Mitochondrial DNA Polymorphism of the Coding Region in Thai Population
ความหลากหลายของลำดับเบสดีเอ็นเอของไมโตคอนเดรีย บริเวณที่เป็นรหัสในการสร้างโปรตีนในประชากรไทย

Arida Klomkliang (อาริดา กล่อมเกลี้ยง) [1]*
Dr. Budsaba Rerkamuaychoke (ดร.บุษบา ฤกษ์อำนวยโชค) **

ABSTRACT

Genetic variations of the mitochondria DNA (mtDNA) control region are a molecular target for degraded sample analysis in forensic casework investigation, however the low power of discrimination is the major limitation of the analysis. The coding region is a much larger subunit than the control region and may provide the possibility to cover rare type of mtDNA. In this study, we analyzed the polymorphism of MTATP6, MTCO3, MTND4L, MTND4, MTND5, MTND6 and MTCYB genes in 28 DNA samples who shared common HVI/HVII types from 200 unrelated Thai individuals. The results showed that a randomly matched probability was decreased, while discrimination power and haplotype diversity was increased, when using the control region together with these seven coding genes. Randomly matched probability was decreased from 0.62% to 0.52% and haplotype diversity was increased from 0.99881 to 0.99984. These result suggested that MTND4 has the most potential for increasing the power of the forensic testing.

บทคัดย่อ

ความผันแปรของลำดับเบสดีเอ็นเอของบริเวณที่ไม่ได้เป็นรหัสในการสร้างโปรตีนในไมโตคอนเดรีย เป็นเป้าหมายระดับโมเลกุล ของการวิเคราะห์ตัวอย่างที่เสื่อมสภาพในการตรวจพิสูจน์ทางนิติเวช อย่างไรก็ดี การมีอำนาจในการแยกแยะต่ำเป็นข้อจำกัดของการตรวจวิเคราะห์ ส่วนบริเวณที่เป็นรหัสในการสร้างโปรตีนนั้น มีขนาดใหญ่กว่าในส่วนแรกมาก จึงมีความเป็นไปได้ที่จะพบลำดับเบสที่สามารถเพิ่มอำนาจในการแยกแยะได้มากขึ้น ในการศึกษาครั้งนี้ได้ตรวจวิเคราะห์ ความหลากหลายบริเวณยีน MTATP6, MTCO3, MTND4L, MTND4, MTND5, MTND6 และ MTCYB ในตัวอย่าง 28 ตัวอย่างที่มีลำดับเบสบน HVI/HVII เหมือนกัน จากการคัดกรองคนไทย 200 คนที่ไม่มีความสัมพันธ์ทางเครือญาติกัน ผลการวิเคราะห์พบว่า

[1] *Correspondent author :aridant@hotmail.com*

* *Student, Master of Science Program in Clinical Pathology, Faculty of Medicine Ramathibodi Hospital, Mahidol University.*

** *Assoc. Prof. Department of Pathology, Faculty of Medicine Ramathibodi Hospital, Mahidol University.*

ค่าโอกาสในการเข้ากันได้แบบสุ่มมีค่าลดลง ในขณะที่อำนาจในการแยกแยะ และค่า haplotype diversity มีค่าเพิ่มขึ้นมากกว่าการใช้เพียงบริเวณ HVI/HVII ค่าโอกาสในการเข้ากันได้แบบสุ่มมีค่าลดลงจาก 0.62% เป็น 0.52 % และค่า haplotype diversity เพิ่มขึ้นจาก 0.99881 เป็น 0.99984 โดยการศึกษาครั้งนี้พบว่าบริเวณยีน MTND4 สามารถนำไปประยุกต์ใช้ในงานด้านนิติวิทยาศาสตร์ได้

Key Words : Mitochondrial DNA, Coding region, Thai population
คำสำคัญ : ไมโตคอนเดรียดีเอ็นเอ บริเวณที่เป็นรหัสในการสร้างโปรตีน ประชากรไทย

Introduction

Testing of mitochondrial DNA (mtDNA) sequence variation has become well established to assist in forensic identification casework. mtDNA typing is performed mostly under two circumstances: when degradation of DNA or sample type (such as hair shafts, teeth or skin particle) makes recovery of nuclear DNA difficult or impossible and when comparison to matrilineal relatives is desired. Mainly features of mtDNA that underlie its application are the following (1) it is present in high copy number per cell [1] (2) it is maternally inherited [2], (3) it has been shown not to recombine in humans [3] and (4) it has been shown to have a mutation rate up to ten times the mutation rate of single copy nuclear DNA due to the lower efficiency of mtDNA polymerase and the apparent lack of mtDNA repair mechanisms [3].

