Int J Legal Med (2002) 116:74–78

ORIGINAL ARTICLE

Soong Deok Lee · Yoon Seong Lee · Jung Bin Lee

Polymorphism in the mitochondrial cytochrome B gene in Koreans

An additional marker for individual identification

Received: 10 January 2001 / Accepted: 29 May 2001

Abstract Sequencing the mitochondrial control region is very useful for individual identification when conventional DNA typing using autosomal STRs is unavailable. However, low discriminatory power is a problem and another polymorphic locus within the mitochondrial genome is necessary. The cytochrome B (MTCYB) gene, which has undergone several changes during evolution, may be a good candidate for this purpose. Here the sequencing data of the MTCYB gene of 98 unrelated Koreans is presented. A total of 30 polymorphic sites were found which were distributed evenly along the gene. All were nucleotide substitutions and no insertions/deletions were noted. A total of 22 different MTCYB lineages were revealed. Out of 22 different control region lineages with 79 samples which shared the same D-loop sequence with some others within a lineage, 10 lineages with 37 samples could be sub-grouped according to different MTCYB sequences. Some issues concerning the MTCYB gene polymorphism are discussed.

Keywords Mitochondrial DNA · Cytochrome B · Polymorphism · Sequencing · Korea

Introduction

Mitochondrial DNA (mtDNA) has several unique characteristics compared to nuclear DNA [1]. MtDNA exists in numerous copy numbers even within a single cell, it is rather short, 16.5 kb long with a double-stranded circular structure and there is a 1.1 kb long control region (D-loop) which is highly polymorphic. These characteristics of mtDNA make it a useful tool as an individual identification marker, especially when nuclear DNA is unavailable. Furthermore maternal transmission makes the interpretation relatively simple [2].

Individual identification involving samples of poor quality in which autosomal STRs cannot be used, for example degraded samples, old bone fragments or hair shafts without roots, are routinely undertaken by mtDNA sequencing for the control region [3, 4, 5], but poor discriminating power remains a problem. Contrary to conventional DNA typing, which mainly targets autosomal STRs, it is not possible to determine that a particular sample is derived from a specific individual on the basis of mtDNA sequencing results alone. We are obliged to add rather conservatively that "it is possible" that some sample is derived from a specific source, even when the mtDNA control region sequences are identical. Other supportive information is required to achieve specificity, but

Table 1 Amino acid changes resulting from 30 nucleotide substitutions in the MTCYB genes found in this study

Site	Nucleotide change	Result	Site	Nucleotide change	Result
14766	C-T	The-Ile	15244	A-G	Gly-Gly
14783	T-C	Leu-Leu	15301	G-A	Leu-Leu
14867	C-T	Leu-Leu	15314	G-A	Ala-Thr
14927	A-G	Thr-Ala	15323	G-A	Ala-Thr
14944	C-T	Ile-Ile	15326	A-G	Thr-Ala
14979	T-C	Ile-Thr	15341	T-C	Phe-Leu
15037	C-T	His-His	15440	T-C	Leu-Leu
15040	C-T	Ile-Ile	15460	C-T	Ser-Ser
15043	G-A	Gly-Gly	15466	G-A	Met-Met
15067	T-C	Phe-Phe	15487	A-T	Pro-Pro
15071	T-C	Thy-His	15497	G-A	Gly-Ser
15119	G-A	Ala-Thr	15573	T-A	Phe-Tyr
15218	A-G	Thr-Ala	15747	T-C	Ile-Ile
15221	G-A	Asp-Asn	15766	A-G	Gly-Gly
15235	A-G	Trp-Trp	15860	A-G	Ile-Val

Electronic supplementary material to this paper can be obtained by using the Springer Link server located at http://dx.doi.org/10.1007/

S. D. Lee (✉) · Y. S. Lee · J. B. Lee
Department of Forensic Medicine,
Seoul National University College of Medicine,
28 Yeongon-dong, Chongno-gu, Seoul 110-799 Korea
e-mail: sdlee@snu.ac.kr,
Tel.: +82-2-7408353, Fax: +82-2-7648340

Table 2 The control region sequences for all the 98 samples screened in this studied. The sequenced areas were from 16016–16401 in the Anderson sequence for HVI and from 048–388 for HVII

