Arch Pharm Res Vol 29, No 1, 88-95, 2006

Archives of Pharmacal Research
http://apr.psk.or.kr

Phylogenetic Analysis by RFLP and Sequencing of Mitochondrial DNA in a Korean Population

Jin-Young Lee, Heui-Soo Kim[1], Bae Jin Ha[2], and Yeong-Hong Park

Department of Biochemistry, College of Medicine, Inje University, Busan 614-735, Korea, [1]Department of Biology, College of Natural Science, Busan National University, Busan 609-735, Korea, and [2]Department of Bioscience and Biotechnology, College of Environmental Engineering and Biotechnology, Silla University, Busan 617-736, Korea

(Received August 5, 2005)

Analysis of molecular nature of mitochondrial DNA (mtDNA) could be powerful marker for anthropological studies of modern populations. While population genetic studies on mtDNA have been reported for several ethnic groups, no such study has been documented for the Korean population. We surveyed mtDNA polymorphisms in the HVS I of noncoding D-loop region and its upstream region from 430 unrelated healthy Korean population by polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) and direct sequencing analysis. PCR product with 2,790 bp spanning the specific mtDNA region (mt13715-16504) was subjected to RFLP analysis using 6 restriction enzyme (*Hinf* I, *Hae* III, *Alu* I, *Dde* I, *Mbo* I, *Rsa* I). On the PAUP analysis of PCR-RFLP results, 38 mtDNA haplotypes (Hap 1-38) were detected in the Korean populations, which were classified into 11 haplogroups (Grp 1-11) of related haplotypes encompassing all 38 haplotypes. In comparison of sequencing data with Anderson's reference sequence, the transition type was more prevalent than the transversion type. Insertions or deletions were not found. In addition, three of the polymorphic sites (A16240C, A16351G, G16384A) in HVS-I region are determined newly. The polymorphic sites were distributed randomly in the region, though the frequency at each site was variable. Thus, this research might be required for the genealogical study of Orientals.

Key words: mtDNA, Korean, Noncoding D-loop, PCR-RFLP, Haplotype

INTRODUCTION

In the last decade, the research on mitochondrial biology has demonstrated that the central role played by the mitochondrion in crucial cellular processes, and the importance of the mitochondrial genome (mtDNA) in these processes is progressively emerging. Human mtDNA has been completely sequenced as early as 1981 (Anderson *et al.*, 1981). The mtDNA is a small circular genome located within the mitochondria in the cytoplasm of the cell. Each human cell contains hundreds of mitochondria and thousands of mtDNAs. Thus, mitochondrial genetics is a population genetics, both at the level of the intracellular colony of symbionts and the human population. Thirty seven genes are known to be encoded by mtDNA. Twenty-four encode for the translational machinery of the mtDNA itself (22 tRNAs and 2 rRNAs). The additional 13 genes encode for subunits of the oxidative phosphorylation (OXPHOS) system (ND 1-6 and ND 4L for complex I (NADH:ubiquinone oxidoreductase); cytochrome b for complex III (ubiquinol;ferricytochrome c oxidoreductase); CO I-III for complex IV (ferrocytochrome c:oxygen oxidoreductase); ATPase 6 and 8 for complex V (H^+-translocating ATP synthase)). The noncoding stretch of the mtDNA sequence is the so called displacement (D)-loop or hypervariable segment (HVS), that is supposed to have a role in the regulation of replication and transcription. Study of the mitochondrial restriction-site polymorphism is a method different from allele-frequency based surveys of nuclear DNA marker loci for several reasons. First, because the mitochondrial genome is maternally inherited and because there is no genetic recombination process in mitochondria, polymorphisms in the mitochondrial genome are tightly linked to each other (Giles *et al.*, 1980). Second, the mtDNA is a small circular genome located within mitochondria in the cytoplasm of the cell. Therefore, the nucleotide sequencing is very simple. Third, compari-

Correspondence to: Yeong-Hong Park, Department of Biochemistry, College of Medicine, Inje University, 633-165, Gaegum-dong, Busanjin-gu, Busan 614-735, Korea
Tel: 82-51-890-6747, Fax: 82-51-894-5801
E-mail: pyh6747@ijnc.inje.ac.kr

