP) analysis *(14, 15)*. The overall flow of operations for variant detection by ESI MS is illustrated in Fig. 1.

PRIMER SELECTION

A set of 444 human mitochondrial genomes were obtained from GenBank and aligned against the Cambridge reference sequence *(16, 17)* represented in Fig. 2. An additional set of 524 human mitochondrial sequences containing only the sequence of the noncontrol region (coordinates 577–16023) was obtained from MitoKor, Inc. *(18)*. For this study, we designed 4 primer pairs to amplify mtDNA regions that contained 7 point variants associated with mitochondrial and/or metabolic diseases (Fig. 2; see Table 1 in the Data Supplement that accompanies the online version of this article at http://www.clinchem.org/content/vol53/issue2). An additional 9 primer pairs were designed to amplify mtDNA surrounding other common pathogenic mtDNA variants located in tRNA genes (Fig. 2; see Table 1 in the online Data Supplement). For each of the 13 primer pairs, we created a database of theoretical PCR products and their molecular masses, based on the 968 human mtDNA sequences in

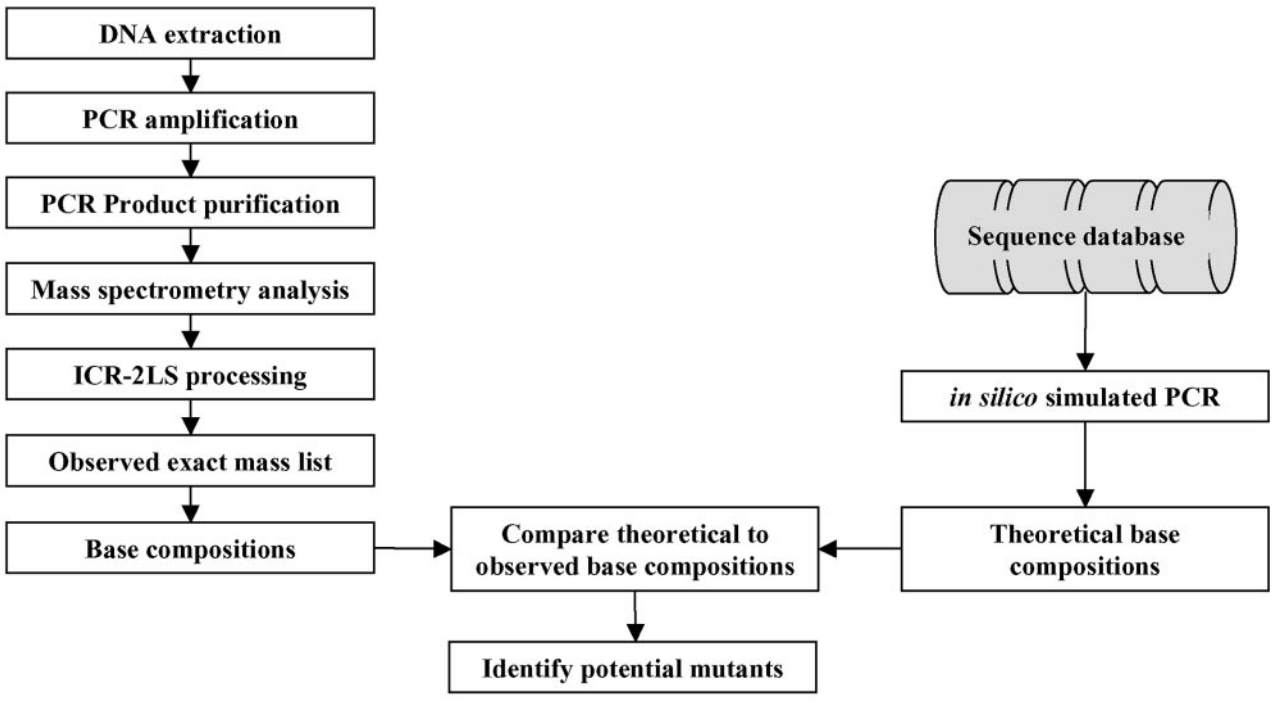

Fig. 1. The ESI-FTICR MS protocol of mtDNA mutation analysis.

After DNA extraction, the PCR amplifies DNA regions of interest. Amplicons are then analyzed by MS, and MS data are converted to a list of exact masses with ICR2LS software *(22)* with sufficient accuracy to compute the base compositions of the PCR products. Observed base compositions are compared with a database of theoretical PCR products derived from a sequence database. A mutation or SNP is recognized as a base composition differing from that of the most commonly observed theoretical amplicon composition.

