Int J Legal Med (2005) 119: 50–53
DOI 10.1007/s00414-004-0487-7

SHORT COMMUNICATION

Ukhee Chung · Hwan Young Lee · Ji-Eun Yoo · Myung Jin Park · Kyoung-Jin Shin

Mitochondrial DNA CA dinucleotide repeats in Koreans: the presence of length heteroplasmy

Received: 24 May 2004 / Accepted: 16 August 2004 / Published online: 17 September 2004

Abstract The mitochondrial DNA HV3 region contains CA dinucleotide repeats which display length polymorphism. To analyze this for forensic purposes, we designed a fluorescence-labelled PCR primer set for short amplification products and carried out genotyping by using capillary electrophoresis. A total of 4 alleles with 4–7 repeat units were observed and the genetic diversity was estimated to be 0.5120 in 500 unrelated Koreans. Interestingly, three individuals showed two or three length variants, i.e. length heteroplasmy.

Keywords Mitochondrial DNA · CA dinucleotide repeats · Length polymorphism · Length heteroplasmy · Koreans

Introduction

The analysis of human mitochondrial DNA (mtDNA) is of central importance for forensic identity testing. The properties of mtDNA that make it valuable for human identification include a high copy number, maternal inheritance and a rapid rate of evolution [1, 2]. Especially, the two non-coding HV1 (first hypervariable region) and HV2 (second hypervariable region) fragments of the control region, which are the most polymorphic regions in mtDNA, generally have been targeted for forensic purposes and analyzed in several kinds of biological evidence [3, 4, 5]. Although HV1/HV2 haplotypes have a high evidential value, other regions within the control region and variations within the mtDNA coding region with forensic utility were sometimes targeted for better differentiation of mtDNA [6, 7, 8, 9, 10].

It is noteworthy that CA dinucleotide repeats between nucleotide position 514 and 523 in HV3 (third hypervariable region) show length variability like STRs (short tandem repeats) in nuclear DNA [11, 12, 13], and the genetic diversity is relatively high. As this length polymorphism was expected to provide additional information for forensic purposes with simple genotyping methods as for STR analysis, i.e. PCR and capillary electrophoresis, we analyzed the CA dinucleotide repeats using a fluorescence-labelled PCR primer set for short amplification products in 500 unrelated Koreans. Also, we found three heteroplasmic samples during the analyses and confirmed the presence of length heteroplasmy by identifying each length variant in an individual.

Materials and methods

DNA samples

Buccal swabs or blood samples were obtained from 500 unrelated Koreans for the present study. DNA was extracted using the QIAamp DNA Mini kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The cell line sample 9948 DNA (Promega, Madison, WI) was used for control DNA to calibrate the allelic ladder.

PCR amplification and genotyping

Primers were designed to amplify CA dinucleotide repeats in the HV3 region using the Primer3 program (http://

U. Chung · H. Y. Lee · J.-E. Yoo · M. J. Park · K.-J. Shin
Department of Forensic Medicine and Brain Korea 21 Project for Medical Science, College of Medicine Yonsei University, 134 Shinchon-Dong Seodaemun-Gu, 120–752 Seoul , Korea

U. Chung · K.-J. Shin
Biometrics Engineering Research Center, Yonsei University, 134 Shinchon-Dong Seodaemun-Gu, 120–749 Seoul, Korea

K.-J. Shin (✉)
Human Identification Research Institute, Yonsei University, 134 Shinchon-Dong Seodaemun-Gu, 120–752 Seoul , Korea
e-mail: kjshin@yumc.yonsei.ac.kr
Tel.: +82-2-361-5775
Fax: +82-2-362-0860

www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi). PCR amplification was carried out in a 10 µl reaction volume containing 10–20 pg of template DNA and the following set of primers: 0.2 µM of F447 (5'-HEX-CATTATTTTCCCCTCCCACTCC) and 0.2 µM of R569 (5'-GGTGTCTTTGGGGTTTGGTTG). Thermal cycling was conducted using a PTC-200 DNA engine (MJ Research, Waltham, MA) under the following conditions: 95°C for 11 min, 25 cycles of 94°C for 1 min, 56°C for 1 min, 72°C for 1 min and a final extension at 60°C for 60 min. The PCR products were analyzed by capillary electrophoresis on an ABI PRISM 310 Genetic Analyzer (Applied Biosystems, Foster City, CA) using GeneScan analysis software version 3.1 (Applied Biosystems, Foster City, CA). Allele typing was carried out based on the sequenced allelic ladder using Genotyper 2.5 software (Applied Biosystems, Foster City, CA).

