

Next Generation Sequencing-Based Mitochondrial Genome Analysis In Relation To The Drug- Induced Toxicity In Korean Subjects

Pham Vinh Hoa

Inje University College of Medicine

Nguyen Van Lam

Inje University College of Medicine

Hye-Eun Jung

Inje University College of Medicine

Yong-Soon Cho

Inje University College of Medicine

Jae-Gook Shin (✉ phshinjg@gmail.com)

Inje University College of Medicine

Research Article

Keywords: Mitochondria, polymorphism, drug-induced toxicity, Korean population.

Posted Date: November 1st, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-968276/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Mitochondrial variants have been investigated to be associated with many diseases, which was reported largely from European populations. Few studies, however, have annotated the whole mitochondrial DNA (mtDNA) genome associated with drug responses including adverse drug reactions (ADRs), especially in Asian populations. This study was performed to characterize mtDNA genetic profiles, especially the distribution and frequency of well-known genetic biomarkers associated with diseases and drug-induced toxicity, in the Korean population by using high throughput next-generation sequencing.

Results: A total of 681 variants was identified among all 118 Korean subjects. The MT-TRNP and displacement loop (D-loop) showed the highest numbers of variants (113 and 74 variants, respectively). In the D-loop, 25.4% of the subjects were identified to have m.16189T>C allele, which was known to reduce the mtDNA copy number variation in human cells. The variant m.1095T>C, annotated to aminoglycoside-induced ototoxicity, was found from only one subject in this study, while the frequency of m.2706A>G and m.3010A>G, which are associated with antibiotic-induced toxicity, were 99.15% and 30.51%, respectively. The 5 subjects were identified to have the variant m.2150T>TA, a genotype associated with highly disruptive effects on mitochondrial ribosomes. In addition, the mitochondrial haplogroups of 118 Korean subjects were found that D and M groups were the most dominant groups with the frequency of 34.74% and 16.1%, respectively.

Conclusions: Our finding was constant with Korean 1K project and well reflected the unique profile of mitochondrial haplogroup distribution. It was the first study to annotate the whole mitochondrial genome with drug-induced toxicity to predict the ADRs event in clinical implementation for Korean population. This approach could be extended for further study for validation the potential ethnic specific mitochondrial genetic biomarkers in Korea population.

Background

Mitochondria are organelles that play a central role in cellular energy suppliers. They have their own genome and genetic code, and an exceptionally high mutation rate [1]. At least 100 mutations are expected to be observed in almost 90% of non-proliferating cells, while no other cell types have fewer than 10 mutations by the age of 70 years [2]. Polymorphism of mitochondrial DNA (mtDNA) is associated with various pathophysiology, and could explain diverse vulnerability to the diseases or drug toxicity [3]. The high variability of human mitochondria has been studied in the context of common diseases. Ninety-five mitochondrial markers in the Mitomap database have been confirmed to be pathogenic as of 2020 [4]. Pathogenic mtDNA mutations were reported to be common in the general population, and are also present in some major haplogroups. Therefore, many healthy individuals carry potential harmful variants. At least 1 in every 200 healthy subjects harbours a pathogenic variant that can be a potential cause of disease in the next generation [5].

In addition, mitochondrial polymorphisms were reported to be associated with drug responses, and can reveal the role of mtDNA variation in susceptibility to drug toxicity. It has been reported that variations in mtDNA can result in differences in mitochondrial function that, in turn, may lead to idiosyncratic drug-induced toxicity [6]. Limited information, however, is available about the roles of mitochondrial haplogroups in susceptibility to drug-induced toxicity. Anti-retroviral therapy (ART), antibiotics and chemotherapeutic agents are well known drug classes in relation to the mitochondria mediated drug toxicities [7]. The mtDNA mutations and copy numbers have been proposed as potential biomarkers to monitor therapeutic responses and prognosis in cancer treatments, although the mechanisms are not well established [8, 9]. There are reports that specific mtDNA haplogroups are associated with ART-induced peripheral neuropathy and metabolic disorders from clinical studies conducted in European patients [10–15]. The susceptibility of antibiotics toxicity is also well known to associated with mtDNA polymorphisms, especially in mitochondrial 12S rRNA and 16S rRNA genes which play an important role in preventing or inducing toxicity [16].

The mtDNA has a unique genetic feature known as 'heteroplasmy', which allows mutant and wild-type mtDNAs to coexist. It is well known that healthy individuals harbour relatively low levels (< 1%) of mtDNA heteroplasmy in general. It is known that the mtDNA heteroplasmy increases throughout the lifespan, and the initial mutant mtDNA will be the predominant type within a specific cell. Once a certain threshold level of mutant mtDNA is above the normal range, the mutant load influences on the cellular and patient phenotypes. Hence, a low level of heteroplasmy would be linked to late-onset of related diseases in an individual subject [17]. Since the Next-generation sequencing (NGS) technology is a tool to identify and quantify the level of heteroplasmy at the 1% range, this high throughput technology has been implemented extensively towards decoding the mitochondrial genome in many ethnic populations [18–23]. Two reports are available so far for the NGS-based mitochondrial genome study in a Korean subject [24, 25], and they found the similar profile of mitochondrial genetic variants with a Northern Chinese Han and Japanese populations. However, those reports are not issued on the potential implementation of those mtDNA variants as potential biomarkers for the prediction of drug induced toxicity.

In this study, therefore, the mtDNA genetic profiles of 118 Korean subjects were analysed by using the NGS approach, in order to investigate the distributions and frequencies of known mtDNA variant biomarkers associated with diseases and drug-induced toxicity in a Korean population.

Methods

Sample collection and DNA extraction

DNA samples were harvested from peripheral blood that were selected from biobank in Pharmacogenomics Research Center, Inje University, Korea. In totally, 118 Korean subjects, all of whom provided written informed consent to participate in the genotyping analysis, were adults with a mean age of 37.1 years (range: 19–65 years). The study population consisted of 65 males (55%) and 88 (74.6%) non-smokers. Sixty-eight subjects were diagnosed with drug-susceptible tuberculosis and treated with

first-line anti-tuberculosis drugs (rifamycin, isoniazid, pyrazinamide and ethambutol). These subjects had no adverse drug reactions, no comorbidities and normal laboratory test results during treatment. Blood samples were collected at the time of recruitment and frozen at -80°C . Total DNA was isolated from peripheral blood samples using a QIAamp DNA Blood Mini Kit (Qiagen, German) according to the manufacturer's instructions. DNA purity and concentration were measured using a NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific, USA).

