Int J Legal Med (2007) 121:57–67
DOI 10.1007/s00414-006-0117-7

TECHNICAL NOTE

Liquid chromatography–electrospray ionization mass spectrometry for simultaneous detection of mtDNA length and nucleotide polymorphisms

Herbert Oberacher · Harald Niederstätter · Walther Parson

Received: 15 February 2006 / Accepted: 5 July 2006 / Published online: 6 September 2006

Abstract We demonstrate the applicability of ion-pair reversed-phase high-performance liquid chromatography—electrospray ionization time-of-flight mass spectrometry (ICEMS) for the simultaneous characterization of length and nucleotide polymorphisms. Two sections within the first (HVS-I) and second (HVS-II) hypervariable segments of the mitochondrial (mt)DNA control region were selected as targets, both containing *poly*-cytosine (C) tracts, which display length heteroplasmy at a substantial frequency in the population. The two mtDNA sections were simultaneously amplified and analyzed by ICEMS in 90 maternally unrelated mother–offspring pairs from Austria. The findings were confirmed by direct sequencing of the polymerase chain reaction products. For the detailed characterization of present-length heteroplasmic variants, the results retrieved through ICEMS were more informative compared with those derived from direct sequencing. Hence, ICEMS represents an interesting option for successful application in forensic science.

Keywords Mitochondrial DNA · Length heteroplasmy · Mass spectrometry · Electrospray ionization · Screening method

Electronic supplementary material Supplementary material is available in the online version of this article at http://dx.doi.org/10.1007/s00414-006-0117-7 and is accessible for authorized users.

H. Oberacher · H. Niederstätter · W. Parson (✉)
Institute of Legal Medicine, Innsbruck Medical University,
Müllerstrasse 44,
6020 Innsbruck, Austria
e-mail: walther.parson@i-med.ac.at

Introduction

Mass spectrometry (MS) has emerged as a powerful tool for the analysis of biological molecules aided by the development of soft ionization techniques such as electrospray ionization (ESI) [1] and matrix-assisted laser desorption/ionization (MALDI) [2], which enable the transfer of polar and thermally labile molecules into the gas phase. In nucleic acids research, MS has gained attention in applications such as quality control of short oligonucleotides [3–5], characterization of short tandem repeat markers [6–11], and the genotyping of single-nucleotide (NT) polymorphisms [12–20].

In contrast to fragment-length estimation derived by electrophoretic mobility, the molecular mass represents an intrinsic feature, which can be used for a more direct discrimination of nucleic acids differing in length and/or NT composition. MALDI, being a well-established mass-based technology for nucleic acid analysis, is restricted to relatively small DNA molecules (approximately 30 NTs [21–23]). Consequently, MALDI can hardly be applied to the direct genotyping of polymerase chain reaction (PCR) products. As a consequence, a number of different strategies were applied to shorten the allele-specific product to a length compatible with the technical limitations of MALDI, including hybridization, ligation, cleavage, and primer extension [13, 14, 24, 25]. However, they all involve post-PCR manipulation, which lengthens the laboratory process and introduces increased risk of contamination. In contrast to MALDI, ESI-MS can be applied to the detection of sequence variants within nucleic acids with lengths up to several hundred NTs [19, 26]. Hence, the direct characterization of sequence variants within PCR amplicons is feasible.

Sample purification is of utmost importance for the successful mass spectrometric analysis of PCR products. PCR mixtures contain different amounts of DNA polymerase, 2′-deoxynucleoside triphosphates (dNTPs), oligonucleotide primers, detergents, albumin, and other ionic buffer components, including metal ions such as sodium, potassium, and magnesium, which can interfere with the detection of amplicons and therefore need to be removed [27, 28]. Commonly applied off-line and online sample preparation techniques include multiple ethanol precipitation [29–31], solid-phase extraction [32, 33], microdialysis [11, 27], and cation exchange [5, 34]. We propagate the use of ion-pair reversed-phase high-performance liquid chromatography online hyphenated to electrospray ionization time-of-flight mass spectrometry (ICEMS). Chromatography represents an excellent sample preparation method, which simultaneously purifies, denatures, and fractionates PCR-amplified nucleic acids before online detection by high-performance ESI-MS. Because no biochemical modifications (e.g., fluorescent labeling) and additional post-PCR reactions are necessary, ICEMS represents an accurate, fast, and—despite the costs for the instrumentation—inexpensive tool for the characterization of genomic markers [15, 17, 19].

In this work, we demonstrate the applicability of ICEMS for the simultaneous characterization of length and NT polymorphisms within human mitochondrial DNA (mtDNA). Two sections within the first and second hypervariable segments (HVS-I, 16144–16237; and HVS-II, 220–381) of the mitochondrial control region were selected as targets. Both contain homopolymeric cytosine tracts (poly C tracts, HVS-I, 16184–16193 [35]; HVS-II, 303–315 [36]), which represent hotspots for mutation and replication slippage causing the co-occurrence of two or more populations of mtDNA length and sequence variants in the same individual, tissue, or cell (heteroplasmy). The two sections were simultaneously amplified and analyzed by ICEMS from 90 maternally unrelated mother–offspring pairs, which served as an Austrian population sample. Sequencing results adapted from Ref. [37] were used to evaluate the plausibility of the ICEMS results.

Materials and methods

Chemicals and oligonucleotides

A stock solution (0.50 M) of butyldimethylammonium acetate (BDMAA) was prepared by titration of a 0.5-M solution of butyldimethylamine with acetic acid (Fluka) at 5°C until pH 8.3 was reached. For preparation of HPLC solvents, HPLC-grade water and acetonitrile (Merck, Darmstadt, Germany) were used. All oligonucleotides were obtained from Eurogentec (Seraing, Belgium).

DNA samples, extraction, and mtDNA quantification

Blood samples were obtained from 90 unrelated Austrian Caucasian females and their children (one child per woman). Total DNA was extracted as described previously [37], and mtDNA was quantified according to Köchl et al. [38]. For the designation of samples, an alphanumeric nomenclature was used, denoting mothers (m)/children (c), and plate number (1 or 2), followed by the sample position in the 96-well plate (e.g., H4). Mother/child pairs are characterized by sharing the same plate number and well coordinate (e.g., m1-A4/c1-A4).