Sequence analysis

To analyze sequences including CA dinucleotide repeat units in the HV3 region, mtDNA was amplified in a total volume of 25 µl using an unlabelled primer set. PCR was performed under the same conditions of GeneScan PCR except for 35 thermal cycles. The PCR products were purified using a QIAquick purification kit (Qiagen, Hilden, Germany) and sequenced from both ends with a BigDye Terminator Cycle Sequencing v2.0 Ready Reaction kit (Applied Biosystems, Foster City, CA). The resulting sequences were analyzed using Sequence Analysis Software Version 3.4 and Sequence Navigator 1.01 (Applied Biosystems, Foster City, CA). The genetic diversity of CA dinucleotide repeats was calculated according to Tajima [14].

Confirmation of length heteroplasmy by cloning

The PCR products of heteroplasmic samples which showed consecutive multiple peaks with two base differences in the GeneScan electropherogram were cloned into the pGEM-T Easy vector (Promega, Madison, WI) according to the manufacturer's instructions. Each clone was sequenced as previously described. In addition, clones with various repeat unit lengths were subjected to GeneScan analysis to investigate whether stutter occurs due to insertion or deletion during the PCR amplification of CA dinucleotide repeats.

Results and discussion

The constructed allelic ladder for CA dinucleotide repeats in the HV3 region and an example of genotyping results are shown in Fig. 1A. The 9948 control DNA displayed the genotype 5. Allele frequencies, genetic diversities and sequence structures of CA dinucleotide repeats are shown in Table 1. There are some variant allele sequence structures. The A to C transversion and C to T transition in CA dinucleotide repeat units were observed in alleles 4 and 5, respectively. Also, a G to A transition in the flanking region was observed in alleles 4 and 5. The genetic diversity calculated with consideration of the flanking region sequence structure is 0.5120, which is higher than 0.4951 when considering only the amplification product size. According to previous reports, the genetic diversities of Germans, Russians, Hungarians and from the Cameroons were 0.3600, 0.3400, 0.4000 and 0.5200, respectively. The genetic diversity of the Korean population (0.4951) is relatively higher than those of Germans, Russians and Hungarians, while closer to the Cameroons [11]. The most abundant allele in Koreans (allele 5) is the same as in Germans, Russians and Hungarians, but not in the Cameroons (allele 4). Although there are some distinctions in genetic diversities and allele distributions among geographically different populations, it can be easily inferred that the combined use of length polymorphism in CA dinucleotide repeats in the HV3 region with sequence polymorphism in HV1 and HV2 regions will increase the forensic utility of mitochondrial DNA to some extent. Moreover, the measuring method for CA dinucleotide repeats designed in this study is simpler and faster than sequencing analysis, and the PCR product size is small enough (122–128 bp) to analyze degraded or low copy DNA or hair shaft DNA without nuclear DNA with ease.

On the other hand, we found three heteroplasmic samples during the course of analysis. One had length variants with 4 and 5 CA repeat units, another had 5 and 6

Table 1 The frequencies, sequence structures and genetic diversities of the CA dinucleotide repeats in the mtDNA HV3 region

Number of CA repeats	N[a]	Frequency	Sequence structure	N[a]	Frequency
4	186	0.3720	CCAG $(CA)_4$ CCGC	181	0.3620
			CCAA $(CA)_4$ CCGC	4	0.0080
			CCAG $(CA)_3$CC CCGC[b]	1	0.0020
5	303	0.6060	CCAG $(CA)_5$ CCGC	299	0.5980
			CCAA $(CA)_5$ CCGC	1	0.0020
			CCAG TA $(CA)_4$ CCGC[c]	3	0.0060
6	7	0.0140	CCAG $(CA)_6$ CCGC	7	0.0140
7	4	0.0080	CCAG $(CA)_7$CCGC	4	0.0080
Total	500			500	
Genetic diversity		0.4951			0.5120

[a]The number of individuals.
[b, c]Classified as allele 4 and allele 5 as the lengths of PCR products are identical to those of $(CA)_4$and $(CA)_5$, respectively.