Amplification of mitochondrial long-range fragments

A well-established method for exclusive amplification of mitochondrial long-range fragments as described by Gould *et al.* [24] was used to avoid the co-amplification of nuclear genomic DNA. The PCR primers designed for the study included MT-L1_F (5'-AAATCTTACCCCGCCTGTTT-3'), MT-L1_R (5'-AATTAGGCTGTGGGTGGTTG-3'), MT-L2_F (5'-GCCATACTAGTCTTTGCCGC-3'), and MT-L2_R (5'-GGCAGGTCAATTTCACTGGT-3'). We used an Applied Biosystems (GeneAmp PCR system 9700, Thermo Fisher Scientific, USA) for analysing long-range fragments 1 and 2 (MT-L1 and MT-L2, respectively) of mitochondrial genomes (~8.5 kb each), as described by McElhoe *et al.* The PCR was performed using 50- μl reaction mixtures containing 200 ng of genomic DNA with LA Taq® and 2 \times GC buffer (TaKaRa, Japan) [25]. The PCR conditions for both the MT-L1 and MT-L2 fragments consisted of an initial denaturation step at 94°C for 1 min followed by 30 cycles of denaturation at 94°C for 30 s, annealing at 59°C for 40 s, extension at 72°C for 9 min, and a final extension step of 72°C for 10 min. The DNA template (9947A) from the National Institute of Standards and Technology (NIST, USA) [26] was utilised as the mitochondrial standard DNA for sequencing.

Mitochondrial whole-genome sequencing

DNA libraries with an expected insert size of ~200 bp were prepared using NEBNext (UK) in accordance with the instruction manual provided [27], and optimised for our samples. The concentrations of mitochondrial libraries were measured using the QuantiFluor® ONE dsDNA system (Promega, USA). Equimolar amounts of the 60 indexed libraries were pooled to obtain a 4-nM mixture. After denaturation and further dilution, the final 16 pM of this mixture was loaded into an Illumina cartridge. Sequencing was performed using the Illumina MiSeq Reagent kit v2 (300 cycles) and MiSeqDx instrument (Illumina, USA) in accordance with the manufacturer's instructions.

Analysis of mitochondrial genomes

A bioinformatics pipeline was developed to reconstruct and analyse human mtDNA from high-throughput sequencing data. The sequencing data obtained from the Illumina MiSeqDx were quality-controlled using FastQC (version 0.11.9) [28] and MultiQC (version 16) [29]. All adapters and sequences with a quality score < 30 and sequence length < 70 bp were removed using Trimmomatic (version 0.39) [30]. In the alignment process, to minimise the nuclear mitochondrial DNA segment (NUMT), the qualified data of each sample were aligned simultaneously to human genome reference hg19 based on the revised Cambridge Reference Sequence (rCRS) of human mitochondrial DNA [31] using BWA (version 0.7.17-r1188) [32]. Regarding downstream alignment processes such as conversion from Sequence

Alignment/Map (SAM) to Binary Alignment/Map (BAM) format and marking and removing duplicates were performed using GATK (version 4.1.8.1). The final alignment assessment was performed using Qualimap (version v.2.2.1) [33]. Variants and heteroplasmy were called using GATK Mutect2 (version 4.1.8.1) in mitochondrial mode. All variants with a heteroplasmy level < 5% were filtered for further analysis. Other filtering processes were in accordance with the instructions of the GATK best practices for mitochondrial genomes [34]. The results of the variant-calling pipeline were confirmed based on the NIST mtDNA standard (GM09947A) [26] before applying to real samples. Finally, all variant-calling pipelines were combined. The variants were annotated using HmtNote (version 0.7.2) [35] against the HmtVar database [36]. The haplogroups of all samples in the cohort were also checked using HaploGrep2 [37].

Results

Distribution of mitochondrial variants in the sequenced mtDNAs

The bioinformatics pipeline was confirmed to be consistent with the variant-calling pipeline for the mtDNA standard sequence [26]. According to NIST [38], the percentage of heteroplasmy for m.1393G>A was $17.4 \pm 1.7\%$ and that for m.7861T>C was $74.6 \pm 14.5\%$, while the values in our study were $17.4 \pm 0.6\%$ and $89.5 \pm 0.1\%$, respectively. These results confirmed the accuracy of the bioinformatics pipeline and experimental procedure for mtDNA sequencing and analysis. The average mtDNA coverage for the examined samples was $2,248X \pm 599X$.

We identified a total of 681 variants called from 118 individuals (Figure 1). Most of the variable positions were transitions (86.5% overall), while the overall percentage of transversions was about 3%. Our study showed that transitions were dominant, as reported previously for human mitochondrial genomes [39]. Indel polymorphism was detected at a rate of 8.5% among all individuals. The median of the mtDNA variants was 38 (range: 26–55). No variants were found in 9/22 tRNA-coding genes. MT-TRNP and the displacement loop (D-loop) had the most polymorphisms (113 and 74 variants, respectively) (Supplementary Table 1). No disease mutations with ‘confirmed’ status in the current MITOMAP database [39] were found in any individuals.

The lowest heteroplasmy level of mtDNA variants was 10% for several non-coding variants. The levels of heteroplasmy of some mutations (m.3206C>T, m.4824A>G, m.8473T>C, m.16179CA>C, and m.16183A>C) varied from approximately 10–100% (homoplasmy) and their allele frequency in the studied population was 11%. The m.16189T>C polymorphism causing a lower mtDNA copy number was present in 29/118 (24.58%) samples as homoplasmy, and in one healthy individual at a heteroplasmy level of 54.6% (Supplementary Table 2).