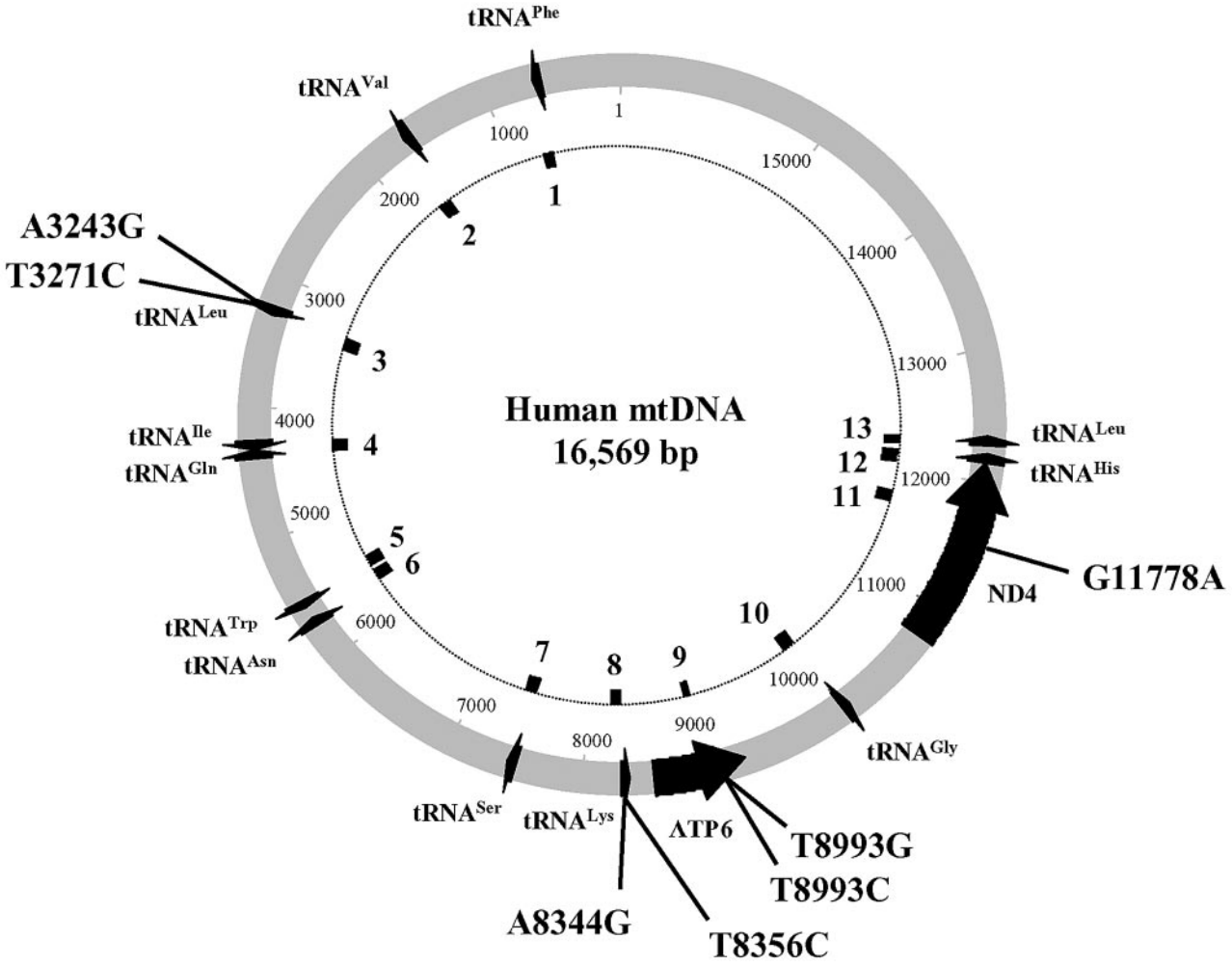

Fig. 2. Morbidity map of human mtDNA.

The 13 pairs of PCR primers used in this study span the majority of tRNA and coding mutations known to be associated with mitochondrial disease and are indicated in black on the inner circle of the map (numbered 1–13; see Table 1 in the online Data Supplement). The 7 most common mutations are highlighted in *bold* around the outer circle.

the alignment and assessed the frequency of each expected PCR product (see Table 1 in the online Data Supplement). The frequency of occurrence of each base composition was defined as the frequency that each expected base composition occurred over the set of 968 sequences in the database. For one region (tRNA-Phe, primer pair 1; see Table 1 in the online Data Supplement), the amplified region contained 15 bases 5′ of the first sequenced base represented in the MitoKor sequence set. For this reason, expected base-composition frequencies for tRNA-Phe were based on only the 444 mtDNA genomes from GenBank. Because human mtDNA outside of the hypervariable D-loop regions is highly conserved *(19)*, usually only a very small number of variants occur within any given PCR product. The frequency of the most common base composition in the database of 968 sequences was >90% for all regions except tRNA-Leu-2.2 (primer pair 13, amplification coordinates 12294–12321; see Table 1 in the online Data Supplement). We devised an assay amenable to high-throughput automation that screened for heteroplasmic base-composition states consistent with 1 or more mtDNA variants relative to the most common expected PCR product for each PCR primer pair.

To verify the presence of mtDNA mutants within the 6 primary test positions (A3243G, T3271C, A8344G, T8356C, T8993G/C, G11778A), we developed a panel of 6 primer pairs that narrowly target the position of interest (see Table 2 in the online Data Supplement). Each of these primer pairs targets an amplified region of 1–4 highly conserved bases surrounding the position of interest (see Table 2 in the online Data Supplement).

In addition to the 7 point mutations described above, the ESI-FTICR MS methods we describe have the potential for detecting most of the pathogenic and polymorphic

mtDNA changes currently catalogued in the Mitomap database (http://www.mitomap.org).