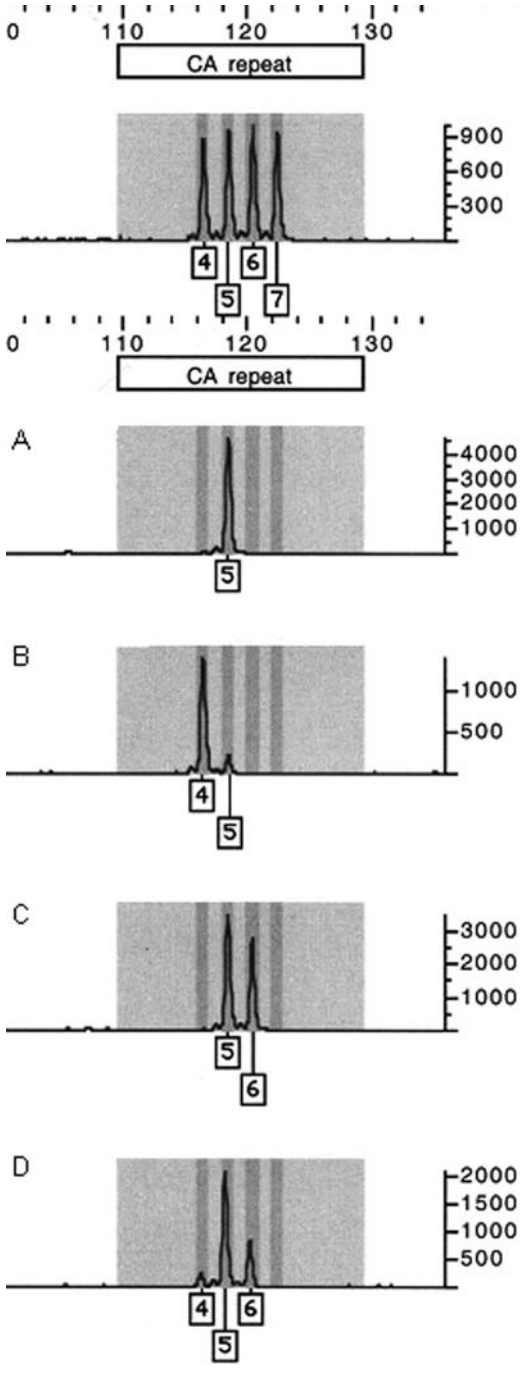

Fig. 1A–D The electropherograms showing genotype results of normal and heteroplasmic mtDNA. Constructed allelic ladder is shown at the top. **A** 9948 male standard DNA of which sequence is CCAG $(CA)_5$ CCGC in CA dinucleotide repeat area. **B** Heteroplasmic mtDNA which contains $(CA)_4$ and $(CA)_5$. **C** Heteroplasmic mtDNA which contains $(CA)_5$ and $(CA)_6$. **D** Heteroplasmic mtDNA which contains $(CA)_4$, $(CA)_5$ and $(CA)_6$

CA repeat units and the other had 4, 5 and 6 CA repeat units (Fig. 1B–D). The presence of length heteroplasmy was established if multiple peaks with two base differences were found in the GeneScan electropherograms. Also, we analyzed the presence of length heteroplasmy by cloning, and the results coincided with those of the GeneScan analysis. According to previous reports [15, 16, 17], direct sequencing, fluorescence-labelled restriction fragment analysis and cloning analysis have been used to detect and identify length heteroplasmy in mitochondrial HV1 and HV2 regions. In comparing the detection limit of the three methods, the level of detection was found to increase in the order from direct sequencing of the mtDNA to restriction fragment analysis, to the cloning assay [17]. However, restriction fragment analysis was thought to be a sufficiently accurate alternative for most applications involving heteroplasmy analysis. In the present study, our previous method that is similar to the fluorescence-labelled restriction fragment analysis with some modifications, was employed [18].