PCR amplification

Duplex PCR amplification was performed on a Gene Amp PCR System 9700 (AB, Applied Biosystems, Foster City, CA, USA) in a total volume of 20 μl containing 1× Advantage 2 SA buffer (BD Biosciences Clontech, Uppsala, Sweden), 200 μM each dNTP, 500 nM each of the primers F16144 (TGACCACCTGTAGTACATAA), R16237 (TGTGTGATAGTTGAGGGTTG), and R381 (GCTGGTGTTAGGGTTCTTTG), 1 μM primer F220 (TGCTTGTARGACATAATAAT), 1× BD Advantage 2 Polymerase Mix (BD Biosciences Clontech), and 1 μl total genomic DNA. The thermal cycler protocol consisted of initial denaturation at 95°C for 1 min, 40 cycles of 95°C for 15 s, 56°C for 30 s, and 68°C for 1 min. The final extension step was carried out at 68°C for 20 min.

Ion-pair reversed-phase high-performance liquid chromatography—electrospray ionization time-of-flight mass spectrometry

An Ultimate fully integrated capillary HPLC system (LC-Packings, Amsterdam, The Netherlands) was used for all chromatographic experiments. A Famos microautosampler (LC-Packings) equipped with a 1.0-μl loop was used for sample injection. The 60×0.2-mm i.d. monolithic capillary column was prepared according to the published protocol [39], and the flow rate was set to 2.0 μl/min. PCR amplicons were injected onto the column without any prior sample preparation step and chromatographed as single strands under completely denaturing conditions (68°C). After each injection, the column was rinsed with 25 mM BDMAA solution containing 2.0% acetonitrile for 3 min. Subsequently, nucleic acids were eluted using a gradient of 2.0 to 28.0% acetonitrile in 25 mM BDMAA and detected by ESI-MS.

ESI-MS was performed on a QSTAR XL mass spectrometer (Applied Biosystems) equipped with a modified TurboIonSpray source. The modifications included the replacements of the Peek tubing transfer line and of the

stainless steel sprayer capillary by fused silica capillaries (transfer line, 375 μm, o.d., 20 μm i.d.; sprayer capillary, 90 μm, o.d., 20 μm, i.d.; Polymicro Technologies, Phoenix, AZ, USA). The exit of the capillary column was connected directly to the transfer line by means of a microtight union (Upchurch Scientific, Oak Harbor, WA, USA). During the first 5 min of each run, the sprayer was shifted from its optimal position to a more peripheral position to reduce the degree of contamination of the ion transfer optics with components of the PCR buffer. Mass calibration and optimization of instrumental parameters were performed in the negative ion mode as described [5, 19]. The spray voltage was typically in the range of 3.0–3.5 kV, and gas flows of 10–15 (ion source gas 1) and 40 arbitrary units (ion source gas 2) were employed. The temperature of ion source gas 2 was adjusted to 200°C. The accumulation time was set to 1 s, and ten time bins were summed up. Reconstructed ion chromatograms and mass spectra (scan, 800–1,500 amu) were recorded on a personal computer operating with the Analyst QS software (service pack 8, Applied Biosystems). Deconvolution of raw mass spectra was performed with Bayesian Protein Reconstruct (BioAnalyst 1.1.1, Applied Biosystems).

Results and discussion

Two sections within the first and second hypervariable segments (HVS-I, 16144–16237; HVS-II, 220–381) of the mtDNA control region were selected as targets to evaluate the applicability of ICEMS for the simultaneous characterization of length and NT polymorphisms. Both segments contain poly C tracts. In HVS-I, the poly C stretch is located between positions 16184 and 16193. In the majority of samples, the homopolymeric tract is interrupted by a thymine (T) residue at position 16189 [40], resulting in a C_5TC_4 sequence string. A T to C transition at position 16189 may result in an uninterrupted stretch of C's. Such a mononucleotide repeat usually induces instability, giving rise to length heteroplasmy, which represents a mosaic of poly C tracts varying in length from 8 to 14 NTs [35]. In HVS-II, the poly C stretch is located between positions 303 and 315. According to the revised Cambridge reference sequence (rCRS [40]), seven C's precede a T at position 310, which is subsequently followed by five additional C's (C_7TC_5). In this respect, the rCRS sequence represents a rare variant because the vast majority of sequences investigated so far carry six C's between positions 310 and 316, which is conventionally designated as insertion at position 315 (315.1C). Hence, C_7TC_6 represents the root of length variants. In this segment, heteroplasmic variants are typically observed in two different versions. Similar to the poly C stretch in HVS-I, the transition of the T residue at position 310 to C results in a long uninterrupted stretch of C's. Generally, this is a relatively rare event and only more frequent in haplogroups in which this transition is a phylogenetic character (e.g., B2, T2, and U4). A much more frequent mechanism for the generation of length heteroplasmy in HVS-II (up to 50% in west Eurasian populations) is the insertion of up to three C's between positions 303 and 310, which are then referred to as 309.1C, 309.2C, and 309.3C.

Identification of length and sequence variants

The two targets were simultaneously amplified in a duplex PCR. As shown in Fig. 1, the rCRS displays these amplicons with lengths of 94 (HVS-I) and 162 bp (HVS-II). For the analysis of the PCR products by ICEMS, no additional preparative post-PCR manipulation step was needed. The online chromatographic purification step efficiently removed interfering components of the amplification reaction mixture such as dNTPs, PCR primers, polymerase, and ionic buffer constituents from the amplicons. The use of an elevated column temperature (68°C) initiated the denaturation of the double-stranded PCR products during their chromatographic separation. Since A–T and G–C bp have very similar molecular masses of 615.4 and 616.4 amu, and A<>T or G<>C substitutions do not cause any mass shift at all, base substitutions are difficult to identify in double-helical DNA via molecular mass measurements. Within single-stranded nucleic acids, however, single-base substitutions cause

16144-16237

tGaccacctgtagtacatAAaaacccAaTCcacatCaaAAccCccTccCcatgcttaCaagcaagtacagCaatcaACc
CTcaactaTCACaca

220-381

tgcttgtaGgacataataaTaacaattGaaTgtctgcacagccActttccacacagacatcataaCaaaaaatttCcaccaaacccc
ccctcccccgcttCtggccacAgcacttaaacaCatctctgcca aaccccaaaaacaaagaaccctaacaccagc

Fig. 1 Sequences of the PCR products covering the homopolymeric cytosine tracts within HVS-I and HVS-II. Sequences of primers are *underlined*. *Capital letters* indicate positions where base exchanges were observed within our sample set

measurable molecular mass differences ranging between ±9.01 (A<>T) and ±40.02 amu (G<>C). Hence, denaturation represents a basic requirement to tap the full potential of the mass spectrometric discrimination capacity.