son of DNA sequences in different organisms reveals that the rate of nucleotide substitution during evolution has been 10 times greater in the mitochondrial genome than in the nuclear genome, which may be due to reduced fidelity of mitochondrial DNA replication, DNA repair, or both (Wilson *et al.*, 1985). Fourth, it is also characterized by a compact structure where there are very few intergenic spacing sequences, no introns and overlapping genes. Last, mtDNA is haploid, contrary to nuclear DNA which exists as diploid. Collectively, mtDNA includes high copy number per cell, lack of recombination, high substitution rate, and uniparental inheritance, that have made mtDNA an attractive source of information for phylogenetic studies (Bhat *et al.*, 1990). The control region of mtDNA, which includes the origin of H strand replication, the D-loop and both origin of transcription, is the most variable region of the human mtDNA genome (Cann *et al.*, 1987). Variation in the mtDNA sequence have been analyzed in human populations, both in terms of evolution and population dispersion and in term of the role that mtDNA mutations play in human disease (Torroni *et al.*, 1996; Ingman *et al.*, 2000; Howell, 1999; Wallace *et al.*, 1999). However, the generation of comprehensive and unambiguous phylogenetic data, especially for the mtDNA coding regions, is limited by the availability of a relatively small number of polymorphisms that have been identified on the basis of the presence or absence of restriction-enzyme recognition sites (Torroni *et al.*, 1996; Wallace *et al.*, 1999). Alternatively, the sequence of D-loop has been used to establish phylogenetic networks (Richards *et al.*, 1996; Richards *et al.*, 1998), and a combination of both methods was recently used for a comparative analysis (Macaulay *et al.*, 1999). Reliance on hypervariable control-region sequences, however, is not without controversy, particularly because of the effects that homoplasy and saturation have at sites with high mutation rates (Howell *et al.*, 1996). The variation of mtDNA in human population has been initially studied by means of RFLP by using five "rare cutter" restriction enzymes (Johnson *et al.*, 1983). Afterward, when the PCR technique allowed to better analyze the fragments of restriction digestions, the analyses were carried out with 14 enzymes, some of which were "fragments cutter" (Cann *et al.*, 1997; Ballinger *et al.*, 1992). The sequence analysis of the D-loop was soon proposed (Richards *et al.*, 1998; Vigilant *et al.*, 1991; Graven *et al.*, 1995). In very recent years, the progress in sequencing analyses has brought to the complete sequencing of the mtDNA molecules for searching the variation (Ingman *et al.*, 2000; Rose *et al.*, 2001; Finnila *et al.*, 2001). Consequently, now are known a number of mtDNA lineage, defined by many mutations, that characterize the human populations. The phylogenetic analysis of these lineages has allowed to discover that they can be clustered in groups (called haplogroups) where all the lineages share some of the mutations. Nowadays, the millenarian history of these human groups is reflected in the fact that mtDNA haplogroups are continent specific (Torroni *et al.*, 1996). The sequence differences between mtDNAs can then be compared by using various phylogenetic procedures including parsimony, neighbor-joining, and unweighted pair-group analysis. In contrast to West Eurasia (Macaulay *et al.*, 1999), the Americas (Torroni *et al.*, 1993; Brown *et al.*, 1998), and Siberia (Torroni *et al.*, 1993; Schurr *et al.*, 1999), the mitochondria of East Asia have not yet been classified satisfactory. A few Korean control region sequences have been published, but it is difficult to assign them into haplogroups, which are defined by RFLP markers. In the present study, we have therefore used mtDNA direct sequencing and RFLPs of upstream region including HVS-I from Korean population. Sequence data were further used to compare Korean mtDNAs with mtDNAs from other Asian populations.

MATERIALS AND METHODS

Sample collection

A total of 430 blood samples were obtained from Inje University Busan Paik Hospital. The sampling was random and only subjects unrelated in maternal lineage were selected.

DNA extraction from blood samples

The blood was centrifuged at 3,000 rpm for 10 min in 4 °C. The obtained white blood cell (WBC) and serum were used for DNA extraction. DNA was extracted from 5 mL of WBC and 195 mL of serum by QIAamp DNA Mini Kit.

PCR amplification of mtDNA spanning nucleotides 13715-16504

The selected region of the mtDNA includes the entire cyt b and ND 6 genes, a small section of the ND 5 gene (240 bp) and almost half of the hypervariable D-loop, which was amplified by two sets of primers of 21 bp. The fragment was 2790 bp. The name of primers was designated to carry a base number (Anderson *et al.*, 1981) of first nucleotide with a letter indicating the direction of the primer, F (forward) or R (reverse) (Table I). PCR amplification was performed in 20 μL reaction volumes containing 2 μL of 10×buffer (Tris-HCl (pH 8.0), 100 mM KCl, 0.5% (w/v) Tween 20), 2.5 mM of each dNTP, 5 pmole of each primer, 5 units of *Taq* polymerase (TaKaRa R001A), and 2 μL of template DNA. An initial denaturation step at 94°C for 5 min was followed by 30 cycles comprising a denaturation at 94°C for 1 min, annealing at 55°C for 1 min, and extension at 70°C for 2 min. The final extension was subsequently performed at 70°C for 10 min, and then

Table I. The types, positions, sequences of Primers used for PCR and sequencing

Primer type	Position*	5'sequences3'
PCR amplification primer for 2790 mtDNA region		
First F1	(13527-13547)	AAC CGC AAA CAT ATC ATA CAC
First R1	(16500-16520)	TCT GGT TCC TAC TTC AGG GTC
Nested F2	(13715-13735)	GAA GCC TAT TCG CAG GAT TTC
Nested R2	(16484-16504)	TGA ACT GTA TCC GAC ATC TGG
Sequencing primers		
SEQ 1F	(13715-13735)	GAA GCC TAT TCG CAG GAT TTC
SEQ 2F	(14118-14138)	ACT CAT CCT AAC CCT ACT CCT
SEQ 3F	(14515-14535)	TAA ACC CAT ATA ACC TCC CCC
SEQ 4F	(14803-14823)	CGA CCT CCC CAC CCC ATC CAA
SEQ 5F	(15100-15120)	CCT CCT GCT TGC AAC TAT AGC
SEQ 6F	(15400-15420)	CAC CTT CCA CCC TTA CTA CAC
SEQ 7F	(15658-15678)	CCC CAT CCT CCA TAT ATC CAA
SEQ 8F	(15980-16000)	CCA TTA GCA CCC AAA GCT AAG

