Int J Leg Med (2006) 120: 18–23
DOI 10.1007/s00414-005-0024-3

ORIGINAL ARTICLE

Harald Niederstätter · Michael D. Coble · Petra Grubwieser · Thomas J. Parsons · Walther Parson

Characterization of mtDNA SNP typing and mixture ratio assessment with simultaneous real-time PCR quantification of both allelic states

Received: 10 March 2005 / Accepted: 4 July 2005 / Published online: 9 August 2005

Abstract We performed a study on the forensic utility of allele-discriminatory quantitative real-time PCR (rtPCR) using Minor Groove Binder TaqMan probes, targeting the highly variable mitochondrial single nucleotide polymorphism 16519T/C. The apparent single-cycle PCR efficiency was virtually 100% for both 16519 alleles. The allele designations made by rtPCR were concordant with the results obtained in a previous study by sequencing analysis. In heteroplasmic samples, minor allele proportions down to 9% were unambiguously detected and quantified. The variation in allele proportion estimates was essentially the same within and between different rtPCR runs, and the differences between total copy number estimates found for rerun samples were comparable to those found with non-allele-discriminatory quantitative rtPCR assays.

Keywords Heteroplasmy · Allele proportion assessment · TaqMan assay · Forensic science

H. Niederstätter · P. Grubwieser · W. Parson (✉)
Institute of Legal Medicine, Innsbruck Medical University,
Müllerstrasse 44,
6020 Innsbruck, Austria
e-mail: walther.parson@uibk.ac.at
Tel.: +43-512-5073303
Fax: +43-512-5072764

M. D. Coble · T. J. Parsons
Armed Forces DNA Identification Laboratory,
Armed Forces Institute of Pathology,
Building 101, 1413 Research Blvd.,
Rockville, MD, 20850, USA

M. D. Coble
Biotechnology Division,
National Institute of Standards and Technology,
100 Bureau Drive, Mail Stop 8311,
Gaithersburg, MD, 20899, USA

Introduction

The analysis of mitochondrial DNA (mtDNA) has proven to be very useful in forensic casework, especially when nuclear short tandem repeat markers cannot be typed. However, the principal limitation associated with forensic mtDNA typing is the low power of discrimination that is obtained when common mitochondrial types are present [3]. Current mtDNA testing typically targets variable base positions in one or two hypervariable parts (HV1, HV2) of the noncoding control region by sequencing both strands of PCR-amplified segments. It is now increasingly recognized that assays targeting single nucleotide polymorphisms (SNPs) are well suited for gaining additional information in mtDNA testing (e.g., [7]; for review, see [8]).

Here we investigate the forensic applicability of quantitative real-time PCR (rtPCR) using TaqMan hybridization probes targeted to the highly discriminatory mitochondrial control region SNP 16519T/C [3, 6]. Our interest in the TaqMan assay for forensic SNP detection relates to a number of potential advantages such as good availability of chemistry and instrumentation, broad linear dynamic range, and homogeneous assay format. The multicolor capability of rtPCR instruments enables the simultaneous interrogation of both base states of the SNP under investigation, which is particularly important because of the potential of mtDNA to manifest heteroplasmic mixtures in continuously varying proportions.

Materials and methods

DNA samples

A total of 405 genomic DNA (gDNA) samples from healthy west Eurasian donors from Austria were classified as either homoplasmic (n=403) or heteroplasmic (n=2) at

the mitochondrial nucleotide position (ntp) 16519 on the basis of sequence information taken from a previous study ([2]; and unpublished data). The sample set included the heteroplasmic reference samples B and C (both T>C), and the homoplasmic samples A (16519T) and D (16519C).

Design of primers and hybridization probes

Primers and probes for the allele-discriminatory TaqMan approach were designed with the ABI PRISM Primer Express software package (version 1.5, AB, Applied Biosystems, Foster City, CA) using GenBank entry J01415 as template.