Analysis of an exemplary sample is depicted in Fig. 2. In ion–pair reversed-phase chromatography, nucleic acids are primarily separated according to chain length. Hence, single strands were fractionated into two peaks representing the HVS-I and HVS-II amplicons (Fig. 2a). Raw mass spectra, which were extracted from the reconstructed ion chromatogram (Fig. 2b,c), consisted of a series of peaks, each of which represented a multiply charged ion of an intact nucleic acid having a specific number of protons removed from the phosphodiester groups. All signals belonging to one series were deconvoluted to yield the molecular mass of the corresponding nucleic acid (Fig. 2b,c and Tables 1 and 2). For the identification of length and sequence variants, the measured molecular mass was compared with

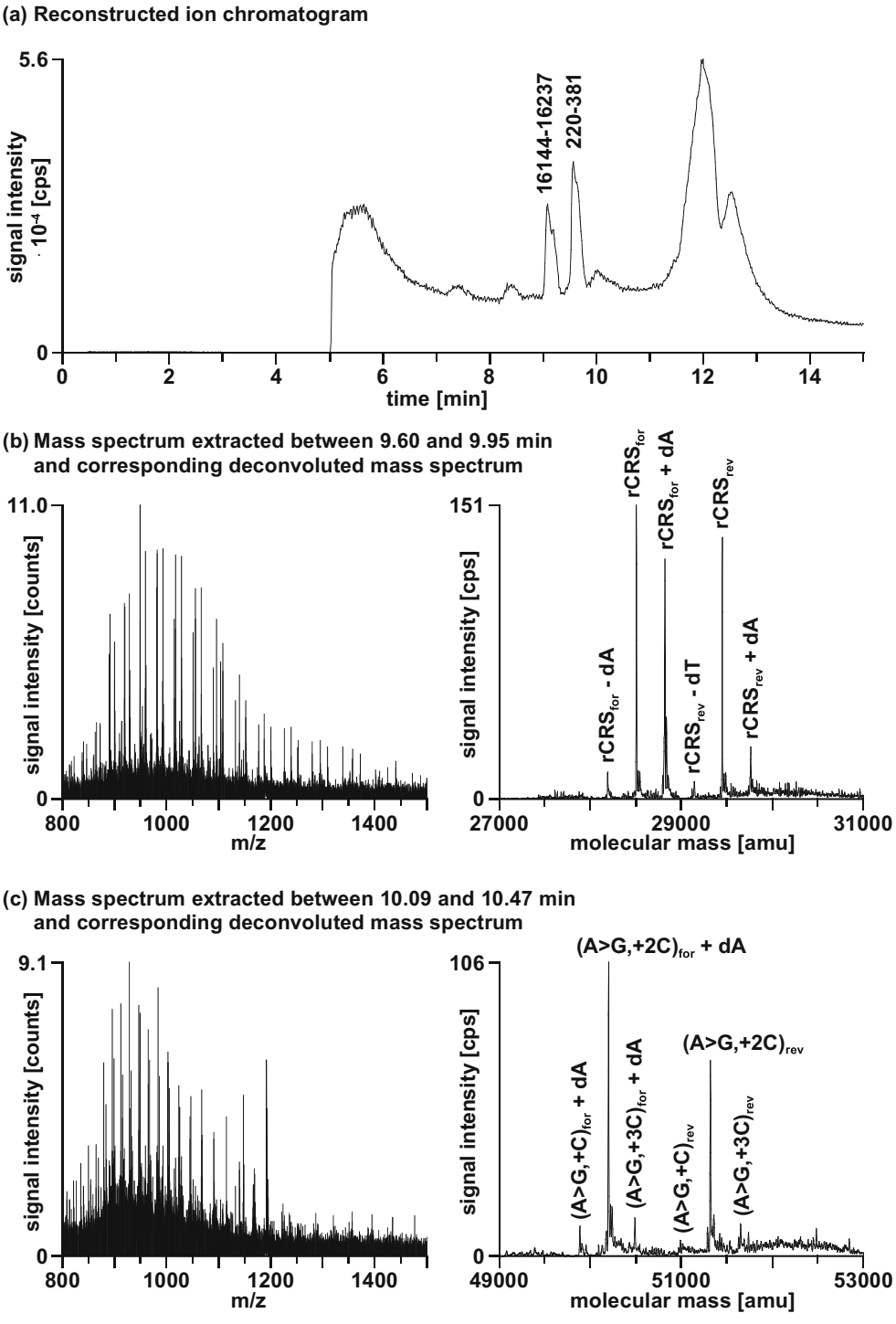


Fig. 2 **a** Reconstructed ion chromatogram and **b**, **c** mass spectra obtained from the analysis of a sample showing length heteroplasmy in HVS-II

Table 1 Simultaneous identification of length and nucleotide polymorphisms within amplicons covering the homopolymeric stretches of cytosines in HVS-I via the comparison of calculated and measured molecular masses

Measured mass (amu)	Assignment	Theoretical mass (amu)	Deviation (amu)	Differences to rCRS
28,507.23	$rCRS_{for}$	28,506.60	0.63	None
29,447.75	$rCRS_{rev}$	29,447.04	0.71	
28,819.98	$rCRS_{for}$ + dA	28,819.80	0.18	
29,760.81	$rCRS_{rev}$ + dA	29,760.25	0.56	
28,192.70	$rCRS_{for}$ − dA	28,193.39	−0.69	
29,142.45	$rCRS_{rev}$ − dT	29,142.85	−0.40	

the theoretical molecular mass derived from the rCRS. In the case of a match, the deviation between measured and theoretical molecular masses was typically within ±25 ppm due to the high accuracy of the time-of-flight mass analyzer.

For one example, the process of allele identification is briefly outlined here, and six different single strands belonging to the HVS-I target were identified (Fig. 2b, Table 1). The two most abundant species were assigned to the forward and reverse single strands lacking any difference to the rCRS. Expectedly, the analysis of a certain HVS-I variant amplified with a non-proofreading polymerase can result in the identification of up to four variant-specific single strands, forward and reverse strand with and without the nontemplate addition of a 2′-deoxyadenosine. In our case, we observed two additional species obviously originating from two strands, which were one NT shorter than the variant-specific forward and reverse strands (Fig. 2b, Table 1). Based on the measured molecular mass differences, we identified a truncated reverse primer population being responsible for the two additional signals. This parallels a one-NT smaller shoulder peak, which is sometimes observed in conventional capillary electrophoresis, with the difference that ICEMS unambiguously identified the source of the phenomenon. This example demonstrates another advantage of mass-based analysis of nucleic acids, the quality assessment of primer sequences, which have a crucial effect on the reliability of DNA results.