Human mtDNA is an extrachromosomal genome that is separated and distinct from the nuclear DNA composing of 16,569 base pair double-stranded .There are two regions including coding and control regions. The coding region encodes 13 subunits of the mitochondrial oxidative phosphorylation system, 2 ribosomal RNAs used for protein synthesis and 22 transfer RNAs [4]. The control or non-coding region, also known as the displacement loop (D-loop), contains sequence analysis of hypervariable region I (HVI) and hypervariable region II (HVII)

Although, analysis of human mtDNA is a useful tool for forensic identity testing, its correctable weaknesses have low power of discrimination obtained when common HVI/HVII are encountered in a casework. Therefore, the analysis of additional information in the coding region seems to be a good strategy to pass this limitation. Because of fifteen times larger than the control region, the coding region is potentially a rich source of additional sequence variation for the discrimination of mtDNA test.

In this study, we investigated mtDNA coding region encompassing nucleotide positions 8527-9207, 9207-9990, 10470-10766, 10766-12137, 12337-14148, 14149-14673 and 14747-15887 in 28 unrelated individuals who matched 1 of 11 common HVI/HVII types from Thai population.

Materials and Methods

DNA sample

The twenty-eight common HVI/HVII DNA samples of unrelated Thai individuals were chosen from 200 DNA samples which were from healthy individuals who subjected to DNA typing assay. These samples were the remaining materials from the routine diagnosis in Human Genetics Unit, Department of Pathology, Ramathibodi Hospital.

PCR condition

The PCR primer used to amplify MTATP6 gene, MTCO3 gene, MTND4L gene, MTND4 gene, MTND5 gene, MTND6 gene and MTCYB are listed in table 1 and 2. Each 50 μl volume reaction contained 1.0 U of Taq DNA polymerase (AppliedBiosystem, USA), 0.2 μM each of primer, 200 μM each dNTP, 10X buffer (100 mM Tris-HCl, pH 8.3, 500mM KCl ; AppliedBiosystem, USA), 1.5 mM $MgCl_2$ (AppliedBiosystem, USA) and amplification was conducted in Thermo cycler 9700 (Perkin Elmer Cetus, USA)

Thermal cycling conditions of reactions using primer in table 1 were 94°C for 10 minutes, followed by 35 cycles each of denaturation at 94°C for 30 seconds, annealing at 57°C for 30 seconds (for primers in table 2, annealing temperature was 55°C), extension at 72°C for 30 seconds and final extension at 72°C for 10 minutes. Then, 2% agarose gel electrophoresis was used to separate amplified products.

Cycle sequencing condition

The PCR products were purified with High pure PCR product purification kit (Roche Diagnostic, Germany) and subsequently to be sequenced using ABI Prism Bigdye Terminator version 1.1 cycle sequencing kit (AppliedBiosystem, USA) and primers (Table 1-2). Cycle sequencing was started at 96°C for 1 minute, followed by 25 cycles of 96°C for 15 seconds, 50°C for 5 seconds and 60°C for 4 minutes. Electrophoresis was performed using Applied Biosystems DNA Sequencer Model ABI Prism 3130. DNA sequences were aligned and compared with the revised Cambridge reference sequence (rCRS) by SeqScape software version 2.5.

Statistic calculation

The probability of two randomly selected individuals from a population having identical mtDNA type was calculated from $P = \Sigma Xi^2$. The power of discrimination and genetic diversity were calculated from $Pd = 1 - \Sigma Xi^2$ and $h = (1-\Sigma x^2)\ n\ /\ (n-1)$, respectively [5].