Sample no.	Control region sequence
F408.2	16223T, 16325C, 73G, 150T, 263G, 315.1C
F413.1	16223T, 16325C, 73G, 150T, 263G, 315.1C
F414.2	16223T, 16325C, 73G, 150T, 263G, 315.1C
F452.1	16075C, 16223T, 16325C, 73G, 150T, 263G, 315.1C
F496.1	16223T, 16298C, 16319A, 16362T, 73G, 152C, 263G, 315.1C
F505.2	16223T, 16234T, 16316G, 73G, 263G, 309.1C, 315.1C
F515.2	16223T, 16234T, 16316G, 73G, 263G, 309.1C, 315.1C
F516.1	16129A, 16223T, 73G, 152C, 263G, 309.1C, 315.1C
F531.2	16182C, 16183C, 16189C, 16232A, 16249C, 16304C, 16311C, 16362T, 73G, 249d, 263G, 309.2C, 315.1C
F562.1	16223T, 73G, 146C, 194T, 263G, 309.1C, 315.1C
F569.1	16129A, 16223T, 73G, 152C, 263G, 309.1C, 315.1C
F587.1	16172C, 16223T, 16257A, 16261T, 16362T, 73G, 150T, 263G, 309.1C, 315.1C
F596.1	16172C, 16223T, 16257A, 16261T, 16362T, 73G, 150T, 263G, 309.1C, 315.1C
F640.1	16223T, 73G, 263G, 315.1C
F645.1	16223T, 73G, 263G, 309.1C, 315.1C
F650.1	16223T, 16298C, 16319A, 16362T, 73G, 152C, 263G, 315.1C
F656.2	16223T, 73G, 150T, 194T, 205A, 263G, 315.1C
F657.1	16093C, 16129A, 16223T, 16249C, 73G, 152C, 263G, 315.1C
F694.1	16223T, 73G, 146C, 194T, 263G, 309.1C, 315.1C
F700.1	16223T, 73G, 263G, 315.1C
F701.1	16075C, 16223T, 16325C, 73G, 150T, 263G, 315.1C
F708.1	16129A, 16223T, 73G, 152C, 263G, 309.1C, 315.1C
F718.1	16129A, 16223T, 73G, 152C, 263G, 309.1C, 315.1C
F739.1	16223T, 16234T, 16316G, 73G, 263G, 309.1C, 315.1C
F740.2	16223T, 16234T, 16316G, 73G, 263G, 309.1C, 315.1C
F741.2	16223T, 16298C, 16319A, 16362T, 73G, 152C, 263G, 315.1C
F748.1	16223T, 16298C, 16319A, 16362T, 73G, 152C, 263G, 315.1C
F761.2	16223T, 73G, 263G, 315.1C
F796.5	16223T, 73G, 146C, 194T, 263G, 309.1C, 315.1C
F797.1	16129A, 16223T, 73G, 152C, 263G, 309.1C, 315.1C
F797.2	16223T, 73G, 263G, 315.1C
F820.2	16166G, 16172C, 16223T, 16295T, 16319A, 16362T, 73G, 146C, 199C, 263G, 315.1C
F828.1	16172C, 16223T, 16257A, 16261T, 16362T, 73G, 150T, 263G, 309.1C, 315.1C
F833.2	16129A, 16223T, 16309G, 73G, 146C, 263G, 309.1C, 315.1C
F844.1	16075C, 16223T, 16325C, 73G, 150T, 263G, 315.1C
F852.1	16172C, 16223T, 16257A, 16261T, 16362T, 173G, 150T, 263G, 309.1C, 315.1C
F861.2	16223T, 73G, 263G, 309.1C, 315.1C
F884.2	16093C, 16129A, 16223T, 16249C, 73G, 152C, 263G, 315.1C
F907.1	16223T, 16311C, 16319A, 73G, 263G, 315.1C
F908.1	16172C, 16223T, 16257A, 16261T, 16362T, 73G, 150T, 263G, 309.1C, 315.1C
F912.1	16223T, 73G, 263G, 309.1C, 315.1C
F916.2	16223T, 16325C, 73G, 150T, 263G, 315.1C
H108	16172C, 16223T, 16257A, 16261T, 73G, 150T, 263G, 309.1C, 315.1C
H120	16223T, 16325C, 16362C, 73G, 150T, 183G, 184A, 263G, 309.1C, 315.1C
H140	16187T, 16223T, 16290T, 16319A, 73G, 235G, 263G, 315.1C
H153	16129A, 16223T, 16362C, 73G, 152C, 263G, 309.1C, 315.1C
H159	16129A, 16140C, 16187T, 16189C, 16266G, 73G, 93G, 210G, 263G, 309.1C, 315.1C
H166	16129A, 16140C, 16187T, 16189C, 16266G, 73G, 93G, 210G, 263G, 309.1C, 315.1C
H17	16223T, 16234T, 16316G, 73G, 263G, 309.1C, 315.1C
H17	16223T, 16311C, 73G, 263G, 315.1C
H173	16223T, 16362C, 73G, 263G, 315.1C
H195	16129A, 16140C, 16187T, 16189C, 16266G, 73G, 93G, 210G, 263G, 309.1C, 315.1C
H252	16223T, 16362C, 73G, 263G, 315.1C
H255	16209C, 16223T, 16324C, 73G, 263G, 315.1C
H259	16223T, 16362C, 73G, 263G, 315.1C
H303.1	16187T, 16223T, 16290T, 16319A, 73G, 235G, 263G, 309.1C, 315.1C