Similarly, six different single strands were identified in the deconvoluted mass spectrum obtained from the HVS-II target (Fig. 2c). In contrast to the HVS-I target, the measured molecular masses of the two most abundant species did not match with the molecular masses calculated from the rCRS (Table 2, step 1). The nature of these length and sequence alterations was determined by adjusting the NT composition of the reference sequence (Table 2, steps 2 and 3). In a first step, the length of the reference sequence was adapted. The obtained molecular mass deviations of approximately 644 and 908 amu suggested the addition of three NTs to the forward strand and two NTs to the reverse

Table 2 Simultaneous identification of length and nucleotide polymorphisms within amplicons covering the homopolymeric stretches of cytosines in HVS-II via the comparison of calculated and measured molecular masses

	Measured mass (Da)	Assignment	Theoretical mass (Da)	Deviation (amu)	Differences to rCRS
Step 1	50,202.25	$rCRS_{for}$	49,294.07	908.18	Length and nucleotide polymorphisms
	51,317.97	$rCRS_{rev}$	50,673.73	644.24	
Step 2	50,202.25	$(rCRS, +2C)_{for}$ + dA	50,185.65	16.60	Nucleotide polymorphism
	51,317.97	$(rCRS, +2C)_{rev}$	51,332.15	−14.18	
Step 3	50,202.25	$(A>G, +2C)_{for}$ + dA	50,201.65	0.60	A>G, +2C
	51,317.97	$(A>G, +2C)_{rev}$	51,317.14	0.83	
	50,202.25	$(C>T, +2C)_{for}$ + dA	50,200.66	1.59	
	51,317.97	$(C>T, +2C)_{rev}$	51,316.15	1.82	
Step 4	49,911.45	$(A>G, +2C)_{for}$ + dA	50,201.65	−290.20	Length variation
	50,988.45	$(A>G, +2C)_{rev}$	51,317.14	−328.69	
	50,202.25	$(A>G, +2C)_{for}$ + dA	50,201.65	0.60	A>G, +2C
	51,317.97	$(A>G, +2C)_{rev}$	51,317.14	0.83	
	50,491.45	$(A>G, +2C)_{for}$ + dA	50,201.65	289.80	Length variation
	51,648.45	$(A>G, +2C)_{rev}$	51,317.14	331.31	
Step 5	49,911.45	$(A>G, +C)_{for}$ + dA	49,912.47	−1.02	A>G, +C
	51,317.97	$(A>G, +C)_{rev}$	51,317.14	0.83	
	50,988.45	$(A>G, +2C)_{for}$ + dA	50,987.93	0.52	A>G, +2C
	50,491.45	$(A>G, +2C)_{rev}$	50,490.84	0.61	
	50,202.25	$(A>G, +3C)_{for}$ + dA	50,201.65	0.60	A>G, +3C
	51,648.45	$(A>G, +3C)_{rev}$	51,646.35	2.10	

strand. We concluded that in this case, the insertion of two C's was variant-specific. The additional increase in the forward strand length originated from the nontemplate addition of 2′-deoxyadenosine. In a subsequent step, NT polymorphisms were identified by comparing measured and "length-corrected" theoretical molecular masses. Molecular mass differences of 16.60 amu for the forward single strand and 14.18 amu for the reverse single strand suggested the presence of an A>G substitution (Table 2, step 2; theoretical molecular mass differences for A>G, 15.999 and −15.011). The implementation of the sequence variation improved the correlation between measured and calculated molecular masses significantly. Deviations within ±0.83 amu or ±16.1 ppm were obtained (Table 2, step 3). Because of its nearly identical molecular mass shifts, a C>T substitution represented another potential sequence variation (theoretical molecular mass differences for C>T, 15.011 and −15.999). However, taking advantage of the high mass accuracy of the applied time-of-flight mass spectrometer, the C>T substitution was unequivocally excluded in this case (Table 2, step 3).

The identification of the most abundant sequence variant facilitated the assignment of the four other species found in the deconvoluted mass spectrum (Fig. 2c). The obtained molecular mass differences suggested the presence of two additional length variants (Table 2, step 4). Conclusively, three different length variants were identified, which were 1, 2, or 3 C's longer than the rCRS variant (Table 2, step 5).

Comparison of ICEMS and sequencing results

The two mtDNA segments were genotyped in 90 maternally unrelated mother–offspring pairs. A summary of the results can be found as "Electronic supplementary material (Table S1)" on the home page of the journal. Generally, results obtained by ICEMS correlated well with the results obtained from direct sequencing of PCR products [37]. Sequencing electropherograms showed a characteristic "out-of-phase" pattern downstream of the length heteroplasmic C stretches. This is a well-known phenomenon, which can make data interpretation difficult in some cases. Here, ICEMS offered the considerable advantage that length variants were clearly separated by MS. Deconvoluted mass spectra resulted in easily distinguishable peak patterns directly related to observed heteroplasmic fragment lengths. Hence, treatment of mass spectrometric data was a rather simple task, which enabled the unequivocal identification of the length variants.

Regarding NT variations, sequencing results can be more informative. ICEMS can only indicate the presence of a sequence variation relative to a given reference sequence but does not specify its exact location because the molecular mass of an intact nucleic acid solely reflects the NT composition but not the base order. Consequently, sequences with multiple substitutions cannot be distinguished from the reference sequence if the variations compensate each other in their summative molecular mass shift (e.g., the parallel transition of A>G and G>A on the same fragment), as observed in 13 of the 90 analyzed mother–offspring pairs.

Among the 180 samples, both targets failed to amplify only in 1 case (m1-D5). In five more samples (HVS-I: c1-C2; HVS-II: m1-F4, c1-F4, m2-D1, and c2-D1) a detectable amplicon amount was obtained only for one of the two targets. In some cases, heteroplasmic (HVS-I: m1-C2, c1-C2, m2-D2, and c2-D2; HVS-II: m1-G5) or homoplasmic base exchanges (HVS-I: m1-E1, c1-E1, m1-G2, c1-G2, m1-H3, c1-H3, m2-D6, and c2-D6; HVS-II: c1-G5, m2-B4, and c2-B4) were observed where sequencing did not support their presence [37]. In each of these samples, mismatches in the primer binding sequences located close to the 3′-end of the forward primers (16162G, 16163G, and 239C) were identified as the sources for the observed PCR aberrations or miscallings. We believe that the small amount of proofreading polymerase present within the Advantage 2 polymerase mix (BD Biosciences Clontech) at least partially corrected these mismatches, leading to detectable amounts of "mutated" amplicons, which mimicked the presence of homoplasmic or heteroplasmic variations within the sequences framed by the amplification primers.