Results

Each DNA sample was amplified using specific primer for 7 genes, MTATP6, MTCO3, MTND4L, MTND4, MTND5, MTND6 and MTCYB genes in mtDNA coding region and the PCR products were observed in agarose gel (Figure 1). After sequencing were compared with the revised Cambridge reference sequence, and the result showed variation of 70 SNPs from 28 DNA samples

investigated (Table 3).

For statistic calculation of random match probability, discrimination power and haplotype diversity, the random match probability was decrease. While discrimination power and haplotypes diversity were increase when using the control region together with these seven coding genes. Randomly matched probability decreased from 0.62% to 0.52%. Discrimination power and haplotypes diversity increase from 0.99374 to 0.99477 and 0.99881 to 0.99984, respectively. (Table 4)

Discussion and Conclusion

Analysis of mtDNA is a valuable tool in casework investigation. Previously we performed mtDNA sequencing in 200 samples of unrelated Thai individuals and the result showed that 28 samples had the same sequences of HVI and HVII regions. Therefore, in this study, we analyzed the polymorphism of 7 genes for MTATP6, MTCO3, MTND4L, MTND4, MTND5, MTND6 and MTCYB genes in these 28 samples which shared common HVI/HVII types. The results showed a total of 70 nucleotide positions that varied from Anderson sequence (Table 3). In MTATP6 gene, the nucleotide variation of A8860G is found in all sequences. The confirmation by [6] who reanalyzed the original CRS and demonstrated that A at position 8860 was a rare polymorphism. For ATPase6 protein, the amino acid sequence at position 112 changed from theronine to alanine. The most common SPNs besides A8806G was at the position A8701G (42.86%) which was the same as Taiwan population [7]. While German population mostly had position G8697A [8]. The most common variant in MTND4 gene is G11719A (100%) and this variant did not affect the protein sequence as determined by [9]. In addition for MTCYB gene, C14766T (100%) was the most frequent SNPs but this genetic variation resulted in amino acid change from isoleucine to threronine. The second common nucleotide variation from rCRS was A15326G (92.86%) which occurred in the Cytb protein for changing the amino acid sequence at position 194 from threonine to alanine. On the other hand, in Taiwan population the most common and the second common SNPs were A15326G and C14766T, respectively [10].

From our previous data, the genetic diversity of the HVI and HVII was 0.99881. The random match probability and discrimination power were 0.62 % and 0.99374, respectively. In this study, some genes in coding region have been shown to increase the genetic diversity and power of discrimination when combined with HVI and HVII. Combination of HVI, HVII and MTND4 showed the most increased value of genetic diversity and power of discrimination. Moreover, the genetic diversity and the power discrimination also increased when combined HVI, HVII with MTATP6 or MTND5 or MTCYB. In addition, the random match probability was decreased to 0.54% when combined HVI, HVII with MTND4 and 0.59% when combined HVI, HVII with

MTATP6 or MTND5 or MTCYB. On the other hand, the genetic diversity, the random match probability and the discrimination power were not different when combined HVI and HVII with MTND4L or MTND6 or MTCO3.

Combining all of these 7 coding genes, MTND4 has the most potential for increasing the power of the forensic testing which sometime had no data or unsuccessful results of autosomal and Y STR. Moreover the forensic casework usually has limitation in nuclear DNA analysis because of sample degradation and small amounts of target DNA. Therefore, mtDNA usually could be the only source to provide useful information of the case.

Acknowledgements

I would like to thank all staffs at Human Genetics Unit, Faculty of Medicine Ramathibodi hospital, Mahidol University for their invaluable help. This research grant was supported by Mahidol University.