Classification of mitochondrial haplogroups in the Korean population

We examined all haplogroups previously reported in the Korean population in Korea1K ($n = 1,094$) [40], and listed in the Korean National Standard Reference Variome Database of whole genomes (KoVariome) [41]. Eleven variants (m.73A>G, m.263A>G, m.750A>G, m.1438A>G, m.2706A>G, m.4769A>G, m.7028C>T, m.8860A>G, m.11719G>A, m.14766C>T, and m.15326A>G) with an overall frequency $\geq 50\%$ that are widespread across all lineages (L, M, and N) [42] were predominant in our study. Among these, five mtDNA variants present at $\geq 80\%$ in lineages L, M, or N were found in all individuals. The rare mutation m.73A>C was observed in one individual in our population. The most dominant haplogroup for the Korean population was found to be D (34.74%), followed by M (16.1%). These haplogroups are prevalent in Asian populations [43] and were diverse in our study subgroups (Figure 2A, Table 1). The D2 sub-haplogroup, defined by mutation m.16271T>C [44] and possibly equating to D4e1b [45], was observed in one subject. Consequently, 87.8% of haplogroup D belonged to the D4 subgroup, which had the highest frequency in our Korean population; the D5 subgroup was observed in all remaining subjects. These results were consistent with the data of Korea1K, which includes sequence data from 1,094 whole genomes.

Table 1
The sub-haplogroups observed in our study (n=118) by using Haplogrep analysis

Haplogroup	Sub-haplogroup	Number of subjects (n=118)
A	A+152+16362, A1, A11, A5a, A5a1a, A5b1a, A5b1b, A6,	12 (10.17%)
B	B4a1b1a, B4a1c1a1, B4a4, B4b1a2, B4b1a2a, B4c1a1a, B4c1a1b, B4d3a1, B4f1, B5a2a1+16129, B5b, B5b1, B5b2a2,	14 (11.86%)
C	C4a1a, C7a1c,	2 (1.69%)
D	D2, D4, D4+195, D4a, D4a1, D4a1c, D4a1h, D4a3b, D4a3b1, D4b2, D4b2a2a1, D4b2b, D4b2b1, D4c1a, D4c2c, D4e1a, D4e2, D4f1, D4g1, D4g1a, D4g1b, D4g1c, D4h1a, D4h1c1, D4i, D4n, D5a, D5a2a1+16172, D5b, D5b1, D5b1b,	41 (34.75%)
F	F1a1, F2, F2f, F2i,	4 (3.39%)
G	G1a1, G1a1a, G1a1a1, G2a+152, G2a1, G2a1+16189+16194, G2a1b, G2a1e, G2a5, G2b2,	14 (11.86%)
M	M10, M10a1a1b, M10a1b, M11b1a, M7a1a, M7a1a1, M7a1a5, M7b1a1a1, M7c1a2a1, M7c1a3, M7c1a5, M7c1b, M8a3a, M8a3a1, M9a1a1, M9a1a1a,	19 (16.1%)
N	N9a2, N9a2a, N9a2c, N9a2d,	5 (4.24%)
R	R+16189, R11a,	2 (1.69%)
Y	Y1,	4 (3.39%)
Z	Z4a,	1 (0.85%)

The distributions of the haplogroups in the Korean population were obtained from Korea1K [40], KoVariome [41] and five super-populations in the 1000 Genomes Project (1KG) [31] and then compared to our data. Our study showed the concordance of the haplogroups of the Korean and East Asian populations (1KG-EAS) (Figure 2B). Other super-populations had distinct mtDNA haplogroups, as described elsewhere [46]. Haplogroups M, A, H and L were most prevalent in South Asian (1KG-SAS), American (1KG-AMR), European (1KG-EUR) and African (1KG-AFR) populations, respectively.

Correlations between annotated mtDNA variants and drug-induced toxicity

The D-loop is 1,122 bp in length, and consists of two hypervariable regions (HVI at nucleotides 16024–16383 and HVII at nucleotides 57–372) and a tandem repeat of poly(C) at nucleotides 303–315 [47]. In all subjects, the mtDNA D-loop region had 74 polymorphisms. All 118 samples had the polymorphism m.263A>G in HVII and one subject carried m.73A>C, which was reported to be rare in Asian populations (0.01338509% in healthy individuals and 0.08976661% in patients). In 30/118 (25.42%) of the samples in the present study, a thymine to cytosine (T>C) transition was found at position 16189, producing an uninterrupted homopolymeric tract at mtDNA segment 16180–16195 in HVI. In total, 3 of the 118 samples (2.5%) had 12 uninterrupted cytosine residues, with the heteroplasmy level of m.16189T>C varying from 54.6–99.9%. Meanwhile, 2 of the 118 (1.7%) samples had length heteroplasmy in the mitochondrial HVII segment at positions 303–315 due to the insertion of cytosine at position 309, and one subject showed replication slippage in the case of a T>C transition at position 310.

We observed a total of 40 mitochondrial variants in mitochondrially encoded 12S rRNA (MT-RNR1) and mitochondrially encoded 16S rRNA (MT-RNR2) genes. Several mtDNA mutations associated with antibiotic-induced toxicity were detected in these two genes (Table 2). The m.1555A>G and m.1494C>T mutations are well known to be associated with ototoxicity in patients treated with aminoglycoside antibiotics, but were not observed in our population. We may not have been able to detect these rare variants, reported to be found at a rate of 2 per 2,922 individuals (0.07%) in the Korean population [48], due to the small sample size. Other aminoglycoside-induced hearing loss variant, m.1189T>C, was observed in 1.69% of all samples; the rate was previously reported as 0.34% for the Korean population [49, 50]. Variant at position 961 was identified at a frequency of 5.93%; these are more likely to be found in Korean than Asian populations (1–3% in the 1KG). The polymorphism of m.2706A>G transition, a variation previously suggested to lead to a predisposition to linezolid-associated lactic acidosis [51, 52], was dominant in all subjects. The m.3010G>A allele, which is associated with linezolid-induced mitochondrial toxicity, was more common in the Korean population (30.51%), consistent with previous studies [44,]. Lastly, we observed a high frequency of m.10398A>G variant that was reported to influence metabolic ART effects [53] at homoplasmic level in our study.