Quantitative analysis of length heteroplasmy

Besides the qualitative molecular mass information, deconvoluted mass spectra contained quantitative information in the form of peak intensities which was used for the determination of relative contents of heteroplasmic-length variants. Data treatment was performed as demonstrated in Ref. [28], and the peak intensities were extracted from the deconvoluted mass spectra (Fig. 2c, Table 3). Within each class of single strand, relative signal intensities were given as a percentage of the highest signal intensity (Table 3).

Table 3 Determination of the relative contents of length variants identified within a sample showing length heteroplasmy in HVS-II

Allele	Single strand	Intensity (cps)	Relative intensity	Relative content
A>G, +C	$(A>G, +2C)_{for}$ + dA	3.9	0.037	0.059
	$(A>G, +2C)_{rev}$	6.0	0.081	
A>G, +2C	$(A>G, +2C)_{for}$ + dA	105.6	1.000	1.000
	$(A>G, +2C)_{rev}$	73.6	1.000	
A>G, +3C	$(A>G, +2C)_{for}$ + dA	14.2	0.134	0.148
	$(A>G, +2C)_{rev}$	11.9	0.162	

Finally, for each of the identified length variants, the strand-specific relative signal intensities were averaged to obtain the variant-specific relative content within the mtDNA mixture (Table 3). In this way, relative contents of 0.059, 1.000, and 0.148, respectively, were determined for the three different length variants in HVS-II of the sample described (Table 3).

Regarding the HVS-I target, length heteroplasmy was detected in 12 (13.3%) of 90 mother–offspring pairs, which corresponded well with the sequencing results [37] and with values reported for other European populations [41–44]. In all cases, length heteroplasmy was induced by a transition at position 16189. Typically, expansion of the C stretch up to a length of 14 NTs was more likely than contraction. We observed a small fraction of mitochondrial genomes with an eight-NT-long C tract in only one sample. Exemplary results obtained from the typing of the HVS-I targets of a mother–offspring pair are shown in Fig. 3.

Length heteroplasmy in HVS-II was observed in 46 (52.9%) individuals of the mother population and in 45 (51.1%) individuals of the offspring population. Typically, much lower values were reported in the literature for different populations in HVS-II. In Germany [41], the mtDNA of 12% of the population sample represented a mixture of different length variants; in Italy [42], 17%; in Slovenia [43], 8%; and in the Czech Republic [44], 23%. The ICEMS approach offered a clearer picture of the length variants included in a heteroplasmic sample, and the results suggest that length heteroplasmy in HVS-II is very likely to be present in many more individuals than so far expected and described. This theory is supported by data presented by Lutz–Bonengel et al. [45], who determined an HVS-II heteroplasmy rate of 60% via the electrophoretic analysis of fluorescently labeled and cleaved PCR products in a German population sample. Exemplary results obtained from the analysis of HVS-II targets are depicted in Fig. 4.

Inheritance of length heteroplasmic mtDNA variants

The genotyping results obtained from the 90 maternally unrelated mother–offspring pairs gave some insights into the inheritance of heteroplasmic length variants. Regarding the HVS-I target, length heteroplasmy was inherited in all cases. A closer look at the 12 intergenerational events revealed that in all cases, maternal and filial distributions of length variants had nearly similar shapes, indicating the presence of a rather wide bottleneck. In all but one case, the offspring displayed overall lower contents of heteroplasmic minority components compared with their mothers (Fig. 5a). The rCRS variant was revealed to be the most abundant length variant observed in our sample population (Fig. 5b). Four mothers and three children carried a one-NT larger variant, which may have become stable by the A-C substitution at position 16183. Only in one case within the mother population was the most abundant variant two NTs larger than the rCRS. For two intergenerational events, the length of the most abundant variant differed between mother and child, and in both cases, we observed a decrease in length of the C stretch.

In HVS-II, length heteroplasmy was observed in 49 mother–offspring pairs. In 41 of these, length heteroplasmy was inherited from the mother to the corresponding child. Within five mother–offspring pairs, heteroplasmy was solely detected in the mitochondrial genome of the mother, whereas within three mother–offspring pairs, heteroplasmy was exclusively detected in the mitochondrial genome of the offspring. Similar to HVS-I, a decrease in the averaged relative contents of the length heteroplasmic minority components from mothers to their offspring was observed (Fig. 5a). In all but two cases, the offspring carried overall lower levels of heteroplasmy than their mothers. The most prominent length variant was one NT longer than the rCRS as a result of the omnipresent 315.1C (Fig. 5b). We did not find contractions of the C stretch below C_7, whereas the