Table 2 (continued)

Sample no.	Control region sequence
H315	16187T, 16223T, 16290T, 16319A, 73G, 235G, 263G, 315.1C
H334	16223T, 16362C, 73G, 263G, 315.1C
H348	16129A, 16223T, 73G, 152C, 263G, 309.1C, 315.1C
H361	16223T, 16311C, 73G, 263G, 315.1C
H362	16129A, 16232A, 16304C, 16311C, 16344T, 73G, 152C, 249d, 263G, 309.2C, 315.1C
H366.1	16129A, 16148T, 16223T, 73G, 152C, 263G, 309.1C, 315.1C
H377.1	16172C, 16223T, 16257A, 16261T, 73G, 150T, 263G, 309.1C, 315.1C
H473.2	16129A, 16223T, 73G, 152C, 263G, 309.1C, 315.1C
H64	16223T, 16396G, 73G, 150T, 194T, 205A, 263G, 315.1C
H75	16223T, 16311C, 73G, 263G, 315.1C
H81	16187T, 16223T, 16290T, 16319A, 73G, 235G, 263G, 309.1C, 315.1C
H84	16209C, 16223T, 16324C, 73G, 207A, 263G, 315.1C
H98	16187T, 16223T, 16319A, 73G, 235G, 263G, 315.1C
P13	16172C, 16362C, 73G, 194T, 215G, 263G, 315.1C
P149	16223T, 16311C, 73G, 263G, 315.1C
P157	16174T, 16223T, 16362C, 73G, 146C, 183G, 263G, 309.1C, 315.1C
P165	16172C, 16362C, 73G, 199T, 215G, 263G, 315.1C
P173	16223T, 16325C, 16362C, 73G, 150T, 183G, 184A, 263G, 309.1C, 315.1C
P174	16209C, 16223T, 16291T, 16324C, 16362C, 73G, 263G, 315.1C
P181	16223T, 16362C, 73G, 146C, 194T, 263G, 309.1C, 315.1C
P184	16223T, 16362C, 73G, 263G, 315.1C
P189	16223T, 16362C, 73G, 146C, 194T, 263G, 309.1C, 315.1C
P19	16223T, 73G, 146C, 194T, 263G, 309.1C, 315.1C
P191	16129A, 16223T, 73G, 152C, 263G, 309.1C, 315.1C
P200	16129A, 16223T, 73G, 152C, 263G, 309.1C, 315.1C
P213	16129A, 16223T, 16362C, 73G, 152C, 263G, 309.1C, 315.1C
P222	16093C, 16129A, 16223T, 16249C, 16362C, 73G, 152C, 263G, 315.1C
P263	16187T, 16223T, 16290T, 16319A, 73G, 235G, 263G, 309.1C, 315.1C
P39	16223T, 16294T, 16295T, 73G, 146C, 199C, 263G, 315.1C
P62	16223T, 16290T, 73G, 263G, 315.1C
P86	16223T, 16362C, 73G, 263G, 315.1C
P9	16223T, 16311C, 73G, 263G, 309.1C, 315.1C
P90	16093C, 16129A, 16223T, 16249C, 16362C, 73G, 152C, 263G, 315.1C
P98	16223T, 16362C, 73G, 152C, 263G, 315.1C
SB41	16129A, 16232A, 16304C, 16311C, 16344T, 73G, 152C, 263G, 309.2C, 315.1C
SB59	16174T, 16223T, 16362C, 73G, 146C, 183G, 263G, 309.1C, 315.1C
SB63	16129A, 16148T, 16223T, 73G, 152C, 263G, 309.1C, 315.1C
SB67	16129A, 16223T, 16362C, 73G, 152C, 263G, 309.1C, 315.1C
SB77	16223T, 16290T, 16319A, 73G, 152C, 235G, 263G, 309.1C, 315.1C
SB8	16223T, 73G, 263G, 309.1C, 315.1C
SD41	16129A, 16223T, 16362C, 73G, 152C, 263G, 309.1C, 315.1C
SD99	16129A, 16140C, 16187T, 16189C, 16266G, 73G, 93G, 210G, 263G, 309.1C, 315.1C

given the situation that nuclear DNA is unavailable for testing, it is not easy to increase the discriminating power of the test. The only solution would appear to be to find more polymorphism within mtDNA [6].