* Position defined according to the published mitochondrial genome sequence of Anderson *et al.* (1981).

at 4°C until removed from the PCR machine. Successful amplification was confirmed using gel electrophoresis on a 1% agarose gel stained with ethidium bromide (10 mg/mL). Lambda DNA/Hind III (G171A, Promega) was used as size marker. Amplified DNAs were visualized and photographed on UV illuminator.

mtDNA RFLP analysis

The polymorphisms were genotyped by PCR-RFLP (Polymerase Chain Reaction-Restriction Fragment Length Polymorphism). Following successful amplification of the mtDNA region (mt13715-16504) from all of the DNA samples, each of 2790 bp PCR products were independently subjected to RFLP analysis using 6 different restriction enzyme (*Hinf* I (R6201, Promega), *Hae* III (R6171, Promega), *Alu* I (R6281, Promega), *Dde* I (R6171, Promega), *Mbo* I (R6711, Promega), *Rsa* I (R6371, Promega)) as outlined previously (Ballinger *et al.*, 1992). The condition and buffers used were those recommended by the manufacturer's instructions. Restriction reactions were performed in a 20 μL volume composed of 4 μL of PCR product, 2 μL of 10×reaction buffer and 5 units of endonuclease for 2 h 30 min at 37°C. The digested samples were resolved by electrophoresis on 3% agarose gel (IB 70042, IBI) with ethidium bromide (10 mg/mL) in 0.5×TAE buffer. 100 bp DNA ladder (G2101, promega) was used as size marker. Visualization of fragments was carried out under UV light and developed on film using a polaroid camera. In all RFLP experiments, negative control (without restriction enzyme) was simultaneously examined to the accuracy of the experiment.

Sequencing analysis

Selected PCR products having different haplotypes were purified and sequenced. Purified PCR products were sequenced on a ABI PRISM (model 3700, Version 3.6) automated DNA sequencer, using the appropriate sequencing primers shown in Table I. The purification kit of PCR product was purchased from QIAGEN.

Data analysis

Distinct restriction morphs were identified by 0, 1 character. Each specimen was assigned multi letter code that described its composit mtDNA genotype (Haplotype). The raw data were fragment profiles, but We are inferred site differences among haplotypes from changes in fragment profiles as the gain (1) and loss (0) of particular restriction sites. The restriction pattern data was analyzed using the PAUP (Phylogenic Analysis Using Parsimony, Version 4.0) computer package. The restriction matrix was bootstrapped 500 times. For sequencing data, Bioedit sequence alignment edit (Version 5.0) program, MEGA program (Molecular Evolutionary Genetics Analysis, Version 2.1), and DNASIS program (Version 2.5) were performed to identification of nucleotide mutations.

RESULTS

Cleavage site variation

PCR-RFLP analysis was performed on the basis of amplification of Korean mtDNA segments using appropriate sets of primers. The amplified segments of the entire cyt b and ND6 genes, a small section of the ND5 gene (240 bp) and almost half of the D-loop has an approximate size of 2790 bp. The 6 restriction enzymes (*Hinf* I, *Hae* III, *Alu* I, *Dde* I, *Mbo* I, *Rsa* I) were used to screen the Korean population. All restriction enzymes were polymorphic for D-loop and cytochrome b, respectively. Six *Hinf* I morphs were seen in this study. Eighty six percent of the Korean population was Morph A. Morph B and morph C were found at following high frequencies, 6.7% and 2.3%, respectively. Morph D was the same as frequencies of morph C (Table II). Four morphs of *Hae* III, morph A, B, C, and D were found in the studied sample. Morph A is the most common (97.0%). The morph B accounts for 1.4% of individuals and morph C and D for 0.9%, 0.7% each (Table II). The high degree of *Alu* I polymorphism described for Korean was confirmed. Nine distinct *Alu* I patterns were observed. *Alu* I morph A was found at a higher frequency (94.2%) than other morphs. Morph B was found in 3.0%. It was also observed in morph D-F and morph G-I at 0.5% and 0.2%, respectively (Table II). A high degree of polymorphism has been found in the *Dde* I restriction sites. 94.2% of the studied sample had the predominant morph A. Morph B was observed in 2.6% and morph C, D, E, and G in 0.7%

each. Morph F and H was also found in 0.2% each (Table II). Seven *Mbo* I morphs were seen in this study. The common morph A accounts for 95.3% of individuals and morphs B for 2.3%. Morph C and G were observed in 0.7% and Morph E and F in 0.2% each (Table II). Four different morphs were observed for *Rsa* I. The *Rsa* I morph A is the most common, found by 96.5% of the subjects, followed by morph B found in 2.1%, morph C in 1.2%, and morph D in 0.2% (Table II).