Dual probe singleplex rtPCR

From the 405 gDNA extracts as well as from the cloned alleles p16519T and p16519C, 65-bp-long fragments spanning ntps 16486–16550 of the mitochondrial genome were amplified in 20 μl cocktails comprising 1× TaqMan Universal PCR Master Mix (AB), 300 nM each primer (16486 F: AACTGTATCCGACATCTGGTTCCTA; 16550 R: AGG GGAACGTGTGGGCTATT), 100 nM each fluorescently labeled hybridization probe (MGB16519T: VIC-AGGCTT TATGACCCTGAA-MGBDQ; MGB16519C: FAM-AGG CTTTATGGCCCTGA-MGBDQ, AB) and 5 ng gDNA or 50,000 circular plasmid DNA (pDNA) molecules as target. Amplification comprised 40 cycles of 15 s at 95 °C and 1 min at 62 °C after initial denaturation at 95°C for 10 min. All rtPCR assays were run on an ABI PRISM 7700 Sequence Detector using the ABI PRISM Sequence Detection Systems software package (version 1.7, AB) for data acquisition and analysis. Real-time PCR results were pasted into a Microsoft Excel spreadsheet template to automate the data manipulations necessary for mtSNP typing.

Preparation of plasmid DNA standards

The entire mtDNA control regions of two individuals classified homoplasmic for either 16519T or 16519C were PCR amplified as described previously [2]. The amplicons were cloned and the resulting plasmids p16519T and p16519C were sequence-verified. Linearized plasmids were quantified with the PicoGreen dsDNA quantitation kit (Molecular Probes, Eugene, OR). Based on these quantification results, pDNA copy numbers were calculated, and the concentrations of circular plasmids adjusted. Equality in DNA concentration for p16519T and p16519C was verified with a SYBR Green I rtPCR assay using the PCR conditions outlined above, primers 16486 F and 16550 R and SYBR Green PCR master mix (AB). Dissociation curve analysis following SYBR Green I assays revealed no evidence for the formation of nonspecific PCR products. Finally, to exclude significant contamination of the pDNA preparations with bacterial genomic DNA and to confirm the plausibility of the calculated pDNA copy numbers, tenfold serial dilutions of both cloned alleles down to nominally 0.1 plasmids per assay were analyzed by quantitative allele-discriminatory rtPCR.

PCR efficiency, linear dynamic range, and absolute quantification

For both SNP alleles, the apparent single-cycle PCR efficiency and the linear dynamic range of the TaqMan assay during the exponential phase of the amplification were determined by means of tenfold serial dilutions of pDNA standards and gDNA reference samples. The respective PCR efficiencies were calculated from the slopes of the best fit log-linear regression lines as $E=10^{(-1/\text{slope})}-1$ and expressed as a percentage. The linear dynamic range of the TaqMan assay represents the range of initial target molecule numbers over which a log-linear relationship with the value of the fractional PCR cycle (C_T) at which the PCR-related signal exceeds the background fluorescence significantly is observed. The number of mitochondrial genome equivalents (mtGE) in unknown gDNA samples was calculated from the C_T values found for both alleles with reference to the parameters of the pDNA calibration curves.

Allelic discrimination and allele proportions in mixed/heteroplasmic samples

The base states at the mitochondrial ntp 16519 as well as allelic proportions in heteroplasmic samples were derived from the initial number of mtGE harboring T and C alleles. For this purpose, stored calibration curves for both alleles were used, as only slight variations in the parameters of the standard curves were observed in initial experiments.

Statistical analyses

All statistical calculations were performed using the routines implemented in GraphPad Prism 4.02 (GraphPad Software, San Diego, CA). For statistical hypothesis testing, the threshold P value was set to $\alpha=0.05$.

Results and discussion

PCR efficiency and linear dynamic range

For serial dilutions of both nonlinearized plasmids p16519T and p16519C used as target in the allele-discriminatory TaqMan assay, the observed difference between the slopes of the best fit lines obtained by linear regression analysis was statistically insignificant (ANCOVA, $P=0.87$). Furthermore, there were no significant differences between the slopes of the linear regression lines and a hypothetical calibration curve with a slope of −3.322 (i.e., 100% PCR

efficiency; F test, P_{16519T}=0.75, P_{16519C}=0.64). The calculated single-cycle PCR efficiencies during the exponential phase of the amplification for p16519T and p16519C were 100.4 and 100.8%, respectively (Fig. 1; Table 1). This is of great practical importance, as equal and high PCR efficiencies (ideally 100%) for both alleles are a basic requirement for high assay sensitivity and reliable detection of low-level point heteroplasmy with the allele-discriminatory TaqMan approach.