PCR CONDITIONS

All PCR amplifications were carried in 50-μL reaction volumes in a 96-well microtiter plate format. The PCR mixture consisted of 1× PCR buffer II (Applied Biosystems), 1.5 mmol/L Mg^{2+} (Applied Biosystems), 0.4 mol/L betaine (Sigma-Aldrich), 0.2 mmol/L of each deoxynucleoside triphosphate (Stratagene), 250 nmol/L of each primer, and 4 U AmpliTaq Gold (Applied Biosystems). Total cellular DNA from each sample (1–1.5 ng) was used as the template in the PCR. Primer pair sequences are shown in Table 1 of the online Data Supplement. We used a "touch down" thermal-cycling approach in the PCR protocol. The cycling conditions consisted of (*1*), an initial denaturation step at 95 °C for 10 min, followed by (*2*), denaturation at 95 °C for 20 s, (*3*), annealing at 60 °C for 20 s, and (*4*), elongation at 72 °C for 30 s. Steps 2–4 were repeated for 29 cycles, with the annealing temperature decreasing 1 °C with each cycle until the annealing temperature reached 55 °C. The final step was elongation at 72 °C for an additional 4 min. Thermal-cycling parameter values were identical for the variant-targeted primer pairs (see Table 2 in the online Data Supplement) except that the annealing temperature was fixed at 53 °C, and 35 cycles were carried out. An MJ Dyad thermocycler (MJ Research) was used for all PCR experiments.

PURIFICATION OF PCR PRODUCTS

PCR products were thoroughly purified and desalted before ESI MS analysis, as previously described *(20)*. This step must precede ESI MS analysis because PCR salts and buffer components have a deleterious effect on the ESI process. Even small amounts of salts (<1 μmol/L) will significantly reduce ESI sensitivity, owing to the appearance of multiple cation adducts in the mass spectra. The protocol is based on the weak anion-exchange method, in which amplified DNA is bound to a weak anion-exchange resin; unconsumed deoxynucleoside triphosphates, salts, and other low–molecular-weight species that could interfere with subsequent ESI MS analysis are removed by rinsing the resin with a solution of 40 mmol/L NH_4HCO_3 and 200 mL/L methanol. Rinsing the resin with 25 μL of 1 mol/L NH_4OH eluted the final purified and desalted amplicons. The final electrospray buffer contained 250 mL/L methanol and 25 mmol/L piperidine/imidazole *(21)*. Purification was carried out in an automated and parallel format. A single 96-well plate of PCR products can be desalted in <30 min.

ESI-FTICR MS ANALYSIS OF PCR PRODUCTS

We used a modified Bruker Daltonics Apex II 70e actively shielded ESI-FTICR MS instrument in negative-ionization mode. We used a 600-MHz Pentium II data station running XMASS software (version 7.0.4; Bruker Daltonics) under a Windows NT 4.0 operating system (Microsoft) to control all aspects of pulse sequence and data collection. The FTICR data station triggered a CTC HTS PAL autosampler (LEAP Technologies) to extract a 15-μL sample directly from the 96-well microtiter plate and inject it into a 10-μL sample loop integrated with a fluid-handling system that supplies sample to the ESI source at a flow rate of 100 μL/h *(12)*. Application of an ESI source potential of −6000V generated a stable electrospray plume, and a countercurrent flow of dry nitrogen gas assisted in the desolvation process. Spectral acquisition was performed in continuous duty-cycle mode, whereby ions accumulated in an external hexapole reservoir concurrent with ion detection in the trapped-ion cell *(22)*. To improve the signal-to-noise ratio, we coadded 32 scans for a total data-acquisition time of 74 s. We used custom processing software developed in-house to analyze the raw mass spectra. First, ICR2LS software *(9, 12, 20)* was used to deconvolute the FTICR spectra into monoisotopic neutral masses for all of the observed signals. From these high-precision monoisotopic mass measurements, we calculated base compositions for each amplicon pair *(9, 12, 20)*.

ESI-TOF MS ANALYSIS OF PCR AMPLICONS

We used a Bruker Daltonics MicroTOF instrument in the comparison of the ESI-TOF MS and ESI-FTICR MS methods (illustrated in Fig. 3). The TOF instrument was equipped with the same automated sample-handling and fluidics capabilities as described above; ESI conditions were identical. Each individual scan on the TOF instrument comprised 75 000 data points acquired at 2 GHz after a 37-μs external ion-accumulation event. For each spectrum, 980 000 scans were coadded to maintain the same 74-s data-acquisition time as used for the FTICR analysis. XMASS software and software written in-house were used to process the data.

Results

MS DETECTION OF SNPS IN MTDNA

The 24 blinded DNA samples were analyzed with the 13 broad-range PCR primer pairs. Table 1 shows the base compositions determined for the 13 amplified regions of each sample. Several uncommon base compositions (bold type in Table 1) were observed in different regions of different samples. Table 2 summarizes the mutations, their locations in mtDNA, and the levels of heteroplasmy observed in the 24 samples.