Table II. Frequency of morphs observed in the Korean mtDNA digested with six kinds of restriction enzyme

Enzyme	Morphs	Male (n=247)	Female (n=183)	Total (n=430)
Hinf I	A	205 (83.0)	165 (90.2)	370 (86.0)
	B	23 (9.3)	6 (3.3)	29 (6.7)
	C	6 (2.5)	4 (2.2)	10 (2.3)
	D	5 (2.0)	5 (2.7)	10 (2.3)
	E	5 (2.0)	1 (0.6)	6 (1.4)
	F	3 (1.2)	2 (1.1)	5 (1.2)
Hae III	A	239 (96.8)	178 (97.3)	417 (97.0)
	B	3 (1.2)	3 (1.6)	6 (1.4)
	C	3 (1.2)	1 (0.6)	4 (0.9)
	D	2 (0.8)	1 (0.6)	3 (0.7)
Alu I	A	234 (94.7)	171 (93.4)	405 (94.2)
	B	7 (2.8)	6 (3.3)	13 (3.0)
	C	2 (0.8)	1 (0.6)	3 (0.7)
	D	2 (0.8)	- (0.0)	2 (0.5)
	E	1 (0.4)	1 (0.6)	2 (0.5)
	F	1 (0.4)	1 (0.6)	2 (0.5)
	G	- (0.0)	1 (0.6)	1 (0.2)
	H	- (0.0)	1 (0.6)	1 (0.2)
	I	- (0.0)	1 (0.6)	1 (0.2)
Dde I	A	234 (94.7)	171 (93.4)	405 (94.2)
	B	6 (2.4)	5 (2.7)	11 (2.6)
	C	2 (0.8)	1 (0.6)	3 (0.7)
	D	2 (0.8)	1 (0.6)	3 (0.7)
	E	1 (0.4)	2 (1.1)	3 (0.7)
	F	- (0.0)	1 (0.6)	1 (0.2)
	G	1 (0.4)	2 (1.1)	3 (0.7)
	H	1 (0.4)	- (0.0)	1 (0.2)
Mbo I	A	238 (96.4)	172 (94.0)	410 (95.3)
	B	5 (2.0)	5 (2.7)	10 (2.3)
	C	2 (0.8)	1 (0.6)	3 (0.7)
	D	1 (0.4)	1 (0.6)	2 (0.5)
	E	- (0.0)	1 (0.6)	1 (0.2)
	F	- (0.0)	1 (0.6)	1 (0.2)
	G	1 (0.4)	2 (1.1)	3 (0.7)
Rsa I	A	239 (96.8)	176 (96.2)	415 (96.5)
	B	7 (2.8)	2 (1.1)	9 (2.1)
	C	1 (0.4)	4 (2.2)	5 (1.2)
	D	- (0.0)	1 (0.6)	1 (0.2)

Phylogenetic analysis

Restriction enzyme analysis of the 2790 bp PCR amplified product (mt13715-16504) of the mtDNA derived from 430 Korean individuals used in this study revealed total 38 polymorphic restriction sites. A total of 38 different genetic haplotypes encompassing the entire 430 samples involved in this study were identified on the basis of the presence (1) or absence (0) of each of the restriction sites. Frequency differences found between the 38 haplotypes (Table III) failed to reach statistical significance when analyzed by the Fisher's exact test. To improve the statistical power of the analysis, phylogenetic relationships between the 38 individual haplotypes were analyzed using maximum parsimony criteria. A maximum parsimony tree was generated and 11 groups of related haplotypes (Grp 1-11) were formed from this data (Fig. 1). The frequencies of the haplotypes contained within each group were examined and statistical analysis was again performed. Males and females were also reanalyzed separately

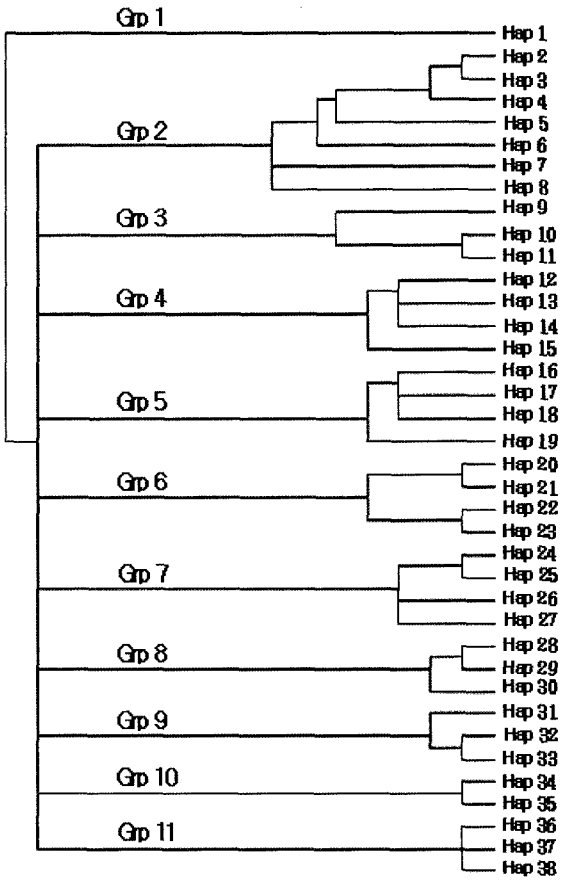

Fig. 1. The unrooted dendrogram showing the relationdhips between the haplotypes using maximum parsimony criteria. This was performed by using the cluster of programmes in the PAUP (Version 4.0).