Table 2

Frequencies of mitochondrial variants in association with drug-induced toxicity in this study (n=118)

Gene	Variant	rsID	Drug	Clinically relevant	Frequency in our study (%)
MT-RNR1	m.663A>G	rs56489998	Aminoglycoside	Ototoxicity	10.17%
	m.961T>C	rs3888511	Aminoglycoside	Ototoxicity	5.93%
	m.961T>C	rs3888511	Linezolid	Mitochondrial toxicity	5.93%
	m.1095T>C	rs267606618	Aminoglycoside	Ototoxicity	0.85%
	m.1189T>C	rs28358571	Aminoglycoside	Ototoxicity	1.69%
MT-RNR2	m.2706A>G	rs2854128	Linezolid	Lactic acidosis, Mitochondrial toxicity	99.15%
	m.3010G>A	rs3928306	Linezolid	Mitochondrial toxicity	30.51%
MT-ND3	m.10398A>G	NA	ART	Metabolic/ cardiovascular complications	70.34%
* MT-RNR1: Mitochondrially encoded 12S ribosomal RNA; MT-RNR2: Mitochondrially encoded 16S ribosomal RNA; MT-ND3: Mitochondrially encoded NADH: Ubiquinone Oxidoreductase Core Subunit 3.					

Discussion

Clinical implications of D-loop variants

The D-loop has a regulatory role in mtDNA transcription and replication [54]. Mutations in the D-loop region may significantly affect mtDNA copy number and the gene expression of mitochondrial genomes, thereby disrupting the function of mitochondria, oxidative phosphorylation and ATP production. Here, we found the largest number of polymorphisms in the D-loop region in all subjects. Notably, 25.42% of the samples harboured the mutation m.16189T>C, which is very close to the mtDNA origin of replication and is significantly associated with various non-communicable diseases [55–58]. This could be explained by the lower binding affinity of proteins to the region with the 16189C variant [55]. A negative correlation between continuous cytosines at the mtDNA segment 16180–16195 in HVI and mtDNA copy number was reported previously in human peripheral blood cells [59]. The reduction of mitochondrial copy number might lead to the mitochondrial dysfunction that is potentially a mechanism of drug-induced toxicity [60]. It suggested that m.16180T>C variant might increase the risk of developing adverse drug reaction. Thus, the impact of this variant should be considered in further investigation.

The non-synonymous m.10398A>G polymorphism (overall frequency = 70.34%) was reported to be associated with long-term ART-induced toxicity. Notably, most of the variants mentioned above showed homoplasmy (> 98.4% of all mtDNA sequences analysed [61]). The phenotypic expression of pathogenic

and heteroplasmic mtDNA mutations could be modulated by homoplasmic mtDNA variants [53]. This suggests that determination of these variants prior to treatment could be useful to identify tuberculosis and HIV patients at risk of drug-induced toxicity.

The frequency of length heteroplasmy in the homopolymeric C-stretch regions was as low as ~4% in our population, and we detected only one individual with possible length heteroplasmy at position 303–315. Compared to point heteroplasmy, length heteroplasmy is more common and less population-specific [62]. The presence of a tandem poly(C) repeat in the D-loop region of nucleotides 303–315 and 16182–16189, as a result of polymerase misincorporation or slippage during mtDNA replication [63], reduced the mtDNA copy number. These length heteroplasmies were reportedly associated with a more than twofold increase in cancer risk [64]. In addition, these length variants have been suggested to be associated with several common diseases, including diabetes mellitus and dilated cardiomyopathy [65], and are therefore of clinical importance.

Clinical implications of mtDNA variants in coding genes

The human mitochondrial genome consists of 13 protein-coding genes, 22 tRNAs and 2 rRNAs. The proteins encoded by the 13 mitochondrial-coding genes are constituents of the enzyme complexes of the oxidative phosphorylation system, enabling mitochondria to act as the cellular powerhouse. They are distributed among various respiratory chain components, in a small part of complex I and throughout complexes III–V (ATP synthase) of the respiratory chain [66].

Mitochondrial rRNA more closely resembles bacterial than human rRNA. However, genetic polymorphisms in these genes can result in greater similarity between mitochondrial and bacterial rRNA, thereby facilitating the binding of anti-microbials (most notably aminoglycosides). In cases of tuberculosis infection treated via multi-antibiotic administration, the associations of mtDNA instability and variation with drug-induced toxicity should be taken into consideration. The polymorphisms in the mitochondrial RNR1 and RNR2 genes were associated with increased mitochondrial and clinical adverse effects, most commonly ototoxicity and peripheral neuropathy [67, 68]. The mutation of mitochondrial 12S rRNA gene is associated with both aminoglycoside-induced deafness and non-syndromic hearing loss. This study detected higher allele frequency in some variants (m.663A>G, m.961T>C, m.3010G>A) compared to previous studies [44, 53]. This suggested that these variants should be considered carefully in studies of the Korean population. It has recently been reported that mtDNA polymorphisms may affect respiratory chain function and first-line anti-tuberculosis drug-induced liver injury. However, the underlying mechanisms have yet to be elucidated [69, 70].

The rare mutations observed in our clinical samples may have highly disruptive effects on mitochondrial ribosomes; based on clinical data, potential disruption is classified as ‘proven’ or ‘not enough evidence’ (NEE) [71, 72]. Among 52 and 145 mutations in the MT-RNR1 and MT-RNR2 genes, respectively, only three (m.1494C>T, m.1555A>G and m.1843T>C) were ‘proven’ in clinical practice. Most of the remaining variants were classified as NEE, so it is not yet clear whether changes in mitochondrial ribosomes affect the binding sites of antibiotic drugs to reduce drug toxicity. Other mtDNA-encoded ribosome components

were suggested to have potential benefits in terms of drug susceptibility and should therefore be investigated further in clinical studies. In the present study, only m.2150T>TA, classified as NEE, was observed (at a rate of 4.24%; 5/118 samples) as a homoplasmic mutant.