The cytochrome B (MTCYB) gene located in the mitochondrial genome, has been found to be rather polymorphic and has undergone several changes during evolution [7]. It has been used to identify species [8, 9, 10] but the degree of polymorphism in humans is not known and should be if it is to be used as an individual identification marker. In this study we determined the nature of the MTCYB gene polymorphism in 98 Koreans and discuss its usefulness for individual identification.

Materials and methods

A total of 21 placenta and 77 blood samples were collected at Seoul National University (SNU) hospital and the Department of Forensic Medicine, SNU Medical College. These 98 samples could be grouped into 41 different mtDNA control region lineages whereby 79 samples shared a control region sequence with some other samples grouped into 22 different lineages but all subjects studied here were maternally unrelated. Total DNA was extracted from placental tissue and peripheral blood as described previously using phenol/chloroform extraction [11]. Amplification for the MTCTB gene (from the Anderson sequence nt 14747–15887) was done using the following primer sets:

- F14737:5'-AAG AAC ACC AAT GAC CCC AA-3'
- R15892:5'-AGG ACA GGC CCA TTT GAG TA-3'

Amplification was carried out in 100 µl reaction volume containing 5 U of Taq DNA polymerase, 1 µM each primer, 0.25 mM each dNTP, 10 mM Tris-HCl (pH 8.8), 50 ml KCl and 0.1% Triton X-100. Thermal cycling conditions were 35 cycles at 95 °C for 30 s, 60 °C for 30 s and 72 °C for 1 min using the GeneAmp PCR System 9600 (Perkin-Elmer, Foster City, Calif.). After amplification the products were confirmed using 1.2% agarose gel electrophoresis and purified using the High Pure PCR product purification kit (Boehringer Mannhein, Mannheim, Germany). After quantification using DyNA Quant 200 (Hoefer Scientific Instruments, San Francisco, Calif.), the eluted DNA was used as the template for sequencing which was performed using the DyeDeoxy Terminator cycle sequencing kit and a 377 model ABI automatic sequencer (Applied Biosystems, Foster City, Calif.) following the supplier's recommendations. The primers used in amplification and the additional internal primers:

- R15096:5'-ATG CCG ATG TTT CAG GTT TC-3'
- F14970:5'-ATG GCT GAA TCA TCC GCT AC-3'
- R15337:5'-GAG GTG GAG TGT TGC TAG GG-3'
- F15256:5'-AGA CAG TCC CAC CCT CAC AC-3'
- R15623:5'-CAA GGA CGC CTC CTA GTT TG-3'
- F15575:5'-ACA CAA TTC TCC GAT CCG TC-3'

were also used as sequencing primers. A total of eight sequencing reactions were performed for each sample.

Results and discussion

With the primers described above, direct sequencing of the PCR products of the MTCYB gene was satisfactory covering both strands and results were obtained for all samples. Compared to the Anderson sequence, 30 sites of different nucleotide sequence were noted (Table S1), among which 16 sites were previously registered in the human mitochondrial genome database, MITOMAP (http://www.gen.emory.edu/mitomap). The majority of changes were noted in less than 10% of the subjects but nucleotide changes at position 14766 (all the 98 samples, this site has been proved that the Anderson sequence was mistyped [12] and the data should be interpreted accordingly), 14783 (79 of 98 samples), 14979 (24 of 98 samples), 15043 (79 of 98 samples), 15301 (79 of 98 samples) and 15326 (96 of 98 samples) were found relatively frequently. Each sample contained on average 5.36 polymorphic nucleotide sites. All the changes noted were nucleotide substitutions and no insertions or deletions were found. As in the control region, the transition type was more prevalent than the transversion type. According to the mitochondrial genetic code 16 variations out of 30 found in this study do not cause an amino acid change and only 14 variations encode different amino acids. Table 1 shows the change of amino acid encoding resulting from the nucleotide variation.