Table III. The distinct haplogroups (Grp1-11) and haplotypes (hap 1-38) observed in a Korean population. Composite genotypes are denoted by capital letters in the following order: *Hinf* I, *Hae* III, *Alu* I, *Dde* I, *Mbo* I, *Rsa* I.

Haplo-group	Haplotype	Composite genotype	Male (n=247)	Female (n=183)	Total (n=430)
Group 1	Hap1	AAAAAA	165 (66.8)	129 (70.5)	294 (68.4)
Group 2	Hap2	FAADBA	2 (0.8)	1 (0.6)	3 (0.7)
Group 2	Hap3	FAAABA	1 (0.4)	- (0.0)	1 (0.2)
Group 2	Hap4	AAAABA	2 (0.8)	4 (2.2)	6 (1.4)
Group 2	Hap5	AAAAGA	1 (0.4)	2 (1.1)	3 (0.7)
Group 2	Hap6	FAAAFA	- (0.0)	1 (0.6)	1 (0.2)
Group 2	Hap7	AAAACA	2 (0.8)	1 (0.6)	3 (0.7)
Group 2	Hap8	AAAFDA	- (0.0)	1 (0.6)	1 (0.2)
Group 3	Hap9	ACAAAA	3 (1.2)	1 (0.6)	4 (0.9)
Group 3	Hap10	ABAAAA	3 (1.2)	2 (1.1)	5 (1.2)
Group 3	Hap11	ABAAEA	- (0.0)	1 (0.6)	1 (0.2)
Group 4	Hap12	AAEEAA	1 (0.4)	1 (0.6)	2 (0.5)
Group 4	Hap13	AAFAAA	1 (0.4)	- (0.0)	1 (0.2)
Group 4	Hap14	DAFAAA	- (0.0)	1 (0.6)	1 (0.2)
Group 4	Hap15	AAIAAA	- (0.0)	1 (0.6)	1 (0.2)
Group 5	Hap16	AABDAA	1 (0.4)	1 (0.6)	2 (0.5)
Group 5	Hap17	AABAAA	5 (2.0)	5 (2.7)	10 (2.3)
Group 5	Hap18	AAGAAA	- (0.0)	1 (0.6)	1 (0.2)
Group 5	Hap19	AADAAA	2 (0.8)	- (0.0)	2 (0.5)
Group 6	Hap20	CAAAAA	5 (2.0)	4 (2.2)	9 (2.1)
Group 6	Hap21	CAAADA	1 (0.4)	- (0.0)	1 (0.2)
Group 6	Hap22	EAAAAA	4 (1.6)	2 (1.1)	6 (1.4)
Group 6	Hap23	EDAAAA	1 (0.4)	- (0.0)	1 (0.2)
Group 7	Hap24	BDAAAA	- (0.0)	1 (0.6)	1 (0.2)
Group 7	Hap25	BDBAAA	1 (0.4)	- (0.0)	1 (0.2)
Group 7	Hap26	BAAAAA	23 (9.3)	3 (1.6)	26 (6.1)
Group 7	Hap27	BAAAAD	- (0.0)	1 (0.6)	1 (0.2)
Group 8	Hap28	AACAAA	2 (0.8)	1 (0.6)	3 (0.7)
Group 8	Hap29	AAHAAC	- (0.0)	1 (0.6)	1 (0.2)
Group 8	Hap30	AAAAAC	1 (0.4)	3 (1.6)	4 (0.9)
Group 9	Hap31	AAAAAB	6 (2.4)	2 (1.1)	8 (1.9)
Group 9	Hap32	DAAAAB	1 (0.4)	- (0.0)	1 (0.2)
Group 9	Hap33	DAAAAA	4 (1.6)	4 (2.2)	8 (1.9)
Group 10	Hap34	AAAHAA	1 (0.4)	- (0.0)	1 (0.2)
Group 10	Hap35	AAAEAA	- (0.0)	1 (0.6)	1 (0.2)
Group 11	Hap36	AAABAA	6 (2.4)	5 (2.7)	11 (2.6)
Group 11	Hap37	AAADAA	1 (0.4)	- (0.0)	1 (0.2)
Group 11	Hap38	AAAGAA	1 (0.4)	2 (1.1)	3 (7.0)

(Table III). It is of interest to note that Grp 2, containing six different haplotypes (hap 2, hap 3, hap 4, hap 5, hap 6, hap 7) is observed at a 4.2% frequency, and Grp 7 (hap 24, hap 25, hap 26, hap 27) is examined at a reduced frequency a 17.2% in the female samples compared to the male samples (82.8%), while inversely Grp 2 is increased in the female samples (55.6%) compared to the males (44.4%).