We subsequently verified the mtDNA mutations detected in the regions of interest by means of 6 SNP-targeted primers. Table 3 lists the base compositions for the PCR products that we generated from each sample with these primer pairs. The uncommon base compositions are in boldface. Samples 1, 4, 22, 23, and 24 had variant A3243G, which is associated with mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes (MELAS). Sample 6 contained variant T3271C,

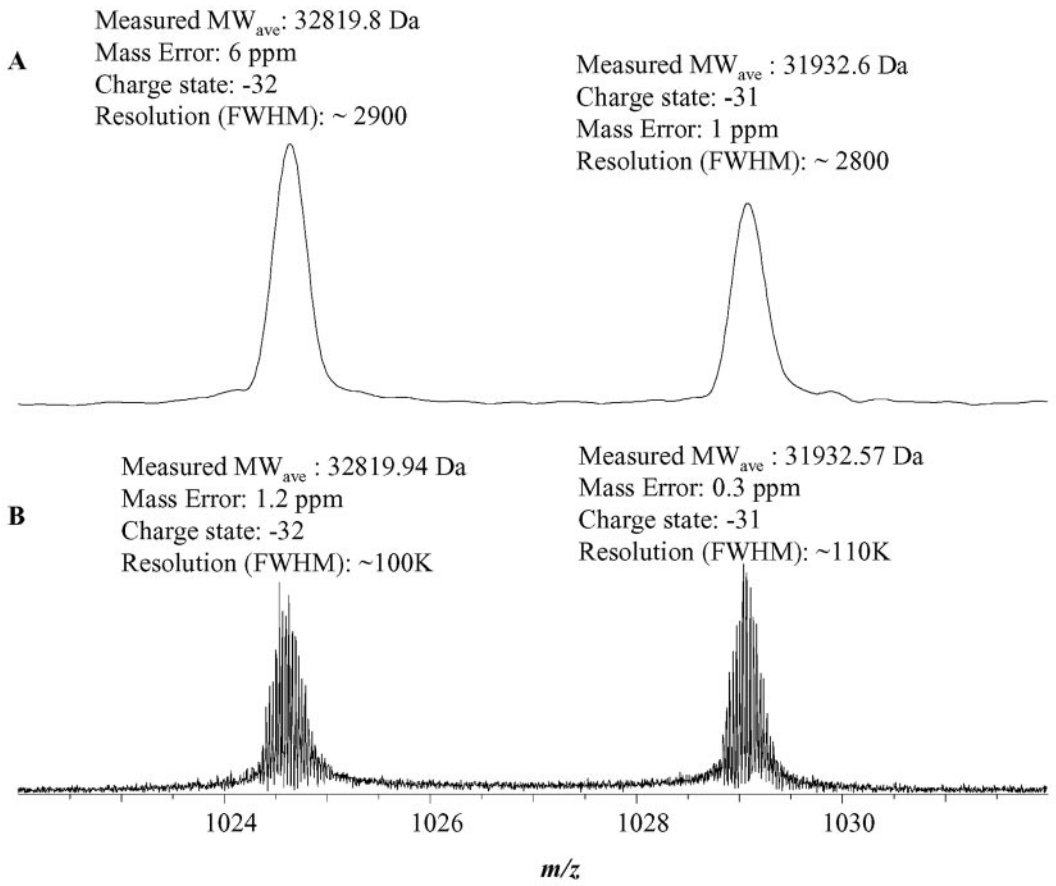

Fig. 3. Comparison of the ESI-FTICR MS and ESI-TOF MS analyses.

Sample 2 was amplified with primer pair tRNA-Phe. Equivalent aliquots were analyzed with (*A*) ESI-TOF MS, and (*B*) ESI-FTICR MS. For clarity, only a narrow region of each spectrum, representing a single charge state of each strand of the amplicon, is shown. FWHM indicates full width at half maximum.

which is also associated with MELAS. Samples 2, 8, 10, and 19 contained variant T8993G, sample 15 contained T8993C, and sample 12 contained G11778A. Samples 3, 5, 7, 9, 11, 13, 14, 16, 17, 18, 20, and 21 lacked any of the 7 pathogenic mutations.

We unblinded the samples after compiling the ESI-FTICR MS results. Fig. 1 in the online Data Supplement compares the results of the ESI MS-based method and the traditional PCR-RFLP method. PCR-RFLP screening of these regions revealed 12 samples positive for a variant and 12 negative samples. Of note is that the ESI-FTICR MS methodology correctly identified all 12 positive samples and all 12 negative samples with no false negatives and no false positives. For samples in which heteroplasmy was detected at a given locus, the extent of heteroplasmy was measured both by relative peak heights in the mass spectra and by conventional PCR-RFLP gel analysis. The correlation between the 2 methods in detecting the extent of heteroplasmy was 87% (see Fig. 1 and Table 3 in the online Data Supplement).

In addition to the 12 samples with 1 of the known 7 pathogenic point mutations, samples 5, 9, 18, and 21 possessed additional sequence variation in region 4243–4363 (primer pair 4, tRNA-Ile-Gln.2) or region 11751–11866 (primer pair 11, ND4) that PCR-RFLP did not identify (Table 2). To further verify the ESI-FTICR MS results, we directly sequenced the tRNA-Ile-Gln.2 regions of samples 5 and 21 and the ND4 regions of samples 9 and 18. We found the following SNPs by direct sequencing: sample 5, T4336C (www.mitomap.org); sample 9, A11812G (GenBank accession no. AY339569); sample 18, T11794C (AY713992); and sample 21, C4312T (AF346999) (Table 2). In all cases, the variants verified by sequencing were consistent with those identified by ESI-FTICR MS.