A total of 22 different genotypes were found, among these 1 type was shared by 34 samples and another was shared by 19. The genetic diversity was found to be 0.825 and the probability of two unrelated persons having the same sequence was 17.4%. The degree of polymorphism in MTCYB gene seems to be much less than that of the control region. The total number of varying nucleotide sequences, the number of such loci in one sample and the number of different lineages are quite low. Nevertheless, the usefulness of the MTCYB gene as a means of individual identification, as an auxiliary tool to the control region and not as a main means of identification, seems to be considerable. Out of 22 different control region lineages with 79 samples which shared the same D-loop sequence with some others within a lineage, 10 lineages with 37 samples could be subgrouped according to different MTCYB sequences. This suggests that if we study additional variations outside of the control region and MTCYB gene, the limitations of mtDNA in terms of discriminating common control region sequences will be reduced. The control region sequences for all the sample screened in this study are listed in Table 2.

Sequencing for a wide range or even the whole mtDNA genome may be necessary for complete discrimination. In the near future specific informative sites can be typed with routine sequencing, or by performing several assays for the single nucleotide polymorphisms [13]. In addition newly developed microchip-based approaches to sequencing have also been devised and these should enable mtDNA sequencing to be performed more easily in the near future [14, 15]. To achieve this more information about the pattern of variation of different mtDNA genes is necessary and this article provides valuable data about the human MTCYB gene polymorphism.

Acknowledgement This paper was supported by the Seoul National University Hospital Research Fund.

References

1. Anderson S, Bankier AT, Barrell GB, Bruijin MHL de, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG (1981) Sequence and organization of the human mitochondrial genome. Nature 290:457–465
2. Case JT, Wallace DC (1981) Maternal inheritance of mitochondrial DNA polymorphisms in cultured human fibroblasts somatic cell. Genetics 7:103–108
3. Corach D, Sala A, Penacino G, Iannucci N, Bernardi P, Doretti M, Fondebrider L, Ginarte A, Inchaurregui A, Somigliana C, Turner S, Hagelberg E (1997) Additional approach to DNA typing of skeletal remains: the search for "missing" persons killed during the last dictorship in Argentina. Electrophoresis 18:1608–1612
4. Gill P, Ivanov PL, Kimpton C, Piercy R, Benson N, Tully G, Evett I, Hagelberg E, Sullivan K (1994) Identification of the remains of the Romanov family by DNA analysis. Nat Genet 6:130–135
5. Anslinger K, Weichhold G, Keil W, Bayer B, Eisenmenger W (2000) Identification of the skeletal remains of Martin Bormann by mtDNA analysis. Int J Legal Med 114:194–196
6. Holland MM, Parsons TJ (1999) Mitochondrial DNA sequence analysis – validation and use for forensic casework. Forensic Sci Rev 11:22–50
7. Esposti MD, Vries SD, Crimi M, Ghelli A, Patarnello T, Meyer A (1992) Mitochondrial cytochrome b: evolution and structure of the protein. Biochim Biophys Acta 1143:243–271
8. Bataille M, Crainic K, Leterreux M, Durigon M, Mazancourt P de (1999) Multiplex amplification of mitochondrial DNA for human and species identification in forensic evaluation. Forensic Sci Int 99:165–170
9. Forrest ARR, Carnegie PR (1994) Identification of gourmet meat using FINS (forensically informative nucleotide sequencing). Biotechniques 17:24–26

10. Parson W, Pegoraro K, Niederstatter H, Foger M, Steinlechner M (2000) Species identification by means of the cytochrome b gene. Int J Legal Med 114:23–28
11. Sambrook J, Fritsch EF, Maniatis T (1987) Analysis and cloning of eukaryotic genomic DNA. Molecular cloning, vol 2. Cold Spring Harbor Laboratory Press, New York, pp 9.1–9.62
12. Andrews RM, Kubacka I, Chinnery P, Lightowlers RN, Turnbull DM, Howell N (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat Med 23:147
13. Lee LG, Connell CR, Bloch W (1993) Allelic discrimination by nick-translation PCR with fluorogenic probes. Nucleic Acids Res 21:3761–3766
14. Kozal MJ, Shah N, Shen N, Yang R, Fucini R, Merigan TC, Richman DD, Morris D, Hubbell E, Chee M, Gingeras TR (1996) Extensive polymorphisms observed in HIV-1 clade B protease gene using high-density oligonucleotide arrays. Nat Med 2:753–759
15. Edman C, Raymond D, Wu D, Tu E, Sosnowski R, Butler W, Nerenberg M, Heller M (1998) Electric field directed nucleic acid hybridization on microchips. Nucleic Acids Res 25:4907–4914