Nucleotide diversity and genetic similarity

We investigated Korean specific nucleotide diversity in 16 Koreans, which are divided into each haplogroup by PCR-RFLP pattern. Near the primer, about 30-40 nucleotide sequences were somewhat ambiguous and sequences from 13754 to 16490 in target region (the numbers designate nucleotide positions cited in Anderson's sequence (Anderson *et al.*, 1981)). Compared to Anderson's data, 59 sites of different nucleotide sites in the selected region were noted. Among the 59 polymorphic nucleotide sites, 27 (45.8%) of them occurred in only one individual. In Table IV, base substitutions detected in at least one individual are shown below the reference sequence. These polymorphic sites were distributed randomly in the region, though the frequency of each site was variable. Most nucleotide changes were noted in less than 5% of the subjects, but nucleotide changes in 14783 (Cyt b region), 15043 (Cyt b region), 15301 (Cyt b region), 15326 (Cyt b region), 16223 (HVS-I region), and 16362 (HVS-I region) were found in about 81.3%, 81.3%, 75.0%, 100.0%, 68.8%, 62.5% of the subjects, respectively. The observed number of substitutions, classified according to the types of mutations, is shown in Table V. In the whole data, transitions between pyrimidines (TC) are more than those between purines (AG). In these results, three of the polymorphic sites (A16240C, A16351G, G16394A) in HVS-I region are determined newly. Among these polymorphic sites, transition at 16351 (AG) instead of transversion at 16351 (AC) was identified in this study. And, transversion at 16240 (AC) and transition at 16384 (GA) were also identified.

DISCUSSION

The molecular nature of mtDNA polymorphism could be of great use anthropological studies of modern populations. The mtDNA is divided into two domains; a coding region constituting over 90% of the genome and a non-coding region which contains both the origin of H-strand replication and the origins of transcription of both strands (Cann *et al.*, 1987). Accordingly, the noncoding region (Anderson *et al.*, 1981) has been known to be highly susceptible to mutation, hence numerous studies have been performed to characterize the sequence variation of this region for its forensic application. But, this has rarely been determined, especially in Orientals and so the present study could provide basic population data for this group. In addition, a few Korean control region sequences have been published,

Table IV. List of variable sites observed in the 13715-16504 (2,790 bp) of mtDNA obtained from 16 Korean subjects. All nucleotides at these positions are shown together with the reference sequence (Anderson *et al.*, 1981). Dots show matches with the reference sequence. Sites shown in bold italics represent those identified in this study, which have not been previously reported.

	13759	13928	14053	14180	14201	14364	14502	14560	14569	14668	14783	14790	14926	14979	15040	15043	15071	15218	15226	15301	15326	15440	15607	15622	15712	15724	15758	15787	16011	16051	16066	16069	16093	16129	16169	16171	16172	16173	16188	16189	16192	16209	16214	16223	16227	***16240***	16241	16249	16266	16269	16278	16291	16292	16304	16311	***16351***	16362	***16384***	16390
	G	*G*	*A*	*T*	*T*	*G*	*T*	*G*	*G*	*C*	*T*	*A*	*A*	*T*	*C*	*G*	*T*	*A*	*A*	*G*	*A*	*T*	*A*	*T*	*A*	*A*	*A*	*T*	*A*	*A*	*A*	*C*	*T*	*G*	*C*	*A*	*T*	*C*	*C*	*T*	*C*	*T*	*C*	*C*	*A*	***A***	*A*	*T*	*C*	*A*	*C*	*C*	*C*	*T*	*T*	***A***	*T*	***G***	*G*
1	·	·	·	·	·	·	C	·	·	·	C	·	·	·	T	A	C	G	G	A	G	·	·	·	·	·	·	·	G	·	G	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·
2	·	·	·	·	·	A	·	·	·	·	C	G	·	·	·	A	·	·	·	A	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	T	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
3	·	·	·	C	·	·	·	·	·	T	C	·	·	C	T	A	·	·	·	A	G	·	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	T	·	·	·	C	·	G	·	·	·	·	·	·	C	·	·
4	·	·	·	·	C	·	·	·	A	·	C	·	·	·	·	A	·	·	·	A	G	·	·	·	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	G	·	·	·	·	·	T	·	·	·	·	·	C	·	·
5	·	·	·	·	·	·	·	C	·	·	C	·	G	·	·	A	·	·	·	A	G	·	G	C	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	C	A	A
6	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	A	·	·	·	A	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·
7	·	·	·	·	·	·	·	·	·	T	C	·	·	·	·	A	·	·	·	A	G	·	·	·	·	·	·	·	·	G	·	·	·	·	·	G	C	·	·	·	·	·	·	T	·	·	·	·	·	·	·	T	·	·	C	G	C	·	·
8	·	·	·	·	·	·	·	·	·	T	C	·	·	C	·	A	·	·	·	A	G	C	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	T	·	·	·	C	T	·	·	·	·	·	·	·	C	·	·
9	·	·	·	·	·	·	·	·	·	T	C	·	·	C	·	A	·	·	·	A	G	C	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	T	·	·	G	C	T	·	·	·	·	·	·	·	C	·	·
10	·	·	·	·	C	·	·	·	A	·	C	·	·	·	·	A	·	·	·	A	G	·	·	·	·	·	G	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	G	·	·	·	·	·	T	·	·	·	·	·	C	·	·
11	·	·	·	·	·	·	·	·	·	·	C	·	G	·	·	A	·	·	·	A	G	·	G	C	·	G	·	·	·	·	·	·	C	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·	·	·	·	C	·	A
12	·	·	·	·	·	·	C	·	·	·	C	·	·	·	·	A	·	·	·	A	G	·	·	·	G	·	·	·	·	·	·	T	·	·	T	·	·	T	T		T	·	·	T	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·
13	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	G	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	A	·
14	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	A	·	·	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
15	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	A	·
16	A	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	A	·	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·

Table V. Kinds of nucleotide changes compared to those of Anderson on control region of mtDNA in a Korean population. Each number represents cumulative number found 16 Koreans. Insertion and deletion are not determined in these results.