Although we performed these studies with a research-grade ESI-FTICR MS instrument, recent advances in ESI-TOF MS technology have led to the development of benchtop instruments that have the requisite sensitivity and mass accuracy to identify PCR amplicons. This capability is illustrated in Fig. 3. We amplified sample 2 with primer pair tRNA-Phe, split the purified amplicon into 2 aliquots, and analyzed the aliquots by ESI-TOF MS or ESI-FTICR MS. The resulting spectra (Figs. 3A and 3B, respectively) illustrate some of the key differences in performance between the ESI-FTICR and ESI-TOF methods. The most striking difference is the presence of "fine structure" in the FTICR spectrum that is absent from the TOF spectrum. The constituent elements comprising the

Table 1. Base compositions of PCR products amplified from 24 blinded samples with 13 broad-range primer pairs.

Primer pair	Description	Amplified region	Wild-type base counts	Observed base counts	Samples (% heteroplasmy)
1	tRNA-Phe	539–643	A36 G12 C38 T19	A36 G12 C38 T19	1–24
2	tRNA-Val	1556–1674	A39 G25 C27 T28	A39 G25 C27 T28	1–24
				A40 G20 C36 T34	1 (50%), 2, 3, 4 (85%), 5, 6 (57%), 7, 21, 22 (15%), 23 (76%), 24 (5%)
3	tRNA-Leu-1	3214–3343	A40 G20 C36 T34	**A39 G21 C36 T34**	1 (50%), 4 (15%), 22 (85%), 23 (24%), 24 (95%)
				A40 G20 C37 T33	6 (43%)
4	tRNA-Ser.2	7411–7528	A44 G19 C33 T22	A44 G19 C33 T22	1–24
				A43 G19 C26 T32	1–4, 6–20, 22–24
5	tRNA-Ile-Gln.2	4243–4362	A43 G19 C26 T32	**A43 G19 C27 T31**	5 (100%)
				A43 G19 C25 T33	21 (100%)
6	tRNA-Asn	5602–5731	A46 G15 C39 T30	A46 G15 C39 T30	1–24
7	tRNA-Trp.2	5467–5599	A43 G16 C34 T40	A43 G16 C34 T40	1–24
8	tRNA-Lys.2	8283–8399	A41 G15 C32 T29	A41 G15 C32 T29	1–24
9	tRNA-Gly	9965–10097	A45 G14 C27 T47	A45 G14 C27 T47	1–24
10	tRNA-His	12096–12230	A43 G19 C37 T36	A43 G19 C37 T36	1–14, 16–22, 23 (5%), 24
				A44 G18 C37 T36	15 (100%), 23 (95%)
11	tRNA-Leu-2.2	12203–12352	A33 G12 C24 T23	A33 G12 C24 T23	1–24
				A32 G13 C39 T32	1–8, 10, 11, 13–17, 19–24
				A31 G14 C39 T32	9 (100%)
12	ND4	11751–11866	A32 G13 C39 T32	**A33 G12 C39 T32**	12 (100%)
				A32 G13 C40 T31	18 (100%)
				A19 G10 C27 T15	1, 2 (8%), 3–7, 9, 10 (8%), 11–14, 16–18, 20–24
13	ATP6	8963–9033	A19 G10 C27 T15	**A19 G11 C27 T14**	2 (92%), 8 (100%), 10 (92%), 19 (100%)
				A19 G10 C28 T14	15 (100%)

Bolded compositions are rare variants further analyzed with confirmatory short-range PCR primers.

amplicon are not monoisotopic (e.g., 98.1% of the carbon atoms are ^{12}C, and ~1.1% are ^{13}C); there exists a naturally occurring isotope distribution (to which $^{14}N/^{15}N$ and $^{16}O/^{18}O$ also contribute). Species containing different numbers of "light" and "heavy" isotopes differ in molecular mass by ~1 atomic mass unit. These small mass differences are resolved by the FTICR method (>100 000 full width at half maximum resolution; Fig. 3B) but not by the TOF method (Fig. 3A). The TOF spectrum reflects only the width of the isotopic envelope, not the individual isotopomers. The centroid of the unresolved isotope envelope in the TOF spectrum can still be used to derive an accurate mass measurement of the analyte, albeit not with the same mass accuracy as the FTICR method. The experimentally determined average molecular masses for the forward and reverse strands of the amplicon were 32 819.94 and 31 932.57 Da, respectively, from the FTICR measurement (Fig. 3B) and 32 819.8 Da and 31 932.6 Da from the TOF measurement (Fig. 3A). The average mass-measurement errors are 0.8 ppm and 4 ppm for the FTICR and TOF spectra, respectively. In general, mass-measurement errors of <20–25 ppm (on both strands) are required to yield an unambiguous base composition. In this example, the correct base composition (A19 G38 C12 T36) is derived from either the FTICR or TOF measurements. The closest base composition in mass (A50 G31 C19 T5) would require a mass-measurement error of >20 ppm on both strands, and the base composition is inconsistent with and significantly different from all database entries for this primer pair. Because the routinely obtained mass-measurement errors of the FTICR and TOF methods are, conservatively, 5 ppm and 15 ppm, respectively, the correct base composition is unambiguously determined on either platform. It is notable that the sensitivities of the FTICR and TOF platforms are comparable for this application. Given the high signal-to-noise ratios of the spectra in Fig. 3, this result suggests that shorter acquisition intervals could be used to further enhance the throughput of the system. We are currently validating this benchtop ESI-TOF MS platform for high-throughput analysis of multiplexed PCR reactions to characterize the hypervariable mtDNA regions for forensics purposes.