In addition, adverse drug reactions were reported in patients inheriting pathogenic mtDNA mutations. Antibiotic drugs were reported to be associated with optic neuropathy in Leber's hereditary optic neuropathy (LHON) carriers. In patients carrying the m.11778G>A (MT-ND4) mutation, erythromycin has the potential to catalyse a bioenergetic crisis at onset of LHON [73], while ethambutol was suggested to have a synergistic and deleterious effect on tissue specificity, as reviewed elsewhere [74]. Therefore, the use of erythromycin and, by extension, other macrolides, should be avoided in patients with pre-existing pathogenic mtDNA mutations.

Haplotype and risk assessment of diseases and drug-induced toxicity

mtDNA polymorphisms have been shown to be associated with various pathophysiological conditions, and could play a role in various physiological and pathological characteristics [3]. The high variability in human mitochondria has been investigated for common diseases [2]. Ninety-five mitochondrial markers in the Mitomap database have been confirmed to be pathogenic, as of 2020 [4]. Pathogenic mtDNA mutations were commonly found in a number of major haplogroups in the general population, indicating that many healthy individuals carry these potentially harmful mtDNA mutations. It was reported that, among 200 healthy individuals, at least one would harbour a pathogenic mtDNA mutation with the potential to cause disease in the next generation [5]. Therefore, mtDNA mutations could serve as biomarkers facilitating early detection and prediction of the prognosis of disease. However, our study identified no pathogenic variants, possibly due to its small sample size.

In European populations, it has been reported that specific mtDNA haplogroups were associated with ART-induced peripheral neuropathy and metabolic disorders [10–15]. In addition, many studies have demonstrated relationships between haplogroups and cancer risk. However, at the time of our study, there had been no investigations of the relationships between haplogroups and drug-induced toxicity for Korean population. Haplogroups vary widely across ethnic groups, even between South and East Asians. Therefore, the characteristics of haplogroups in specific populations should be considered when exploring the susceptibility of these populations to diseases and drug-induced toxicity.

Conclusions

Drug-induced toxicity via mitochondrial dysfunction has been studied extensively. However, the contributions of mitochondrial variants to drug-induced toxicity remain unclear. The present study identified distinct haplotypes in a Korean population, consistent with the results reported previously by Korea1K. Despite the relatively small sample size of our study, we investigated correlations between mtDNA characteristics and risk factors for certain diseases and drug-induced toxicity, through NGS analysis. Further studies including larger cohorts are needed to examine the associations of mtDNA with

adverse drug reactions in the Korean population. For this issue, we are going to extend the investigation in association with antituberculosis drug-induced toxicity such as drug-induced liver injury and peripheral neuropathy.

Declarations

Ethics approval and consent to participate

The study was conducted following the Declaration of Helsinki and institutional criteria. The institutional review boards (IRBs) of all participating sites reviewed and approved the research. Written informed consents were obtained from all participants.

Consent for publication

Not applicable

Availability of data and materials

The datasets generated and/or analysed during the current study are available in the PRJNA774877 repository, <https://www.ncbi.nlm.nih.gov/sra/?term=PRJNA774877>

Competing interests

The authors state that they have no conflict of interest.

Funding

Not applicable

Authors' contributions

Pham Vinh Hoa involved on study design and experimental conduction, data analysis and manuscript preparation; Nguyen Van Lam involved on data analysis and manuscript preparation; Hye-Eun Jung involved on study design, data analysis, and manuscript editing; Yong-Soon Cho involved on data analysis and manuscript revision; Jae-Gook Shin involved on project and study design, and data analysis, and manuscript edition and funding of grant.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by Korean government (MIST) (No.2018R1A52021242). The funding source had no role in the study design, in the data collection or analysis, in the preparation of the manuscript, or in the decision to submit the paper for publication.

Authors' information (optional)

Author's address:

Pham Vinh Hoa – Email: hoaptv@gmail.com

Co-authors' email address:

Nguyen Van Lam – Email: lamnvcnsh@gmail.com

Hye-Eun Jung – Email: jung.he@spmed.kr

Yong-Soon Cho – Email: yscho@gmail.com

References

1. D. C. Wallace, "Mitochondrial DNA Variation in Human Radiation and Disease," *Cell*, vol. 163, no. 1. Cell Press, pp. 33–38, Sep. 2015, doi: 10.1016/j.cell.2015.08.067.
2. H. Li, J. Slone, L. Fei, and T. Huang, "Mitochondrial DNA Variants and Common Diseases: A Mathematical Model for the Diversity of Age-Related mtDNA Mutations," *Cells*, vol. 8, no. 6, p. 608, Jun. 2019, doi: 10.3390/cells8060608.
3. H. A. L. Tuppen, E. L. Blakely, D. M. Turnbull, and R. W. Taylor, "Mitochondrial DNA mutations and human disease," *Biochimica et Biophysica Acta - Bioenergetics*, vol. 1797, no. 2. Elsevier, pp. 113–128, Feb. 01, 2010, doi: 10.1016/j.bbabi.2009.09.005.
4. "ConfirmedMutations < MITOMAP < Foswiki."
5. H. R. Elliott, D. C. Samuels, J. A. Eden, C. L. Relton, and P. F. Chinnery, "Pathogenic Mitochondrial DNA Mutations Are Common in the General Population," *Am. J. Hum. Genet.*, vol. 83, no. 2, pp. 254–260, Aug. 2008, doi: 10.1016/j.ajhg.2008.07.004.
6. U. A. Boelsterli and P. L. K. Lim, "Mitochondrial abnormalities-A link to idiosyncratic drug hepatotoxicity?," *Toxicology and Applied Pharmacology*, vol. 220, no. 1. Toxicol Appl Pharmacol, pp. 92–107, Apr. 01, 2007, doi: 10.1016/j.taap.2006.12.013.
7. S. L. Penman, A. S. Carter, and A. E. Chadwick, "Investigating the importance of individual mitochondrial genotype in susceptibility to drug-induced toxicity," *Biochemical Society Transactions*, vol. 48, no. 3. Portland Press Ltd, pp. 787–797, Jun. 01, 2020, doi: 10.1042/BST20190233.
8. Y. Yuan *et al.*, "Comprehensive molecular characterization of mitochondrial genomes in human cancers," *Nat. Genet.*, vol. 52, no. 3, pp. 342–352, Mar. 2020, doi: 10.1038/s41588-019-0557-x.
9. W. X. Zong, J. D. Rabinowitz, and E. White, "Mitochondria and Cancer," *Molecular Cell*, vol. 61, no. 5. Cell Press, pp. 667–676, Mar. 03, 2016, doi: 10.1016/j.molcel.2016.02.011.
10. T. Hulgan *et al.*, "Mitochondrial haplogroups and peripheral neuropathy during antiretroviral therapy: An adult AIDS clinical trials group study," *AIDS*, vol. 19, no. 13, pp. 1341–1349, Sep. 2005, doi: 10.1097/01.aids.0000180786.02930.a1.