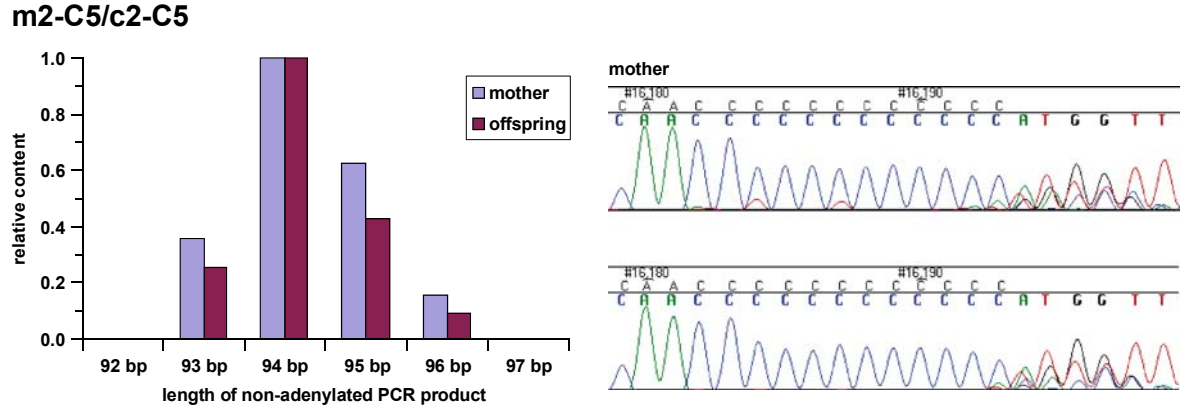
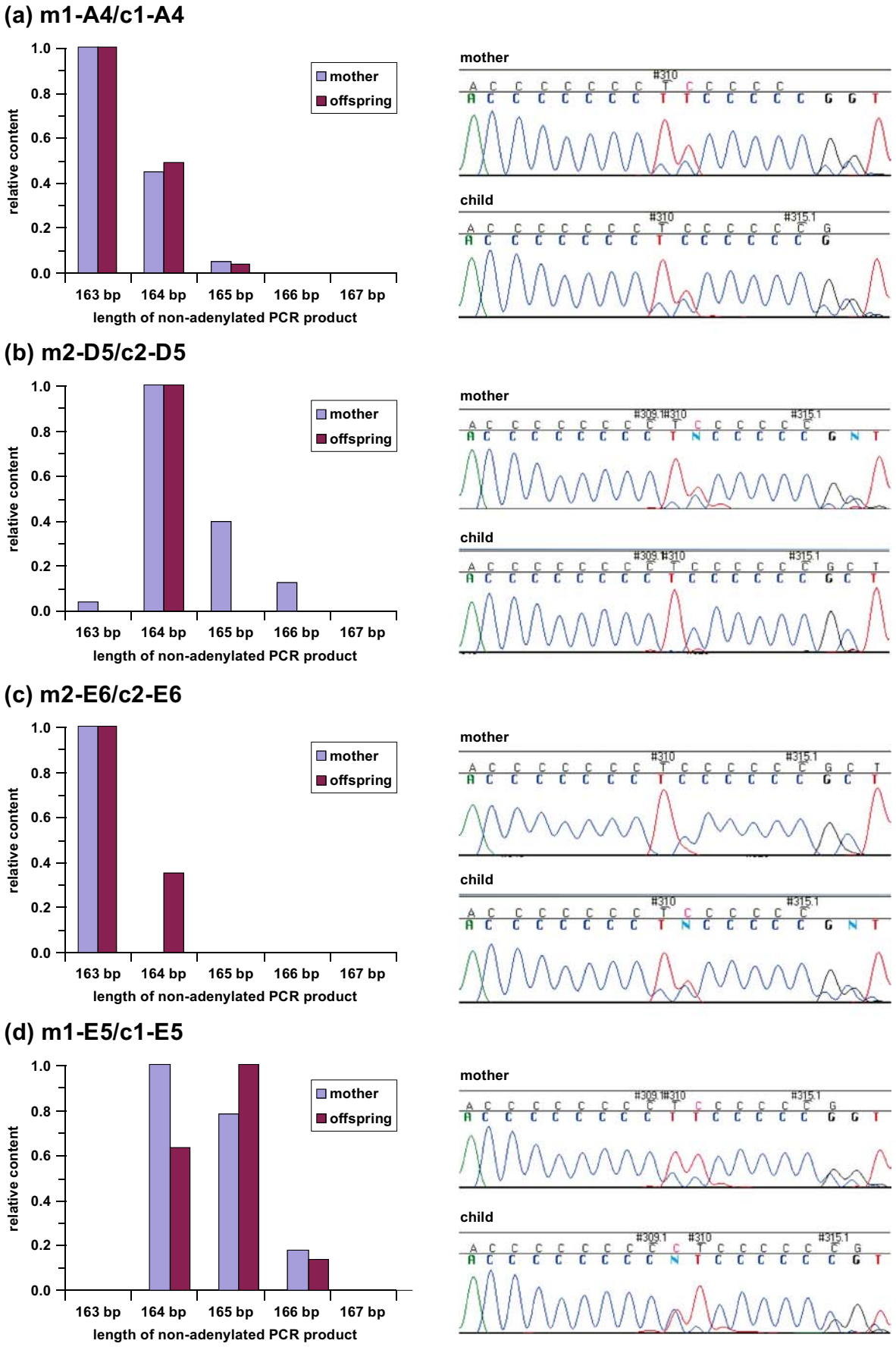


Fig. 3 Genotyping results of a selected mother/child pair showing length heteroplasmy in HVS-I. The amplicon length of the rCRS would have been 94 bp


(a) m1-A4/c1-A4
mother
offspring
relative content
1.0
0.8
0.6
0.4
0.2
0.0
163 bp
164 bp
165 bp
166 bp
167 bp
length of non-adenylated PCR product
mother
child
#310
#315.1
(b) m2-D5/c2-D5
mother
offspring
relative content
length of non-adenylated PCR product
mother
child
#309.1
#310
#315.1
(c) m2-E6/c2-E6
mother
offspring
relative content
length of non-adenylated PCR product
mother
child
#310
#315.1
(d) m1-E5/c1-E5
mother
offspring
relative content
length of non-adenylated PCR product
mother
child
#309.1
#310
#315.1

◀ **Fig. 4** Genotyping results of selected mother/child pairs showing length heteroplasmy in HVS-II. **a** Mother and offspring principally have the same distributions of length variants. **b** Mother is showing length heteroplasmy, whereas the offspring does not. **c** Mother is not showing length heteroplasmy, whereas the offspring does. **d** Mother and offspring have different distributions of length variants. The amplicon length of the rCRS would have been 162 bp

expansions of the C tract up to 11 NTs were observed. Within six mother–offspring pairs, the length of the most abundant variant changed between the generations, resulting in a one-NT larger dominant variant.

Conclusions

The potential of ICEMS as a fast and efficient method for the characterization of mtDNA variations was demonstrated. Sequence information was gained via the comparison of measured molecular masses to theoretical molecular masses derived from the rCRS. Because the molecular mass of an intact nucleic acid solely reflects the NT composition but not the order, ICEMS was only able to indicate the presence of a sequence variation but failed to specify the exact location of the ambiguity. For this purpose, DNA sequencing remains the method of choice. Regarding the characterization of length variants, ICEMS represented a more appropriate technology and offered the considerable advantage that length variants were effortlessly separated by mass spectrometry. Hence, deconvoluted mass spectra were better interpretable than "blurred" electropherograms. Based on these findings, we do not consider ICEMS to be a substitute for sequence analysis of mtDNA in general. However, from the standpoint of a screening method, it represents a very fast, cheap, and reliable technique. The simultaneous detection of sequence and fragment length variation is a feature which is still missing in the toolbox of DNA analysis methods.