Discussion

ESI-FTICR MS, a rapid and accurate method of mtDNA mutation and SNP detection, does not require precisely targeting a specific variant to identify a deviation from the wild-type sequence. Furthermore, this approach allows

Table 2. mtDNA mutations, SNPs, and heteroplasmy detected by ESI MS.

Sample	Mutation	SNP	Percent heteroplasmy
1	A3243G		50
2	T8993G		100
3	No mutation observed		
4	A3243G		10
5		T4336C	100
6	T3271C		44
7	No mutation observed		
8	T8993G		100
9		A11812G	100
10	T8993G		92
11	No mutation observed		
12	G11778A		100
13	No mutation observed		
14	No mutation observed		
15	T8993C		100
16	No mutation observed		
17	No mutation observed		
18		T11794C	100
19	T8993G		98
20	No mutation observed		
21		C4312T	100
22	A3243G		81
23	A3243G		22
24	A3243G		94

the probing of relatively large regions of genome for the presence of substitutions, SNPs, insertions, and deletions that were not originally targeted. Because this broad-survey assay determines the base composition of a PCR product and not the sequence, it does not conclusively determine the exact position of specific mtDNA variants. The purpose of the assay method we have described is primarily to detect sequence variants and heteroplasmic states (the presence of 2 or more mitochondrial variants existing simultaneously within an individual), either of which may be associated with a metabolic disease. In the event that a sample is flagged as possibly having 1 or more characterized mutations, the sample can be further investigated via sequencing. Compared with traditional electrophoresis gel–based methods such as PCR-RFLP, the ESI MS protocol has several advantages: (*a*) The ESI MS-based method can detect both expected and unexpected variants. In contrast, traditional PCR-RFLP methods can detect only expected (i.e., targeted) mutations. Traditional methods answer the question, "Is this specific mutation present in the sample?" The protocol based on ESI MS answers the question, "Which mutations are present in the sample?" (*b*) The MS instrument is a universal detector. It can detect multiple mtDNA variants simultaneously, whereas PCR-RFLP methods require different restriction enzyme digestions for different mutations. (*c*) Data interpretation can be done in an automated and standardized format, which is difficult for traditional methods such as PCR-RFLP. (*d*) Finally, samples can be screened in a high-throughput mode.

All analytic steps in the ESI MS protocol, including PCR amplification, post-PCR sample purification, and MS analysis, were performed in an automated and parallel format, resulting in a significant reduction in the chances of contamination and operator error. The time requirements for this method are as follows: PCR amplification, 1 h 40 min; post-PCR purification, 30 min per 96-well plate; MS analysis, 1–2.5 h per 96-well plate. The capability for performing all of these processes in high-throughput mode enables screening of large numbers of samples for both expected and unexpected variants and facilitates searches for associations between mtDNA variants and diseases. The ESI-MS method can be used as a rapid, low-cost tool to screen samples before the time-consuming and costly

Table 3. Base compositions of PCR products amplified from 24 blinded samples with 6 confirmatory, short-range PCR primer pairs.

Primer pair	Wild-type base counts	Observed base counts	Associated mutation	Samples (% heteroplasmy)
				1 (50%), 2,3,4 (90%), 5–21,22 (19%), 23 (78%), 24 (6%)
3243	A22 G14 C11 T16	A22 G14 C11 T16		
		A21 G15 C11 T16	A3243G	1 (50%), 4 (10%), 22 (81%), 23 (22%), 24 (94%)
				1–5, 6 (56%), 7–24
3271	A18 G8 C14 T18	A18 G8 C14 T18		
		A18 G8 C15 T17	T3271C	6 (44%)
8344	A24 G8 C10 T17	A24 G8 C10 T17		1–24
8356	A26 G7 C13 T14	A26 G7 C13 T14		1–24
		A12 G7 C18 T9		1, 3–7, 9, 10 (8%),11–14,16–18,19 (2%),20–24
8993	A12 G7 C18 T9	**A12 G8 C18 T8**	T8993G	2 (100%), 8 (100%), 10 (92%), 19 (98%)
		A12 G7 C19 T8	T8993C	15 (100%)
				1–11,13–24
11 778	A15 G6 C17 T12	A15 G6 C17 T12		
		A16 G5 C17 T12	G11778A	12 (100%)

Bolded base compositions corresponded to known pathogenic variants.

process of direct DNA sequencing. In most cases, knowing that a mutation or SNP is present in a region of interest (even when the exact coordinates of the mutation are not known), and knowing the type of nucleotide substitution (e.g., A→T) are the key data for discriminating a typical patient sample from a sample with 1 or more abnormalities requiring additional investigation. Moreover, in the relatively rare instance when sequencing is required to determine the exact coordinates of a nucleotide substitution, prescreening by the ESI MS methods allow much smaller regions to be sequenced, because the mass measurement and base-composition analysis provide the approximate location of the SNP.