11. J. A. Canter *et al.*, "African mitochondrial DNA subhaplogroups and peripheral neuropathy during antiretroviral therapy," *J. Infect. Dis.*, vol. 201, no. 11, pp. 1703–1707, Jun. 2010, doi: 10.1086/652419.
12. S. L. Hendrickson *et al.*, "Mitochondrial DNA haplogroups influence lipoatrophy after highly active antiretroviral therapy," *J. Acquir. Immune Defic. Syndr.*, vol. 51, no. 2, pp. 111–116, Jun. 2009, doi: 10.1097/QAI.0b013e3181a324d6.
13. D. Micheloud *et al.*, "European mitochondrial DNA haplogroups and metabolic disorders in HIV/HCV-coinfected patients on highly active antiretroviral therapy," *J. Acquir. Immune Defic. Syndr.*, vol. 58, no. 4, pp. 371–378, Dec. 2011, doi: 10.1097/QAI.0b013e31822d2629.
14. T. Hulgan *et al.*, "European mitochondrial DNA haplogroups and metabolic changes during antiretroviral therapy in AIDS Clinical Trials Group Study A5142," *AIDS*, vol. 25, no. 1, pp. 37–47, Jan. 2011, doi: 10.1097/QAD.0b013e32833f9d02.
15. T. Hulgan *et al.*, "Mitochondrial DNA variation and changes in adiponectin and endothelial function in HIV-infected adults after antiretroviral therapy initiation," *AIDS Res. Hum. Retroviruses*, vol. 29, no. 10, pp. 1293–1299, Oct. 2013, doi: 10.1089/aid.2013.0079.
16. D. Pacheu-Grau, A. Gó Mez-Durá, M. J. Ló Pez-Pé Rez, J. Montoya, and E. Ruiz-Pesini, "Mitochondrial pharmacogenomics: barcode for antibiotic therapy," *Drug Discov. Today*, vol. 15, no. 2, 2010, doi: 10.1016/j.drudis.2009.10.008.
17. J. B. Stewart and P. F. Chinnery, "The dynamics of mitochondrial DNA heteroplasmy: Implications for human health and disease," *Nat. Rev. Genet.*, vol. 16, no. 9, pp. 530–542, 2015, doi: 10.1038/nrg3966.
18. S. Maruyama, T. Komuro, H. Izawa, and H. Tsutsumi, "Analysis of human mitochondrial DNA polymorphisms in the Japanese population," *Biochem. Genet.*, vol. 51, no. 1–2, pp. 33–70, Feb. 2013, doi: 10.1007/s10528-012-9542-0.
19. K. Yamamoto *et al.*, "Genetic and phenotypic landscape of the mitochondrial genome in the Japanese population," *Commun. Biol.*, vol. 3, no. 1, pp. 1–11, Dec. 2020, doi: 10.1038/s42003-020-0812-9.
20. N. T. Duong *et al.*, "Complete human mtDNA genome sequences from Vietnam and the phylogeography of Mainland Southeast Asia," *Sci. Rep.*, vol. 8, no. 1, p. 11651, Dec. 2018, doi: 10.1038/s41598-018-29989-0.
21. M. Stoljarova, J. L. King, M. Takahashi, A. Aaspõllu, and B. Budowle, "Whole mitochondrial genome genetic diversity in an Estonian population sample," *Int. J. Legal Med.*, vol. 130, no. 1, pp. 67–71, Jan. 2016, doi: 10.1007/s00414-015-1249-4.
22. R. S. Just, T. M. Diegoli, J. L. Saunier, J. A. Irwin, and T. J. Parsons, "Complete mitochondrial genome sequences for 265 African American and U.S. 'Hispanic' individuals," *Forensic Sci. Int. Genet.*, vol. 2, no. 3, Jun. 2008, doi: 10.1016/j.fsigen.2007.12.001.
23. J. L. King *et al.*, "High-quality and high-throughput massively parallel sequencing of the human mitochondrial genome using the Illumina MiSeq," *Forensic Sci. Int. Genet.*, vol. 12, pp. 128–135,