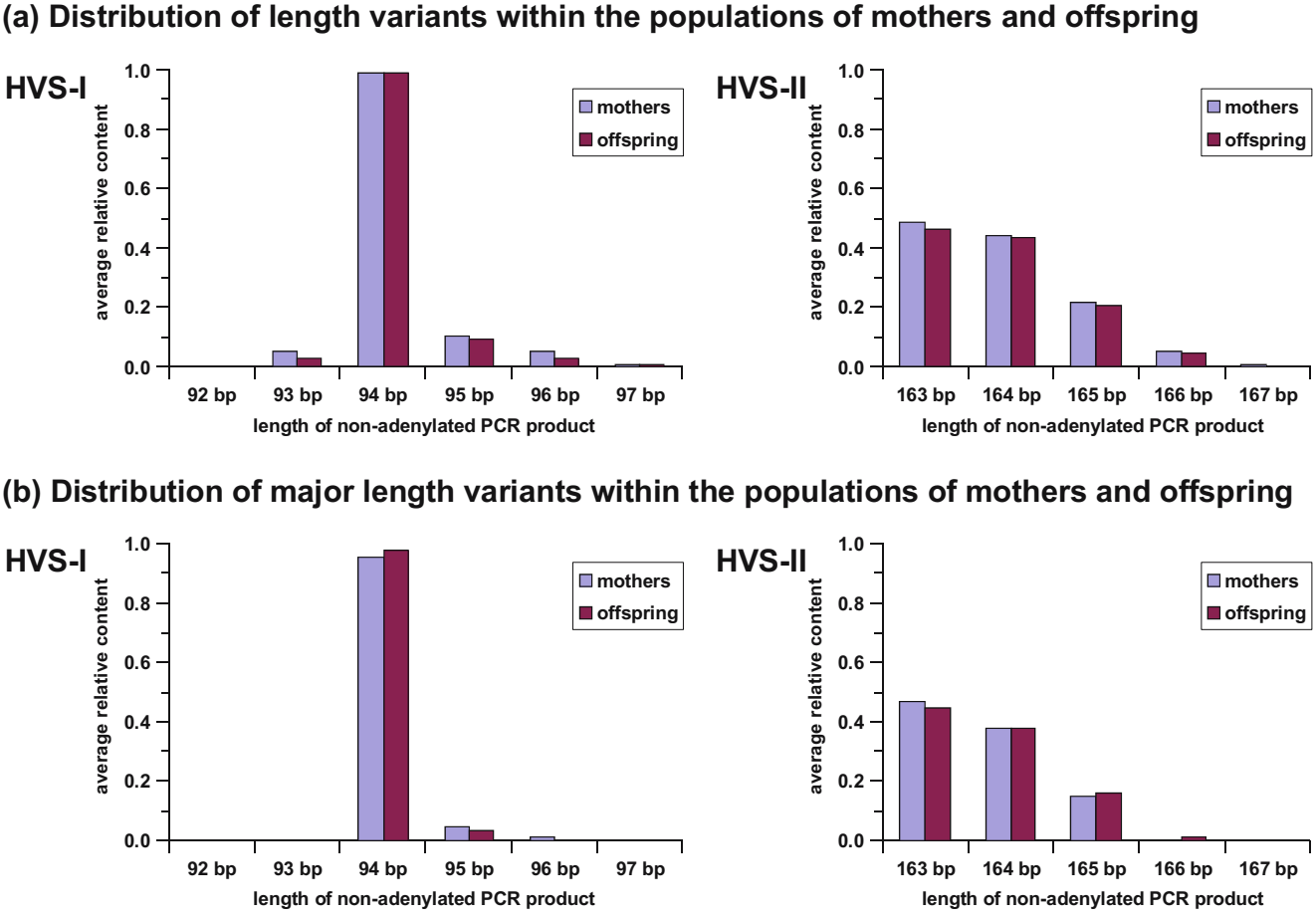


Fig. 5 Distribution of (major) length variants within (sub-) populations of mothers and offspring. The amplicon length of the rCRS would have been 94 bp for the HVS-I target and 162 bp for the HVS-II target

Acknowledgement The authors wish to thank Applied Biosystems for the generous provision of the mass spectrometer and the associated equipment.