Reference

1. Satoh M, Kuroiwa T. Organization of multiple nucleoids and DNA molecules in mitochondria of a human cell. Exp Cell Res. 196: 1991. 137-140.
2. Case JT, Wallace DC. Maternal inheritance of mitochondrial DNA polymorphisms in cultured human fibroblasts. Somatic Cell Genet. 7: 1981. 103-108.
3. Butler JM. Forensic DNA Typing: Biology, Technology, and Genetics of STR Markers. Second Edition. London: Academic Press. 2005.
4. Anderson S, Bankier AT, Barrell BG, de Bruijn, MH, Coulson AR, Drouin J, et al. Sequence and organization of the human mitochondrial genome. Nature. 290: 1981. 457- 465.
5. Tajima F. Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. Genetics. 123: 1989. 585-595.
6. Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat Genet. 23:147. 1999.
7. Tzen CY, Wu TY, Liu HF. Sequence polymorphism in the coding region of mitochondrial genome encompassing position 8389-8865. Forensic Sci Int.120: 2001. 204 -209.
8. Poetsch M, Wittig H, Krause D, Lignitz E. The impact of mtDNA analysis between position nt 8306 and nt 9021 for forensic casework. Mitochondrion. 3: 2003. 133-137.
9. Lutz-Bonengel S, Schmitt T, Pollak S. Sequence polymorphisms within the human mitochondrial genes MTATP6, MTATP8, MTND4. Int J Legal Med. 117: 2003. 133-142.
10. Hwa HL, Ko, TM, Chen YC, Chang YY, Tseng LH, Su YN, Lee JC. Study of the cytochrome b gene sequence in populations of Taiwan. J Forensic Sci. 55: 2010. 167-170.

Table 1 PCR Primers for 6 genes in coding region.

PCR primer position	Sequence
MT-ATP6_ F1 (8451-8472)	5' TAAACACAAACTACCACCTACC 3'
MT-ATP6_ R1 (8788-8809)	5' TTGGTGTAAATGAGTGAGGCAG 3'
MT-ATP6_F2 (8703-8724)	5' CATACACAACACTAAAGGACGA 3'
MT-ATP6_R2 (9073-9091)	5' GTAGAGGGAAGGTTAATGGT 3'
MT-CO3_F1 (9057-9076)	5' CACCCTAGCAATATCAACCA 3'
MT-CO3_R1 (9610-9630)	5' CTCCTGATGCGAGTAATACGG 3'
MT-CO3_F2 (9596-9613)	5' ACTCCTAAACACATCCGT 3'
MT-CO3_R2 (10090-10110)	5' TTAGTAGTAAGGCTAGGAGGG 3'
MT-ND4L_F (10356-10375)	5' CTAAGTCTGGCCTATGAGTG 3'
MT-ND4L_R (10795-10810)	5' TTGGAAAGTCATGTCAGTGGT 3'
MT-ND4_F2 (11144-11162)	5' ACCTTGGCTATCATCACCC 3'
MT-ND4_R2 (11696-11715)	5' TGGGCGATTATGAGAATGAC 3'
MT-ND4_F3 (11655-11674)	5' CAGCCATTCTCATCCAAACC 3'
MT-ND4_R3 (12176-12194)	5' GTCGTAAGCCTCTGTTGTC 3'
MT-ND5_F1 (12284-12304)	5' CTATCCATTGGTCTTAGGCCC 3'
MT-ND5_R1 (12788-12808)	5' ATCAACTGATGAGCAAGAAGG 3'
MT-ND5_F2 (12726-127406)	5' CTTAGTTACCGCTAACAACCT 3'
MT-ND5_R2 (13238-13256)	5' TGAAGTGGAGAAGGCTACGA 3'
MT-ND5_F3 (13155-13174)	5' CCAAACTCTAACACTATGCT 3'
MT-ND5_R3 (13724-13743)	5' AGTAATGAGAAATCCTGCGA 3'
MT-ND5_F4 (13649-13670)	5' CCACCCTTACTAACATTAACGA 3'
MT-ND5_R4 (14043-14063)	5' ATGATGGAGGTGGAGATTTGG 3'
MT-ND6_F1 (13984-14004)	5' CTACTCCTCCTAGACCTAACC 3'
MT-ND6_R1 (14360-14379)	5' GAGGTAGGATTGGTGCTGTG 3'
MT-ND6_F2 (14346-14366)	5' CATACTCTTTCACCCACAGCA 3'
MT-ND6_R2 (14819-14836)	5' TCATGCGGAGATGTTGGA 3'
MT-CYB_ F3 (15581-15594)	5' ACAATTCTCCGATCCGTCC 3'
MT-CYB _R3(15910-15928)	5' CCGGTTTACAAGACTGGTG 3'

Table 2 PCR Primers for 2 genes in coding region.