The results from the analysis of serial dilutions of four gDNA reference samples (A–D) were in excellent agreement with the findings of experiments using pDNA standards as target (Fig. 1; Table 1). None of the observed differences in the apparent PCR efficiencies between pDNA standards and gDNA samples was statistically significant (ANCOVA, Table 1). This enabled the use of non-linearized plasmids as comparative standards for absolute quantification of amplifiable mtDNA molecules in unknown DNA samples. The lower end of the linear dynamic range was approximately 10 double-stranded copies for both cloned alleles. Virtually the same values were obtained for homoplasmic and heteroplasmic gDNA samples (Fig. 1; Table 1).

The highest concentration we used in pDNA standard curves was 10^6 dsDNA molecules per reaction. For the two most concentrated gDNA samples in our study, approximately 13,500,000 (16519C) and 10,000,000 (16519T) mtGE per reaction were estimated. Despite the fact that nondefault settings for baseline determination had to be used for these gDNA samples, valid allele designations were made by means of absolute quantification. Taking these results together, the linear dynamic range of the

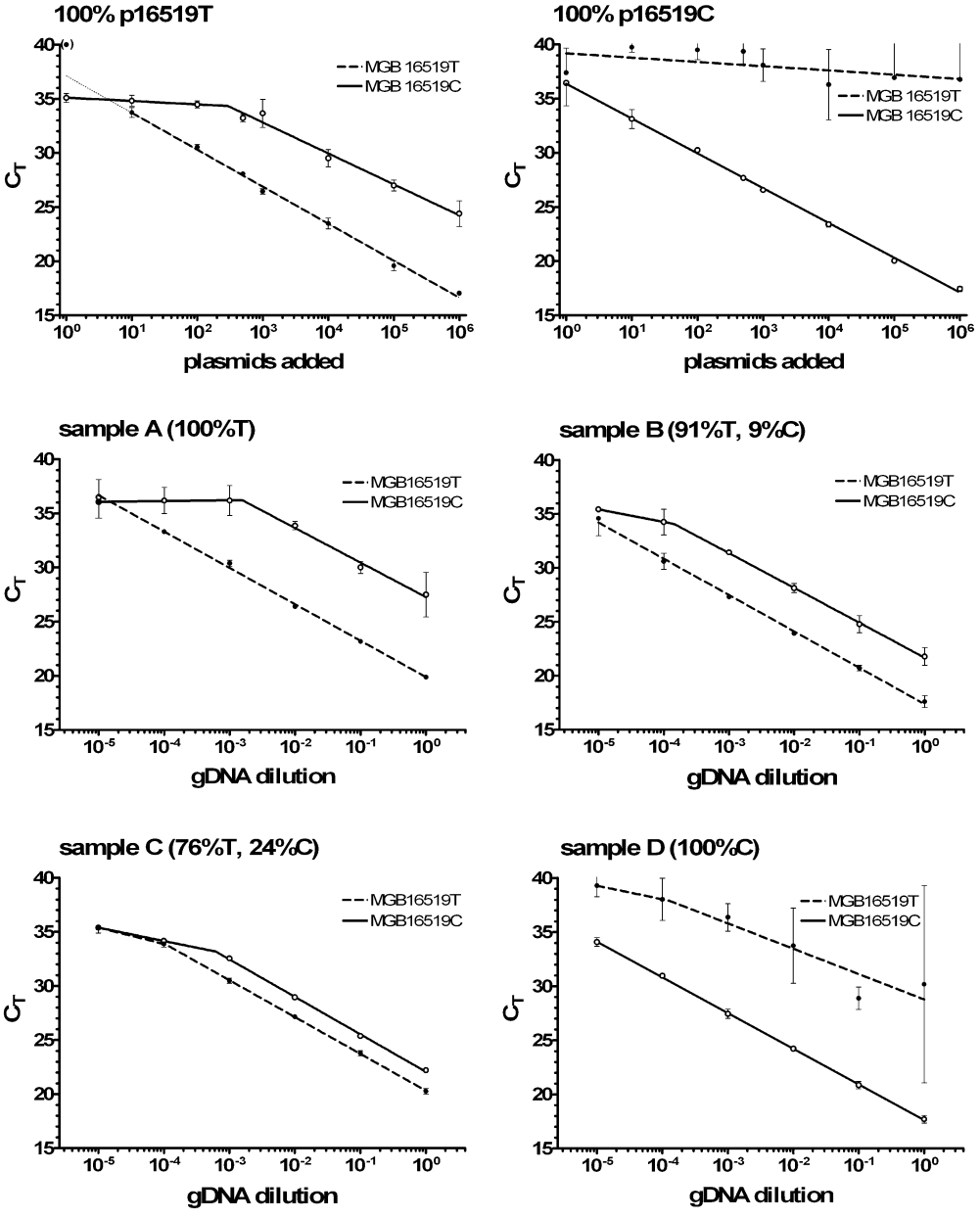

Fig. 1 Dynamic range of the allele-discriminatory rtPCR assay for both cloned 16519 alleles and four gDNA reference samples. Mean values and standard deviations of quadruplicate measurements are shown, and lines represent the results of linear regression or two-segment linear regression analysis