- 2014, doi: 10.1016/j.fsigen.2014.06.001.
24. M. P. Gould, C. M. Bosworth, S. McMahon, S. Grandhi, B. T. Grimer, and T. La Framboise, "PCR-free enrichment of mitochondrial DNA from human blood and cell lines for high quality next-generation DNA sequencing," *PLoS One*, vol. 10, no. 10, pp. 1–13, 2015, doi: 10.1371/journal.pone.0139253.
 25. J. A. McElhoe *et al.*, "Development and assessment of an optimized next-generation DNA sequencing approach for the mtgenome using the Illumina MiSeq," *Forensic Sci. Int. Genet.*, vol. 13, pp. 20–29, 2014, doi: 10.1016/j.fsigen.2014.05.007.
 26. B. C. Levin, D. K. Hancock, K. A. Holland, H. Cheng, and K. L. Richie, "Standard Reference Materials[®] Human Mitochondrial DNA-Amplification and Sequencing Standard Reference Materials-SRM 2392 and SRM 2392-I."
 27. "INSTRUCTION MANUAL NEBNext[®] Ultra[™] II DNA Library Prep Kit for Illumina[®]."
 28. U. H. Trivedi *et al.*, "Quality control of next-generation sequencing data without a reference," *Front. Genet.*, vol. 5, no. MAY, 2014, doi: 10.3389/fgene.2014.00111.
 29. P. Ewels, M. Magnusson, S. Lundin, and M. Källér, "MultiQC: Summarize analysis results for multiple tools and samples in a single report," *Bioinformatics*, vol. 32, no. 19, pp. 3047–3048, Oct. 2016, doi: 10.1093/bioinformatics/btw354.
 30. A. M. Bolger, M. Lohse, and B. Usadel, "Trimmomatic: A flexible trimmer for Illumina sequence data," *Bioinformatics*, vol. 30, no. 15, pp. 2114–2120, Aug. 2014, doi: 10.1093/bioinformatics/btu170.
 31. A. Auton *et al.*, "A global reference for human genetic variation," *Nature*, vol. 526, no. 7571. Nature Publishing Group, pp. 68–74, Sep. 30, 2015, doi: 10.1038/nature15393.
 32. H. Li and R. Durbin, "Fast and accurate short read alignment with Burrows-Wheeler transform," *Bioinformatics*, vol. 25, no. 14, pp. 1754–1760, Jul. 2009, doi: 10.1093/bioinformatics/btp324.
 33. K. Okonechnikov, A. Conesa, and F. García-Alcalde, "Qualimap 2: Advanced multi-sample quality control for high-throughput sequencing data," *Bioinformatics*, vol. 32, no. 2, pp. 292–294, Jan. 2016, doi: 10.1093/bioinformatics/btv566.
 34. "GATK Best Practices Mitochondrial Analysis – GATK." <https://gatk.broadinstitute.org/hc/en-us/community/posts/360056028651-GATK-Best-Practices-Mitochondrial-Analysis> (accessed May 05, 2021).
 35. R. Preste, R. Clima, and M. Attimonelli, "Human mitochondrial variant annotation with HmtNote," *bioRxiv*. bioRxiv, p. 600619, Apr. 10, 2019, doi: 10.1101/600619.
 36. R. Preste, O. Vitale, R. Clima, G. Gasparre, and M. Attimonelli, "Hmtvar: A new resource for human mitochondrial variations and pathogenicity data," *Nucleic Acids Res.*, vol. 47, no. D1, pp. D1202–D1210, Jan. 2019, doi: 10.1093/nar/gky1024.
 37. H. Weissensteiner *et al.*, "HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing," *Nucleic Acids Res.*, vol. 44, no. W1, pp. W58–W63, Jul. 2016, doi: 10.1093/nar/gkw233.
 38. K. Kiesler, "Sequencing and Standards for Characterization of the Mitochondrial Genome," 2018.

39. C. Lanave, S. Tommasi, G. Preparata, and C. Saccone, "Transition and transversion rate in the evolution of animal mitochondrial DNA," *BioSystems*, vol. 19, no. 4, pp. 273–283, 1986, doi: 10.1016/0303-2647(86)90004-3.
40. S. Jeon *et al.*, "Korean Genome Project: 1094 Korean personal genomes with clinical information," *Sci. Adv.*, vol. 6, no. 22, p. eaaz7835, May 2020, doi: 10.1126/sciadv.aaz7835.
41. J. Kim *et al.*, "KoVariome: Korean National Standard Reference Variome database of whole genomes with comprehensive SNV, indel, CNV, and SV analyses," *Sci. Rep.*, vol. 8, no. 1, Dec. 2018, doi: 10.1038/s41598-018-23837-x.
42. "TopVariants < MITOMAP < Foswiki."
43. "SearchAllele < Main < Foswiki." <https://www.mitomap.org/foswiki/bin/view/Main/SearchAllele> (accessed May 06, 2021).
44. "Origin and evolution of Native American mtDNA variation: a reappraisal - PubMed." <https://pubmed.ncbi.nlm.nih.gov/8808611/> (accessed Jun. 21, 2021).
45. M. Tanaka *et al.*, "Mitochondrial genome variation in Eastern Asia and the peopling of Japan," *Genome Res.*, vol. 14, no. 10 A, pp. 1832–1850, Oct. 2004, doi: 10.1101/gr.2286304.
46. K. Yamamoto *et al.*, "Genetic and phenotypic landscape of the mitochondrial genome in the Japanese population," *Commun. Biol.*, vol. 3, no. 1, pp. 1–11, Dec. 2020, doi: 10.1038/s42003-020-0812-9.
47. "Mutect2 – GATK." <https://gatk.broadinstitute.org/hc/en-us/articles/360042477952-Mutect2> (accessed Apr. 30, 2021).
48. "rs28358571 RefSNP Report - dbSNP - NCBI." <https://www.ncbi.nlm.nih.gov/snp/rs28358571> (accessed May 26, 2021).
49. J. H. McDermott *et al.*, "Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for the use of aminoglycosides based on *MT-RNR1* genotype," *Clin. Pharmacol. Ther.*, p. cpt.2309, May 2021, doi: 10.1002/cpt.2309.
50. L. Palenzuela *et al.*, "Does linezolid cause lactic acidosis by inhibiting mitochondrial protein synthesis?," *Clin. Infect. Dis.*, vol. 40, no. 12, 2005, doi: 10.1086/430441.
51. M. Narita, B. T. Tsuji, and V. L. Yu, "Linezolid-Associated Peripheral and Optic Neuropathy, Lactic Acidosis, and Serotonin Syndrome," *Pharmacotherapy*, vol. 27, no. 8, pp. 1189–1197, Aug. 2007, doi: 10.1592/phco.27.8.1189.
52. "rs3928306 RefSNP Report - dbSNP - NCBI." <https://www.ncbi.nlm.nih.gov/snp/rs3928306> (accessed May 06, 2021).
53. T. Hulgan *et al.*, "Mitochondrial DNA variation and changes in adiponectin and endothelial function in HIV-infected adults after antiretroviral therapy initiation," *AIDS Res. Hum. Retroviruses*, vol. 29, no. 10, pp. 1293–1299, Oct. 2013, doi: 10.1089/aid.2013.0079.
54. B. D. Greenberg, J. E. Newbold, and A. Sugino, "Intraspecific nucleotide sequence variability surrounding the origin of replication in human mitochondrial DNA," *Gene*, vol. 21, no. 1–2, pp. 33–49,

Jan. 1983, doi: 10.1016/0378-1119(83)90145-2.