The estimated reagent costs (for primers, polymerase, deoxynucleoside triphosphates, desalting columns, buffers, and so on) for this work in a high-throughput modality are a few dollars per sample, and the time to a preliminary answer is 4–5 h. In comparison, a PCR-RFLP method typically requires 2 days to analyze a batch of 12 DNA samples for 7 different mtDNA variants, and 24 samples for this variant set requires 4 days. The enzyme digestion and gel electrophoresis steps required by the PCR-RFLP method are not easily automated; these steps thus create a bottleneck that limits the number of samples that can be analyzed simultaneously. ESI MS eliminates this bottleneck and avoids the extra costs of restriction enzymes, agarose gels, and toxic fluorescent dyes. Automated DNA sequencing cannot reliably detect heteroplasmy at levels of <30% *(23)*, and several commercial facilities currently charge ~$20 for single-pass sequencing. Therefore, not only is direct DNA sequencing more expensive than ESI MS, mtDNA sequencing is not reliable for detecting <30% heteroplasmy.

In addition to significantly reducing the equipment costs associated with the ESI-FTICR MS platform, full migration of this assay to the benchtop ESI-TOF MS platform requires significantly less space and little or no MS expertise. Multiplexing primer pairs and employing shorter MS acquisition intervals could enable surveys of up to several hundred samples a day for key mtDNA mutations.

The authors thank Jared Drader (Ibis) and Gordon Anderson (Pacific Northwest National Laboratory) for their work on the monoisotopic relative molecular mass calculations, and we thank Fred Hajjar (MassSpectra) for his work on automated base-composition calculation. Support for this work was provided by generous gifts to R.K.N. from Mrs. Dorothy R. Engs, the UCSD Foundation Christini Fund, the Lored Foundation, and the William Wright Family Foundation.

References

1. Doktycz MJ, Hurst GB, Habibi-Goudarzi S, McLuckey SA, Tang K, Chen CH, et al. Analysis of polymerase chain reaction-amplified DNA products by mass spectrometry using matrix-assisted laser desorption and electrospray: current status. Anal Biochem 1995; 230:205–14.
2. Hurst GB, Doktycz MJ, Vass AA, Buchanan MV. Detection of bacterial DNA polymerase chain reaction products by matrix-assisted laser desorption/ionization mass spectrometry. Rapid Commun Mass Spectrom 1996;10:377–82.
3. Krahmer MT, Johnson YA, Walters JJ, Fox KF, Fox A, Nagpal M. Electrospray quadrupole mass spectrometry analysis of model oligonucleotides and polymerase chain reaction products: determination of base substitutions, nucleotide additions/deletions, and chemical modifications. Anal Chem 1999;71:2893–900.
4. Naito Y, Ishikawa K, Koga Y, Tsuneyoshi T, Terunuma H. Molecular mass measurement of polymerase chain reaction products amplified from human blood DNA by electrospray ionization mass spectrometry. Rapid Commun Mass Spectrom 1995;9: 1484–6.
5. Naito Y, Ishikawa K, Koga Y, Tsuneyoshi T, Terunuma H, Arakawa R. Genetic diagnosis by polymerase chain reaction and electrospray ionization mass spectrometry: detection of five base deletion from blood DNA of a familial adenomatous polyposis patient. J Am Soc Mass Spectrom 1997;8:737–42.
6. Hofstadler SA, Griffey RH. Analysis of noncovalent complexes of DNA and RNA by mass spectrometry. Chem Rev 2001;101:377–90.
7. Ecker DJ, Sampath R, Blyn LB, Eshoo MW, Ivy C, Ecker JA, et al. Rapid identification and strain-typing of respiratory pathogens for epidemic surveillance. Proc Natl Acad Sci U S A 2005;102: 8012–7.
8. Hall TA, Budowle B, Jiang Y, Blyn L, Eshoo M, Sannes-Lowery KA, et al. Base composition analysis of human mitochondrial DNA using electrospray ionization mass spectrometry: a novel tool for the identification and differentiation of humans. Anal Biochem 2005;344:53–69.
9. Van Ert MN, Hofstadler SA, Jiang Y, Busch JD, Wagner DM, Drader JJ, et al. Mass spectrometry provides accurate characterization of two genetic marker types in Bacillus anthracis. Biotechniques 2004;37:642–4.
10. Sampath R, Hofstadler SA, Blyn LB, Eshoo MW, Hall TA, Massire C, et al. Rapid identification of emerging pathogens: coronavirus. Emerg Infect Dis 2005;11:373–9.
11. Muddiman DC, Anderson GA, Hofstadler SA, Smith RD. Length and base composition of PCR-amplified nucleic acids using mass measurements from electrospray ionization mass spectrometry. Anal Chem 1997;69:1543–9.
12. Hofstadler SA, Sampath R, Blyn LB, Eshoo MW, Hall TA, Jiang Y, et al. TIGER: the universal biosensor. Int J Mass Spectrom 2005;242:23–41.
13. Ecker JA, Massire C, Hall TA, Ranken R, Pennella TT, Agasino Ivy C, et al. Identification of acinetobacter species and genotyping of *acinetobacter baumannii* by multilocus PCR and mass spectrometry. J Clin Microbiol 2006;44:2921–32.
14. Yoneda M, Tanno Y, Tsuji S, Attardi G. Detection and quantification of point mutations in mitochondrial DNA by PCR. Methods Enzymol 1996;264:432–41.
15. Smith ML, Hua XY, Marsden DL, Liu D, Kennaway NG, Ngo KY, et al. Diabetes and mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes (MELAS): radiolabeled polymerase chain reaction is necessary for accurate detection of low percentages of mutation. J Clin Endocrinol Metab 1997;82:2826–31.
16. Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, et al. Sequence and organization of the human mitochondrial genome. Nature 1981;290:457–65.
17. Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. Reanalysis and revision of the Cambridge reference