References

1. Whitehouse CM, Dreyer RN, Yamashita M, Fenn JB (1985) Electrospray interface for liquid chromatographs and mass spectrometers. Anal Chem 57:675–679
2. Karas M, Hillenkamp F (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal Chem 60:2299–2301
3. Butler JM, Devaney JM, Marino MA, Vallone PM (2001) Quality control of PCR primers used in multiplex STR amplification reactions. Forensic Sci Int 119:87–96
4. Fountain KJ, Gilar M, Gebler JC (2004) Electrospray ionization mass spectrometric analysis of nucleic acids using high-throughput on-line desalting. Rapid Commun Mass Spectrom 18:1295–1302
5. Oberacher H, Niederstätter H, Parson W (2005) Characterization of synthetic nucleic acids by electrospray ionization quadrupole time-of-flight mass spectrometry. J Mass Spectrom 40:932–945
6. Butler JM, Li J, Shaler TA, Monforte JA, Becker CH (1998) Reliable genotyping of short tandem repeat loci without an allelic ladder using time-of-flight mass spectrometry. Int J Legal Med 112:45–49
7. Butler JM, Levin BC (1998) Forensic applications of mitochondrial DNA. Trends Biotechnol 16:158–162
8. Oberacher H, Parson W, Mühlmann R, Huber CG (2001) Analysis of polymerase chain reaction products by on-line liquid chromatography–mass spectrometry for genotyping of polymorphic short tandem repeat loci. Anal Chem 73:5109–5115
9. Null AP, Muddiman DC (2001) Perspectives on the use of electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry for short tandem repeat genotyping in the post-genome era. J Mass Spectrom 36:589–606
10. Null AP, Hannis JC, Muddiman DC (2001) Genotyping of simple and compound short tandem repeat loci using electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. Anal Chem 73:4514–4521
11. Hannis JC, Muddiman DC (2001) Genotyping short tandem repeats using flow injection and electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. Rapid Commun Mass Spectrom 15:348–350
12. Oberacher H, Oefner PJ, Parson W, Huber CG (2001) On-line liquid chromatography mass spectrometry: a useful tool for the detection of DNA sequence variation. Angew Chem Int Ed Engl 40:3828–3830
13. Pusch W, Wurmbach JH, Thiele H, Kostrzewa M (2002) MALDI-TOF mass spectrometry-based SNP genotyping. Pharmacogenomics 3:537–548
14. Tost J, Gut IG (2002) Genotyping single nucleotide polymorphisms by mass spectrometry. Mass Spectrom Rev 21:388–418
15. Berger B, Hölzl G, Oberacher H, Niederstätter H, Huber CG, Parson W (2002) Single nucleotide polymorphism genotyping by on-line liquid chromatography–mass spectrometry in forensic science with the Y-chromosomal locus M9. J Chromatogr B Analyt Technol Biomed Life Sci 782:89–97
16. Oberacher H, Oefner PJ, Hölzl G, Premstaller A, Huber CG (2002) Re-sequencing of multiple single nucleotide polymorphisms by liquid chromatography–electrospray ionization mass spectrometry. Nucleic Acids Res 30:e67
17. Oberacher H, Huber CG, Oefner PJ (2003) Mutation scanning by ion-pair reversed-phase high-performance liquid chromatography–electrospray ionization mass spectrometry (ICEMS). Hum Mutat 21:86–95
18. Van Ert MN, Hofstadler SA, Jiang Y et al (2004) Mass spectrometry provides accurate characterization of two genetic marker types in *Bacillus anthracis*. Biotechniques 37:642–648
19. Oberacher H, Niederstätter H, Casetta B, Parson W (2005) Detection of DNA sequence variations in homo- and heterozygous samples via molecular mass measurements by electrospray ionization time-of-flight mass spectrometry. Anal Chem 77:4999–5008
20. Hall TA, Budowle B, Jiang Y et al (2005) Base composition analysis of human mitochondrial DNA using electrospray ionization mass spectrometry: a novel tool for the identification and differentiation of humans. Anal Biochem 344:53–69
21. Fei Z, Ono T, Smith LM (1998) MALDI-TOF mass spectrometric typing of single nucleotide polymorphisms with mass-tagged ddNTPs. Nucleic Acids Res 26:2827–2828
22. Shahgholi M, Garcia BA, Chiu NH, Heaney PJ, Tang K (2001) Sugar additives for MALDI matrices improve signal allowing the smallest nucleotide change (A:T) in a DNA sequence to be resolved. Nucleic Acids Res 29:e91
23. Mengel-Jorgensen J, Sanchez JJ, Borsting C, Kirpekar F, Morling N (2004) MALDI-TOF mass spectrometric detection of multiplex single base extended primers. A study of 17 Y-chromosome single-nucleotide polymorphisms. Anal Chem 76:6039–6045
24. Graber JH, Smith CL, Cantor CR (1999) Differential sequencing with mass spectrometry. Genet Anal 14:215–219
25. Jurinke C, Oeth P, van den Boom D (2004) MALDI-TOF mass spectrometry: a versatile tool for high-performance DNA analysis. Mol Biotechnol 26:147–164
26. Muddiman DC, Anderson GA, Hofstadler SA, Smith RD (1997) Length and base composition of PCR-amplified nucleic acids using mass measurements from electrospray ionization mass spectrometry. Anal Chem 69:1543–1549
27. Null AP, George LT, Muddiman DC (2002) Evaluation of sample preparation techniques for mass measurements of PCR products using ESI–FT–ICR mass spectrometry. J Am Soc Mass Spectrom 13:338–344
28. Oberacher H, Parson W, Hölzl G, Oefner PJ, Huber CG (2004) Optimized suppression of adducts in polymerase chain reaction products for semi-quantitative SNP genotyping by liquid chromatography–mass spectrometry. J Am Soc Mass Spectrom 15:1897–1906
29. Stults JT, Marsters JC (1991) Improved electrospray ionization of synthetic oligodeoxynucleotides. Rapid Commun Mass Spectrom 5:359–363
30. Portier N, Van Dorsselaer A, Cordier Y, Roch O, Bischoff R (1994) Negative electrospray ionization mass spectrometry of synthetic and chemically modified oligonucleotides. Nucleic Acids Res 22:3895–3903
31. Walters JJ, Muhammad W, Fox KF et al (2001) Genotyping single nucleotide polymorphisms using intact polymerase chain reaction products by electrospray quadrupole mass spectrometry. Rapid Commun Mass Spectrom 15:1752–1759
32. Gilar M, Belenky A, Wang BH (2001) High-throughput biopolymer desalting by solid-phase extraction prior to mass spectrometric analysis. J Chromatogr A 921:3–13
33. Jiang Y, Hofstadler SA (2003) A highly efficient and automated method of purifying and desalting PCR products for analysis by electrospray ionization mass spectrometry. Anal Biochem 316:50–57
34. Huber CG, Buchmeiser MR (1998) On-line cation-exchange for suppression of adduct formation in negative-ion electrospray mass spectrometry of nucleic acids. Anal Chem 70:5288–5295
35. Bendall KE, Sykes BC (1995) Length heteroplasmy in the first hypervariable segment of the human mtDNA control region. Am J Hum Genet 57:248–256

36. Hauswirth WW, Clayton DA (1985) Length heterogeneity of a conserved displacement-loop sequence in human mitochondrial DNA. Nucleic Acids Res 13:8093–8104
37. Brandstätter A, Niederstätter H, Parson W (2004) Monitoring the inheritance of heteroplasmy by computer-assisted detection of mixed basecalls in the entire human mitochondrial DNA control region. Int J Legal Med 118:47–54
38. Köchl S, Niederstätter H, Parson W (2005) DNA extraction and quantitation of forensic samples using the phenol–chloroform method and real-time PCR. In: Carracedo A (ed) Forensic DNA typing protocols (methods in molecular biology). Humana, Totowa, pp 13–29
39. Premstaller A, Oberacher H, Huber CG (2000) High-performance liquid chromatography–electrospray ionization mass spectrometry of single- and double stranded nucleic acids using monolithic capillary columns. Anal Chem 72:4386–4393
40. Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat Genet 23:147
41. Pfeiffer H, Brinkmann B, Hühne J et al (1999) Expanding the forensic German mitochondrial DNA control region database: genetic diversity as a function of sample size and microgeography. Int J Legal Med 112:291–298
42. Tagliabracci A, Turchi C, Buscemi L, Sassaroli C (2001) Polymorphism of the mitochondrial DNA control region in Italians. Int J Legal Med 114:224–228
43. Zupanic Pajnic I, Balazic J, Komel R (2004) Sequence polymorphism of the mitochondrial DNA control region in the Slovenian population. Int J Legal Med 118:1–4
44. Vanecek T, Vorel F, Sip M (2004) Mitochondrial DNA D-loop hypervariable regions: Czech population data. Int J Legal Med 118:14–18
45. Lutz-Bonengel S, Sanger T, Pollak S, Szibor R (2004) Different methods to determine length heteroplasmy within the mitochondrial control region. Int J Legal Med 118:274–281