Type of nucleotide change		Entry of data	Number of loci
	Nucleotide substitution		
Transition	G→A	38	8
	A→G	38	18
	T→C	52	17
	C→T	29	13
Transversion	G→T	–	–
	G→C	3	2
	A→T	–	–
	A→C	1	1
	T→G	–	–
	T→A	–	–
	C→G	–	–
	C→A	–	–

but it is difficult to assign them into haplogroups, which are defined by RFLP markers. Thus, target regions of mtDNA in this study can be utilized for the analysis of individual identity as well as population genetic studies. In this study, we are assessed mtDNA polymorphism in a 2,790 bp region of the genome (mt13715-16504) in Korean population. The identification of the variants by PCR-RFLP analysis in this study created a total of 38 distinct genetic haplotypes within Korean population. This high level of diversity was somewhat surprising considering the relatively homogeneous nature of the Korean community. Considering the high number of variables (38 haplotypes) and their low frequencies in most cases, it was not surprising that the frequency differences observed between the two sex groups failed to identify any significant association. Phylogenetic analysis, however, identified 11 groups (Grp 1-11) of related haplotypes encompassing all 38 haplotypes. The frequency differences among haplogroups would lose their significance. Nevertheless, these findings merit re-examination in a larger study and provide a basis for future work. Although restriction enzyme analyses represent polymorphisms over mitochondrial genome from 430 individuals, some ambiguities remain with respect to the actual number of nucleotide differences and the estimation of genetic distances. Thus, the selected samples from haplotype generated by RFLP analysis were sequenced and compared to Anderson's reference sequence. Our results show that polymorphic sites are randomly distributed in the studied control region, though the degree of polymorphism is higher in HVS-I than in other coding region. According to others (Stoneking *et al.*, 1990; Ward *et al.*, 1991), the variation in 16223 was found in 44.2% of Caucasians and 82.1% of Nuu-Chah-Nulths, the variation in 16362 was found in 15.3% of Caucasians and 64.2% of Nuu-Chah-Nulths. These data are similar to those of the Japanese (Yoshii *et al.*, 1995). The variation in 16223 was found in 75.7% of the Japanese population and the variation in 16362 was found in 50.0% of the Japanese population. In this result, 16223 variation was observed lower than Japanese as 68.8%, while 16362 variation was determined higher than Japanese as 62.5%. Although we could not compare my data with those of others phylogenetically, the differences between races were evident. But if the number of subjects becomes larger and more nucleotides are sequenced, the Korean data would become more similar to Japanese than Chinese. Comparison of mitochondrial genotype frequencies among Korean, Japanese and Chinese populations indicates that the Asian populations share most of the genotypes, although their frequencies differ from area to area. Collectively, our results are consistent three main propositions; (i) Cleavage patterns of *Hinf* I, *Hae* III, *Alu* I, *Dde* I, *Mbo* I, *Rsa* I were seen in six morphs, four morphs, nine morphs, eight morphs, seven morphs, and four morphs, respectively. Among the morphs of each enzyme, the A type was shown the most common in all cleavage pattern of restriction endonucleases. (ii) From the results of RFLP with each restriction enzyme, thirty eight mtDNA haplotypes (Hap1-38) were detected in the Korean population. Moreover, phylogenetic analysis identified 11 groups (Grp 1-11) of related haplotypes encompassing all 38 haplotypes. (iii) 59 nucleotide change sites in the selected region were identified. As the results, nucleotide substitution rather than insertion/deletion accounts for the majority pattern of nucleotide changes. The transition type was more prevalent than the transition type. In addition, three of the polymorphic sites (A16240C, A16351G, G16394A) in HVS-I region are newly determined.

In the future, extensive studies on mitochondrial polymorphisms of Korean population will provide useful diagnostic or predictive information, at least to Asian people. In addition, these racial differences are thought to be caused by the high evolution rate, and these data would be valuable for the genealogical study of Orientals.

ACKNOWLEDGEMENT

This work was supported by the 2002 Inje University research grant.

REFERENCES

Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J., Staden, R., and

Young, I. G., Sequence and organization of the human mitochondrial genome. *Nature*, 290, 457-465 (1981).

Ballinger, S. W., Schurr, T. G., Torroni, A., Gan, Y. Y., Hodge, J. A., Hassan, K., Chen, K. H., and Wallace, D. C., Southeast Asian mitochondrial DNA analysis reveals genetic continuity of ancient mongoloid migrations. *Genetics*, 130, 139-152 (1992).

Bhat, P. P., Mishra, B. P., and Bhat, P. N., Polymorphism of mitochondrial DNA (mtDNA) in cattle and buffaloes. *Biochem. Genet.*, 28, 311-318 (1990).