Table 1 Parameters of the pDNA standard curves for both alleles of the mtSNP 16519 and serial dilutions of the homoplasmic and heteroplasmic gDNA reference samples A–D

	p16519T (standard curve), 100% T	Sample A, 100% T	p16519C (standard curve), 100% C	Sample D, 100% C	Sample C		Sample B	
					76% T	24% C	91% T	9% C
Slope±SE, deviation from standard curve significant? (*P*)	−3.312±0.033 (–)	−3.355±0.097, no (0.520)	−3.303±0.041 (–)	−3.296±0.043 no (0.922)	−3.393±0.042, no (0.247)	−3.462±0.056, no (0.151)	−3.369±0.109, no (0.418)	−3.226±0.162, no (0.517)
Y-intercept±SE	36.57±0.09	(–)	36.83±0.11	(–)	(–)	(–)	(–)	(–)
r^2	0.992	0.992	0.989	0.998	0.999	0.998	0.990	0.985
95% CI slope	−3.377 to −3.246	−3.570 to −3.139	−3.385 to −3.221	−3.393 to −3.199	−3.490 to −3.295	−3.599 to −3.325	−3.610 to −3.127	−3.623 to −2.828
95% CI *Y*-intercept	36.39 to 36.74	(–)	36.61 to 37.05	(–)	(–)	(–)	(–)	(–)
PCR efficiency	100.4%	98.6%	100.8%	101.1%	97.1%	94.5%	98.1%	104.2%
Linear dynamic range	10–1,000,000 copies p16519T	10–(109,689) mtGE 16519T	1–1,000,000 copies p16519C	7–(613,503) mtGE 16519C	7–(83,636) mtGE 16519T	13–(26,908) mtGE 16519C	4–(529,811) mtGE 16519T	7–(35,686) mtGE 16519C

For undiluted gDNA samples, the calculated initial copy numbers were set in parentheses, as they do not necessarily represent the upper limit of the linear dynamic range
SE standard error of the mean, *CI* confidence interval, *P* probability of Type I error, r^2 square of the correlation coefficient, *mtGE* mitochondrial genome equivalents

TaqMan assay covered 5–6 orders of magnitude (Fig. 1; Table 1), starting from approximately 10 mtGE. This is of practical interest for routine applications, as it allows the investigation of a broad range of specimens under identical technical conditions, regardless of DNA concentration. The low detection limit for both alleles strongly reduces also the risk of "allele drop out" when genotyping point-heteroplasmic or multicontributor low copy number DNA samples, particularly when the allele proportions differ markedly from each other. Furthermore, the short amplicon length (65 bp) that is obtained with the presented assay enables the analysis of even severely fragmented DNA.

Concordance study

In a concordance study, 405 gDNA samples with known control region sequences were typed with the allele-discriminatory rtPCR assay. For 15 out of the 405 samples, no automated 16519T/C allele assignments were possible. In all of these instances, the background corrected normalized endpoint fluorescence readings did not reach the preset threshold of 0.4 arb. units. Inspection of the sequence data revealed the presence of additional template: probe mismatches (16518G→T, 16518G→T plus 16527C→T, 16524A→G, 16526G→A, 16527C→T) in all of these samples. No nonmatching bases were found in the amplification–primer binding sequences of the 405 gDNA samples in this study. Thus, only the signal generating process was affected, not the amplification reaction itself. This is consistent with the observation that for the affected samples, the C_T values of the major allele were well below the preset threshold value of 30. This allowed correct 16519T/C allele designations despite the low signal intensities. However, the lowered amount of fluorochrome molecules liberated per PCR cycle clearly impaired the quantification results.

For the remaining 390 samples, the automated genotyping results for the mtSNP 16519T/C (108× 16519T, 280× 16519C, 2× heteroplasmy T>C) were fully concordant with the information obtained by Sanger type sequencing.