As for the mechanism of heteroplasmy formation in the CA dinucleotide repeats, their propensity for insertion-deletion mutation of multiples of repeating unit during replication could explain the phenomena along with the low fidelity of mitochondrial DNA polymerase. The high slippage rate of dinucleotide repeats can be observed in STRs, and resultant stutter production is correlated to the length of repeat stretches consisting of uniform repeats [19]. In this study, it was also demonstrated that the PCR amplification of CA dinucleotide repeats did not produce stutters for up to 6 CA repeat units, while clones with 7 CA repeats showed traces of stutter in the electropherograms. Taken together, polymerase slippage is the driving force not only in stutter artifact genesis during PCR but also causing new mutations (plus/minus one repeat) in dinucleotide repeats.

In summary, the analysis of length variability of CA dinucleotide repeats in the HV3 region will help increase the power of discrimination of mitochondrial DNA sequence analyses. Also, the length heteroplasmy of CA dinucleotide repeats in the HV3 region will increase the chance of forensic identification and help to understand the mechanism of mitochondrial evolution and microsatellite formation.

Acknowledgements This work was supported by a grant from the Korea Science and Engineering Foundation (KOSEF) through the Biometrics Engineering Research Center (BERC) at Yonsei University.

References

1. Anderson S, Bankier AT, Barrell BG et al. (1981) Sequence and organization of the human mitochondrial genome. Nature 290:457–465
2. Cann RL, Stoneking M, Wilson AC (1987) Mitochondrial DNA and human evolution. Nature 325:31–36
3. Holland MM, Parsons TJ (1999) Mitochondrial DNA sequence analysis—validation and use for forensic casework. Forensic Sci Rev 11:21–50
4. Gill P, Ivanov PL, Kimpton C et al. (1994) Identification of the remains of the Romanov family by DNA analysis. Nat Genet 6:130–135
5. Wilson MR, DiZinno JA, Polanskey D, Replogle J, Budowle B (1995) Validation of mitochondrial DNA sequencing for forensic casework analysis. Int J Legal Med 108:68–74

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 Legal Med 111:67–77, Erratum in Int J Legal Med (1998) 111:286, Int J Legal Med (1999) 112:145–150
7. Lee SD, Lee YS, Lee JB (2002) Polymorphism in the mitochondrial cytochrome B gene in Koreans. An additional marker for individual identification. Int J Legal Med 116:74–78
8. Vanecek T, Vorel F, Sip M (2004) Mitochondrial DNA D-loop hypervariable regions: Czech population data. Int J Legal Med 118:14–18
9. Pereira L, Cunha C, Amorim A (2004) Predicting sampling saturation of mtDNA haplotypes: an application to an enlarged Portuguese database. Int J Legal Med 118:132–136
10. 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 power of forensic testing in Caucasians. Int J Legal Med 118:137–146
11. Szibor R, Michael M, Spitsyn VA, Plate I, Ginter EK, Krause D (1997) Mitochondrial D-loop 3' $(CA)_n$ repeat polymorphism: optimization of analysis and population data. Electrophoresis 18:2857–2860
12. Lutz S, Weisser HJ, Heizmann J, Pollak S (1997) A third hypervariable region in the human mitochondrial D-loop. Hum Genet 101:384
13. Lutz S, Wittig H, Weisser HJ et al. (2000) Is it possible to differentiate mtDNA by means of HVIII in samples that cannot be distinguished by sequencing the HVI and HVII regions? Forensic Sci Int 113:97–101
14. Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595
15. Brandstatter A, Niederstatter 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
16. Pfeiffer H, Lutz-Bonengel S, Pollak S, Fimmers R, Baur MP, Brinkmann B (2004) Mitochondrial DNA control region diversity in hairs and body fluids of monozygotic triplets. Int J Legal Med 118:71–74
17. 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 DOI:10.1007/s00414-004-0457-0
18. Lee HY, Chung U, Yoo JE, Park MJ, Shin KJ (2004) Quantitative and qualitative profiling of mitochondrial DNA length heteroplasmy. Electrophoresis 25:28–34
19. Shinde D, Lai Y, Sun F, Arnheim N (2003) Taq DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: (CA/GT)n and (A/T)n microsatellites. Nucleic Acids Res 31:974–980