PCR primer position	Sequence
MT- ND4_F1(10655-10673)	5' ATACTAGTCTTTGCCGCC 3'
MT- ND4_R1(11197-11218)	5' GGTGTAGAATAGGAAGTATGTG 3'
MT-CYB _F1(14736-14753)	5' CAAGAACACCAATGACCC 3'
MT-CYB _R1(15250-15267)	5' GTGGGACTGTCTACTGAG 3'
MT-CYB _F2(15183-15202)	5' TTACAAACTTACTATCCGCC 3'
MT-CYB _R2(15639-15659)	5' GGATTATTGCTAGGATGAGGA 3'

Table 3 Nucleotide variation of 7 genes in mtDNA coding region.

rCRS		Thai	Number	Prevalence %	Location
8581	G	A	1	3.57	MTATP6
8701	A	G	12	42.86	MTATP6
8718	A	G	2	7.14	MTATP6
8853	A	G	2	7.14	MTATP6
8860	A	G	28	100.00	MTATP6
8961	A	G	1	3.57	MTATP6
8989	G	A	1	3.57	MTATP6
9053	G	A	11	39.29	MTATP6
9127	A	G	2	7.14	MTATP6
9274	A	G	2	7.14	MTCO3
9449	C	T	4	14.29	MTCO3
9512	C	T	2	7.14	MTCO3
9540	T	C	12	42.86	MTCO3
9548	G	A	9	32.14	MTCO3
9554	G	A	2	7.14	MTCO3
9644	A	G	2	7.14	MTCO3
9824	T	C	4	14.29	MTCO3
10609	T	C	11	39.29	MTND4L
10679	A	G	2	7.14	MTND4L
10873	T	C	12	42.86	MTND4
10897	C	T	2	7.14	MTND4
11065	A	G	2	7.14	MTND4
11084	A	G	2	7.14	MTND4
11215	C	T	3	10.71	MTND4
11253	T	C	1	3.57	MTND4
11350	A	G	1	3.57	MTND4
11482	T	C	2	7.14	MTND4
11665	C	T	2	7.14	MTND4
11719	G	A	28	100.00	MTND4
11914	G	A	1	3.57	MTND4
11923	A	G	1	3.57	MTND4
11932	C	T	2	7.14	MTND4
12091	T	C	2	7.14	MTND4
12354	T	C	2	7.14	MTND5
12405	C	T	2	7.14	MTND5
12406	G	A	11	39.29	MTND5
12621	C	T	2	7.14	MTND5
12705	G	T	12	42.86	MTND5
12771	G	A	2	7.14	MTND5

Table 3 Nucleotide variation of 7 genes in mtDNA coding region. (Cont.)

rCRS		Thai	Number	Prevalence %	Location
12811	T	C	2	7.14	MTND5
12882	C	T	11	39.29	MTND5
12950	A	G	2	7.14	MTND5
13353	A	G	2	7.14	MTND5
13401	T	C	2	7.14	MTND5
13590	G	A	1	3.57	MTND5
13759	G	A	11	39.29	MTND5
13806	C	T	2	7.14	MTND5
13928	G	C	13	46.43	MTND5
13929	C	T	1	3.57	MTND5
13965	T	C	1	3.57	MTND5
14053	A	G	2	7.14	MTND5
14088	T	C	4	14.29	MTND5
14110	T	C	2	7.14	MTND5
14209	A	G	3	10.71	MTND6
14766	C	T	28	100.00	MTCYB
14783	T	C	13	46.43	MTCYB
14971	T	C	3	10.71	MTCYB
14974	C	T	2	7.14	MTCYB
14979	T	C	1	3.57	MTCYB
15043	G	A	11	39.29	MTCYB
15236	A	G	2	7.14	MTCYB
15257	G	A	2	7.14	MTCYB
15301	G	A	12	42.86	MTCYB
15326	A	G	26	92.86	MTCYB
15338	C	T	2	7.14	MTCYB
15346	G	A	3	10.71	MTCYB
15773	G	A	1	3.57	MTCYB
15412	T	G	2	7.14	MTCYB
15458	T	C	2	7.14	MTCYB
15691	A	G	2	7.14	MTCYB