Detection of mixtures/point heteroplasmy

To determine the detection limit of the allele discriminating rtPCR assay for low-level point heteroplasmy, we performed 384 replicate reactions for each of the two cloned alleles (Fig. 2). The extent of nonspecific cross hybridization of the TaqMan probes (particularly MGB16519C, see also Fig. 1) varied markedly and reproducibly between different wells in the thermal cycler block, showing the worst results in peripheral wells (data not shown). The lowest measured major allele proportions for p16519T and p16519C were 92.1 and 95.1%, respectively (Fig. 2). The analysis of homoplasmic gDNA samples in the concordance study resulted in similar frequency distribution patterns for the obtained allele proportion estimates (Fig. 2).

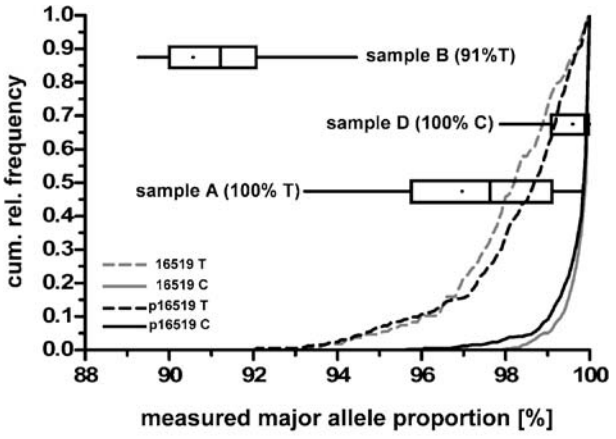

Fig. 2 Cumulative relative frequencies of the measured allele proportions of pure pDNA standards (p16519T and p16519C, 384 replicates each) and the major allele in homoplasmic gDNA samples classified as homoplasmic for 16519T (n=108) or C (n=280) by means of DNA sequencing analysis. Horizontal box-and-whiskers plots (minimum, 25th percentile, median, 75th percentile, maximum) quantify the variability of the major allele proportions measured for two homoplasmic (**A**, **D**) and one heteroplasmic (**B**) gDNA reference sample in the intra-run variability study (8 replicates per sample); dots in the boxes depict the results obtained in the concordance study

Here, the lowest measured major allele proportions were 93.3% T and 98.0% C.

The degree of the intra-run variability in relation to the base state at ntp 16519 was assessed by examining two homoplasmic and two heteroplasmic gDNA reference samples in eight replicate reactions each. All of the single allele proportions found by the replicate testing of the two homoplasmic samples A (mean±standard deviation 97.4± 1.4% T) and D (99.5±0.3% C) fell perfectly in the range of results found for pDNA standards and the homoplasmic gDNA samples in the concordance study (Fig. 2). All replicate reactions of sample C (76.4±3.2% T) were typed invariably as heteroplasmic. However, the nonspecific cross hybridization of the TaqMan probes significantly compromises the unambiguous detection of low-level point heteroplasmy, as was seen with reference sample B (91.0± 1.5% T). For the individual major allele proportion estimates for this sample, a small overlap with pure 16519T targets was observed (Fig. 2). Hence, running samples in replicates are necessary to assure reliable detection and quantification of low-level point heteroplasmy.

Taking all these results together, the quantitative allele-discriminatory TaqMan assay appears to detect point heteroplasmy or mixtures at least as well as direct sequencing, the "gold standard" in forensic mtDNA testing. Contrary to DNA sequencing, SNP genotyping, mixture/heteroplasmy detection, and assessment of allele proportions with allele-discriminatory rtPCR can be easily automated. Furthermore, quantitative PCR permits comparison of the PCR template quantities present in a sample extract and its associated controls. This may permit laboratories the formulation of objective criteria for distinguishing genuine typing results from those that potentially result from sporadic contamination or postmortem DNA damage [4].

Temporal reproducibility

In order to assess the reproducibility of the results obtained with the allele-discriminatory TaqMan assay over time, we reanalyzed a set of 30 samples comprising 28 randomly chosen gDNA preparations and the heteroplasmic reference samples B and C. The timespan between the first and the second rtPCR run ranged between 197 and 456 days. The minimum and maximum ages of the primers and probes were 77 and 533 days post-synthesis when used for analysis.