sequence for human mitochondrial DNA. Nat Genet 1999;23: 147.

18. Hernstadt C, Elson JL, Fahy E, Preston G, Turnbull DM, Anderson C, et al. Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups. Am J Hum Genet 2002;70:1152–71.

19. Saccone C, Gissi C, Lanave C, Larizza A, Pesole G, Reyes A. Evolution of the mitochondrial genetic system: an overview. Gene 2000;261:153–9.

20. Jiang Y, Hofstadler SA. A highly efficient and automated method of purifying and desalting PCR products for analysis by electrospray ionization mass spectrometry. Anal Biochem 2003;316:50–7.

21. Greig M, Griffey RH. Utility of organic bases for improved electrospray mass spectrometry of oligonucleotides. Rapid Commun Mass Spectrom 1995;9:97–102.

22. Senko MW, Hendrickson CL, Emmett MR, Shi SDH, Marshall AG. External accumulation of ions for enhanced electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. J Am Soc Mass Spectrom 1997;8:970–6.

23. Hancock DK, Tully LA, Levin BC. A standard reference material to determine the sensitivity of techniques for detecting low-frequency mutations, SNPs, and heteroplasmies in mitochondrial DNA. Genomics 2005;86:446–61.

Suppl Table 1, ONLINE ONLY

Table 1. Base compositions of PCR products amplified from 24 blinded samples with 13 broad-range primer pairs.

Primer pair	Description	Amplicon coordinates	Amplified nucleotides	Primer sequences (5′→3′)	Normal population survey results[a]	
					Variant base compositions, n	Wild-type base compositions, %
1	tRNA-Phe	539–643	562–620	TAC CCC GAA CCA ACC AAA CCC CA TAT GGG GTG ATG TGA GCC CGT CT	11	91
2	tRNA-Val	1556–1674	1584–1648	CAA GTC GTA ACA TGG TAA GTG TAC TGG A TAG CTC AGA GCG GTC AAG TTA AGT TG	5	98.5
3	tRNA-Leu-1	3214–3343	3243–3316	TCC AAG AAC AGG GTT TGT TAA GAT GGC AG TGG GTA CAA TGA GGA GTA GGA GGT TGG	7	97.3
4	tRNA-Ile-Gln.2	4243–4362	4269–4336	TGC ATT CCC CCT CAA ACC TAA GAA AT TTC TCA GGG ATG GGT TCG ATT CTC AT	7	96.2
5	tRNA-Trp.2	5467–5599	5494–5568	TCA CGC TAC TCC TAC CTA TCT CCC CTT TGG GCA GTC CTT AGC TGT TAC AGA AAT TAA G	5	96.4
6	tRNA-Asn	5602–5731	5626–5703	TCC ACT CTG CAT CAA CTG AAC GC TGT AGA TTG AAG CCA GTT GAT TAG GGT G	6	95.8
7	tRNA-Ser.2	7411–7528	7443–7498	TCC CAT TCG AAG AAC CCG TAT ACA TAA AAT CT TCT GGT ACC TTT TTG AAA AAG TCA TGG AGG	4	97.5
8	tRNA-Lys.2	8283–8399	8313–8371	TCC TCT AGA GCC CAC TGT AAA GCT AAC TTA TGG TGG GCC ATA CGG TAG TAT TTA GTT G	4	99.6
9	ATP6	8963–9033	8991–9009	TCATCAGCCTACTCATTCAACCAATAGC TAGGTGGCCTGCAGTAATGTTAGC	3	94.3
10	tRNA-Gly	9965–10097	9996–10065	TGTCTCCATCTATTGATGAGGGTCTTACTC TAGGAGGGTGTTGATTATTAAAATTAAGGCGA	3	95.8
11	ND4	11 751–11 866	11 777–11 843	TCT CAA ACT ACG AAC GCA CTC ACA GT TAA GGC GAG GTT AGC GAG GCT TG	5	94.1
12	tRNA-His	12 096–12 230	12 121–12 202	TCC TAT CCC TCA ACC CCG ACA TCA T TTA GCA GTT CTT GTG AGC TTT CTC GG	7	97.3
13	tRNA-Leu-2.2	12 261–12 352	12 294–12 321	TCC CAA CTT TTA AAG GAT AAC AGC TAT CCA TTG TGG GGT GCA TGG TTA TTA CTT TTA TTT GGA G	5	84.4

Bolded compositions are rare variants further analyzed with confirmatory short-range PCR primers.

[a] Among 968 sequences (see *Materials and Methods*).