Table 4 Statistic parameter calculation from HVI/HVII only and HVI/HVII/coding gene(s).

mtDNA region	Haplotype diversity	Matching probability (%)	Power of discrimination
HVI + HVII	0.99881	0.62	0.99374
HVI + HVII +MTATP6	0.99917	0.59	0.99410
HVI + HVII +MTCO3	0.99881	0.62	0.99374
HVI + HVII +MTNDL4	0.99881	0.62	0.99374
HVI + HVII +MTND4	0.99933	0.54	0.99423
HVI + HVII +MTND5	0.99917	0.59	0.99410
HVI + HVII +MTND6	0.99881	0.62	0.99374
HVI + HVII +MTCYB	0.99917	0.59	0.99410
HVI + HVII +7 genes	0.99984	0.52	0.99477

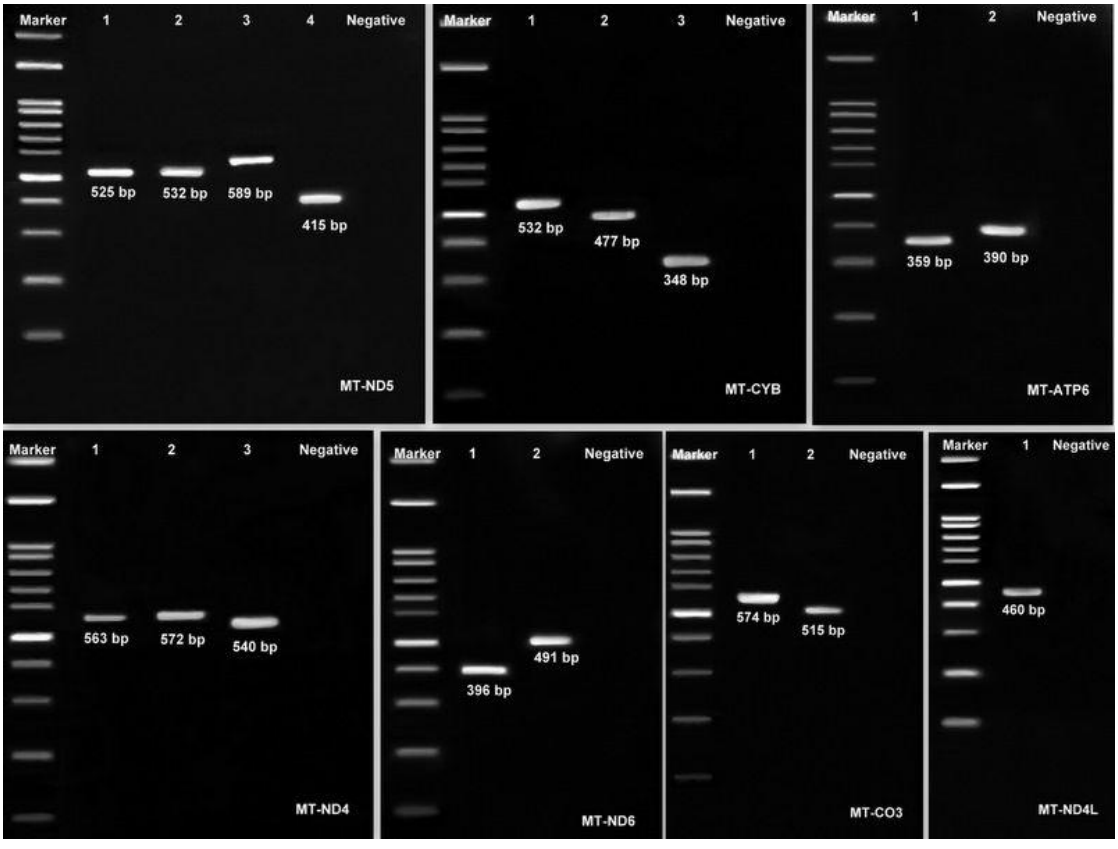

Figure 1 Agarose gel electrophoresis of PCR products amplified from one DNA sample using primers specific for 7 genes. The sizes of PCR products were shown under each band. Marker = DNA ladder marker, Land 1-4 = PCR products from the first to fourth primer pairs respectively, Negative = negative control.