For the homoplasmic samples, all allele proportion estimates obtained in the second round of rtPCR analysis fell perfectly into the range of values we observed in the concordance study (Fig. 2). For homoplasmic samples, the largest observed differences between the first and second typing were 1.2% (100% C vs 98.8% C) and 5.3% (100% T vs 94.7% T), and the median differences were 0.1% for 16519C and 0.9% for 16519T. The observed differences between the two measurements of the two heteroplasmic reference samples B and C were 1.3 and 4.4%, respectively. These differences detected over time were comparable to those found in the intra-run variation study, in which the respective differences between the maximum and minimum allele proportion estimates for these samples amounted to 4.4 and 9.9%.

The variation in the estimated mtGE found in the temporal reproducibility study (at worst about a factor of 2) was similar to the results we obtained with a commercially available rtPCR kit for quantification of amplifiable human nuclear DNA (data not shown) and to data from the literature [5, 9].

Conclusions

Forensic mtDNA typing is in its essence SNP typing, even if complete sequencing of target regions is employed to obtain the SNP data. As we look either to augment mtDNA testing by going beyond HV1 and HV2 or to replace sequencing with techniques that may be faster, better, or cheaper, it is important to rigorously evaluate alternative SNP detection assays, with real world forensic applications in mind. Here we have examined the performance of an allele-discriminatory quantitative TaqMan rtPCR SNP assay with regard to parameters highly relevant to mtDNA forensic typing: sensitivity, specificity, range of suitable template concentrations, mixture/heteroplasmy detection, reproducibility, ease of data analysis, and quantitative performance. In all these areas, the assay performed quite well for forensic purposes, with particularly strong capabilities for quantifying allele proportions, contamination avoidance due to homogeneous assay format, and for absolute template copy number quantitation. The quantitative

nature of the TaqMan assay makes it unique among recent SNP typing methods, providing the potentially very useful means for distinguishing between authentic signal and possible contamination based on measurable values. These features make it a valuable tool for forensic applications, complementing endpoint-signal-based, multistep SNP typing approaches with a high multiplexing capability, such as single nucleotide primer extension assays (e.g., [1, 10]).

Acknowledgements Part of this work was supported by National Institute of Justice grant 2000-1J-CX-K010 to T.J.P. The opinions and assertions contained herein are solely those of the authors and are not to be construed as official or as views of the U.S. Department of Defense, the U.S. Department of the Army, or the U.S. Department of Justice.

Reference

1. Brandstätter A, Parsons TJ, Parson W (2003) Rapid screening of mtDNA coding region SNPs for the identification of west European Caucasian haplogroups. Int J Leg Med 117:291–298
2. 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 Leg Med 118:47–54
3. Coble MD, Just RS, O'Callaghan JE, Letmanyi IH, Peterson CT, Irwin JA, Parsons TJ (2004) Single nucleotide polymorphisms over the entire mtDNA genome that increase the forensic power of testing in Caucasians. Int J Leg Med 118:137–146
4. Gilbert MT, Willerslev E, Hansen AJ, Barnes I, Rudbeck L, Lynnerup N, Cooper A (2003) Distribution patterns of post-mortem damage in human mitochondrial DNA. Am J Hum Genet 72:32–47; erratum in: Am J Hum Genet 72:779
5. Kline MC, Duewer DL, Redman JW, Butler JM (2005) Results from the NIST 2004 DNA quantitation study. J Forensic Sci 50:570–578
6. Lutz S, Weisser HJ, Heizmann J, Pollak S (1998) Location and frequency of polymorphic positions in the mtDNA control region of individuals from Germany. Int J Leg Med 111:67–77; erratum in: Int J Leg Med (1999) 112:145–150
7. Parsons TJ, Coble MD (2001) Increasing the forensic discrimination of mitochondrial DNA testing through analysis of the entire mitochondrial DNA genome. Croat Med J 42:304–309
8. Sobrino B, Carracedo A (2004) SNP typing in forensic genetics: a review. Methods Mol Biol 297:107–126
9. Swillens S, Goffard JC, Maréchal Y, de Kerchove d'Exaerde A, El Housni H (2004) Instant evaluation of the absolute initial number of cDNA copies from a single real-time PCR curve. Nucleic Acids Res 32:e56
10. Vallone PM, Just RS, Coble MD, Butler JM, Parsons TJ (2004) A multiplex allele-specific primer extension assay for forensically informative SNPs distributed throughout the mitochondrial genome. Int J Leg Med 118:147–157