Cann, R. L., Stoneking, M., and Wilson, A. C., Mitochondrial DNA and human evolution. *Nature*, 325, 31-36 (1987).

Finnila, S., Lehtonen, M. S., and Majamaa, K., Phylogenetic network for Eurpean mtDNA. *Am. J. Hum. Genet.*, 68, 1475-1484 (2001).

Giles, R. E., Blanc, H., Cann, H. M., and Wallace, D. C., Maternal inheritance of human mitochondrial DNA. *Proc. Natl. Acad. Sci. U.S.A.*, 77, 6715-6719 (1980).

Graven, L., Passarino, G., Semino, O., Boursot, P., Santachiara-Benerecetti, S., Langaney, A., and Excoffier, L., Evolutionary correlation between control region sequence and restriction polymorphisms in the mitochondrial genome of a large Senegalese Mandenka sample. *Mol. Biol. Evol.*, 12, 334-345 (1995).

Howell, N., Kubacka, I., and Mackay, D. A., How rapidly dose the human mitochondrial genome evolve? *Am. J. Hum. Genet.*, 59, 501-509 (1996).

Howell, N., Human mitochondrial diseases : answering questions and questioning answers. *Int. Rev. Cytol.*, 186, 49-116 (1999).

Ingman, M., Kaessmann, H., Paabo, S., and Gyllensten, U., Mitochondrial genome variation and the origin of modern humans. *Nature*, 408, 708-713 (2000).

Johnson, M. J., Wallace, D. C., Ferris, S. D., Rattazzi, M. C., and Cavalli-Sforza, L. L., Radiation of human mitochondria DNA types analyzed by restriction endonucleases cleavage patterns. *J. Mol. Evol.*, 19, 255-271 (1983).

Macaulay, V., Richards, M., Hickey, E., Vega, E., Cruciani, F., Guida, V., Scozzari, R., Bonne-Tamir, B., Sykes, B., and Torroni, A., The emerging tree of West Eurasian mtDNAs : a synthesis of control-region sequences and RFLPs. *Am. J. Hum. Genet.*, 64, 232-249 (1999).

Richards, M., Corte-Real, H., Forster, P., Macaulay, V., Wilkinson-Herbots, H., Demaine, A., Papiha, S., Hedges, R., Bandelt, H. J., and Sykes, B., Paleolithic and neolithic lineages in the European mitochondrial gene pool. *Am. J. Hum. Genet.*, 59, 185-203 (1996).

Richards, M. B., Macaulay, V. A., Bandelt, H. J., and Sykes, B. C., Phylogeography of mitochondrial DNA in western Euroupe. *Ann. Hum. Genet.*, 62, 241-260 (1998).

Rose, G., Passarino, G., Carrieri, G., Altomare, K., Greco, V., Bertolini, S., Bonafe, M., Franceschi, C., and De Benedictis, G., Paradoxes in longevity : sequence analysis of mtDNA haplogroup J in centenarians. *Eur. J. Hum. Genet.*, 9, 701-707 (2001).

Schurr, T. G., Sukernik, R. I., Starikovskaya, Y. B., and Wallace, D. C., Mitochondrial DNA variation in Koryaks and Itel'men : population replacement in the Okhotsk Sea-Bering Sea region during the Neolithic. *Am. J. Phys. Anthropol.*, 108, 1-39 (1999).

Stoneking, M., Jorde, L. B., Bhatia, K., and Wilson, A. C., Geographic variation in human mitochondrial DNA from Papua New Guinea. *Genetics*, 124, 717-733 (1990).

Torroni, A., Huoponen, K., Francalacci, P., Petrozzi, M., Morelli, L., Scozzari, R., Obinu, D., Savantaus, M. L., and Wallace, D. C., Classification of European mtDNAs from an analysis of three European populations. *Genetics*, 144, 1835-1850 (1996).

Torroni, A., Schurr, T. G., Cabell, M. F., Brown, M. D., Neel, J. V., Larsen, M., Smith, D. G., Vullo, C. M., and Wallacw, D. C., Asian affinities and continental radiation of the four founding Native American mtDNAs. *Am. J. Hum. Genet.*, 53, 563-590 (1993).

Vigilant, M., Stoneking, M., Harpending, H., Hawkes, K., and Wilson, A. C., African populations and the evolution of human mitochondrial DNA. *Science*, 253, 1503-1507 (1991).

Wallace, D. C., Brown, M. D., and Lott, M. T., Mitochondrial DNA variation in human evolution and disease. *Gene*, 238, 211-230 (1999).

Ward, R. H., Frazier, B. L., Dew-Jager, K., and Paabo, S., Extensive mitochondrial diversity within a single Amerindian tribe. *Proc. Natl. Acad. Sci. U.S.A.*, 88, 8720-8724 (1991).

Wilson, A. C., The molecular basis of evolution. *Sci. Am.*, 253, 164-173 (1985).

Yoshii, T., Takada, E., Akiyama, K., and Ishiyama, I., Sequence polymorphism of mitochondrial DNA and its forensic application. *Nippon. Hoigaku. Zassahi.*, 49, 242-250 (1995).