55. K. S. Park *et al.*, "A mitochondrial DNA variant at position 16189 is associated with type 2 diabetes mellitus in Asians," *Diabetologia*, vol. 51, no. 4, pp. 602–608, Apr. 2008, doi: 10.1007/s00125-008-0933-z.
56. B. Kumar *et al.*, "Association of mitochondrial copy number variation and T16189C polymorphism with colorectal cancer in North Indian population," *Tumor Biol.*, vol. 39, no. 11, Nov. 2017, doi: 10.1177/1010428317740296.
57. E. Saldaña-Rivera, M. J. Careaga-Castilla, G. D. Olvera-Cárdenas, E. Pérez-Soto, and V. Sánchez-Monroy, "Mitochondrial T16189C Polymorphism Is Associated with Metabolic Syndrome in the Mexican Population," *Dis. Markers*, vol. 2018, 2018, doi: 10.1155/2018/3981315.
58. L. Jamali, M. M. Banoei, E. Khalili, S. Dadgar, and M. Houshmand, "Association of genetic variations in the mitochondrial d-loop with β -thalassemia," *Mitochondrial DNA Part A DNA Mapping, Seq. Anal.*, vol. 27, no. 3, pp. 1–4, 2016, doi: 10.3109/19401736.2014.958730.
59. C. W. Liou *et al.*, "Association between a common mitochondrial DNA D-loop polycytosine variant and alteration of mitochondrial copy number in human peripheral blood cells," *J. Med. Genet.*, vol. 47, no. 11, pp. 723–728, Nov. 2010, doi: 10.1136/jmg.2010.077552.
60. A. Ramachandran, R. G. J. Visschers, L. Duan, J. Y. Akakpo, and H. Jaeschke, "Mitochondrial dysfunction as a mechanism of drug-induced hepatotoxicity: current understanding and future perspectives," *J. Clin. Transl. Res.*, vol. 4, no. 1, p. 75, 2018, doi: 10.18053/JCTRES.04.201801.005.
61. Y. He *et al.*, "Heteroplasmic mitochondrial DNA mutations in normal and tumour cells," *Nature*, vol. 464, no. 7288, pp. 610–614, Mar. 2010, doi: 10.1038/nature08802.
62. W. W. Hauswirth and D. A. Clayton, "Length heterogeneity of a conserved displacement-loop sequence in human mitochondrial DNA," *Nucleic Acids Res.*, vol. 13, no. 22, pp. 8093–8104, Nov. 1985, doi: 10.1093/nar/13.22.8093.
63. N. Howell and C. B. Smejkal, "Persistent heteroplasmy of a mutation in the human mtDNA control region: Hypermutation as an apparent consequence of simple-repeat expansion/contraction," *Am. J. Hum. Genet.*, vol. 66, no. 5, pp. 1589–1598, 2000, doi: 10.1086/302910.
64. H. Zhao, J. Shen, L. Medico, M. Platek, and C. B. Ambrosone, "Length heteroplasmies in human mitochondrial DNA control regions and breast cancer risk," *Int. J. Mol. Epidemiol. Genet.*, vol. 1, no. 3, pp. 184–192, 2010.
65. G. Avital *et al.*, "Mitochondrial DNA heteroplasmy in diabetes and normal adults: role of acquired and inherited mutational patterns in twins," *Hum. Mol. Genet.*, vol. 21, no. 19, p. 4214, Oct. 2012, doi: 10.1093/HMG/DDS245.
66. S. Anderson *et al.*, "Sequence and organization of the human mitochondrial genome," *Nature*, vol. 290, no. 5806, pp. 457–465, 1981, doi: 10.1038/290457a0.
67. T. Nguyen and A. Jeyakumar, "Genetic susceptibility to aminoglycoside ototoxicity," *Int. J. Pediatr. Otorhinolaryngol.*, vol. 120, pp. 15–19, May 2019, doi: 10.1016/j.ijporl.2019.02.002.

68. G. Garrabou *et al.*, "Influence of mitochondrial genetics on the mitochondrial toxicity of linezolid in blood cells and skin nerve fibers," *Antimicrob. Agents Chemother.*, vol. 61, no. 9, Sep. 2017, doi: 10.1128/AAC.00542-17.
69. A. Mansouri, C. H. Gattolliat, and T. Asselah, "Mitochondrial Dysfunction and Signaling in Chronic Liver Diseases," *Gastroenterology*, vol. 155, no. 3, pp. 629–647, 2018, doi: 10.1053/j.gastro.2018.06.083.
70. L.-N. Lee *et al.*, "Mitochondrial DNA Variants in Patients with Liver Injury Due to Anti-Tuberculosis Drugs," *J. Clin. Med.*, vol. 8, no. 8, p. 1207, Aug. 2019, doi: 10.3390/jcm8081207.
71. P. M. Smith *et al.*, "The role of the mitochondrial ribosome in human disease: Searching for mutations in 12s mitochondrial rRNA with high disruptive potential," *Hum. Mol. Genet.*, vol. 23, no. 4, pp. 949–967, Feb. 2014, doi: 10.1093/hmg/ddt490.
72. J. L. Elson *et al.*, "The presence of highly disruptive 16S rRNA mutations in clinical samples indicates a wider role for mutations of the mitochondrial ribosome in human disease," *Mitochondrion*, vol. 25, pp. 17–27, Nov. 2015, doi: 10.1016/j.mito.2015.08.004.
73. C. C. Luca, B. L. Lam, and C. T. Moraes, "Erythromycin as a potential precipitating agent in the onset of Leber's hereditary optic neuropathy," *Mitochondrion*, vol. 4, no. 1, pp. 31–36, 2004, doi: 10.1016/j.mito.2004.05.002.
74. P. Yu-Wai-Man, P. G. Griffiths, and P. F. Chinnery, "Mitochondrial optic neuropathies - Disease mechanisms and therapeutic strategies," *Progress in Retinal and Eye Research*, vol. 30, no. 2. Elsevier, pp. 81–114, Mar. 2011, doi: 10.1016/j.preteyeres.2010.11.002.

Figures

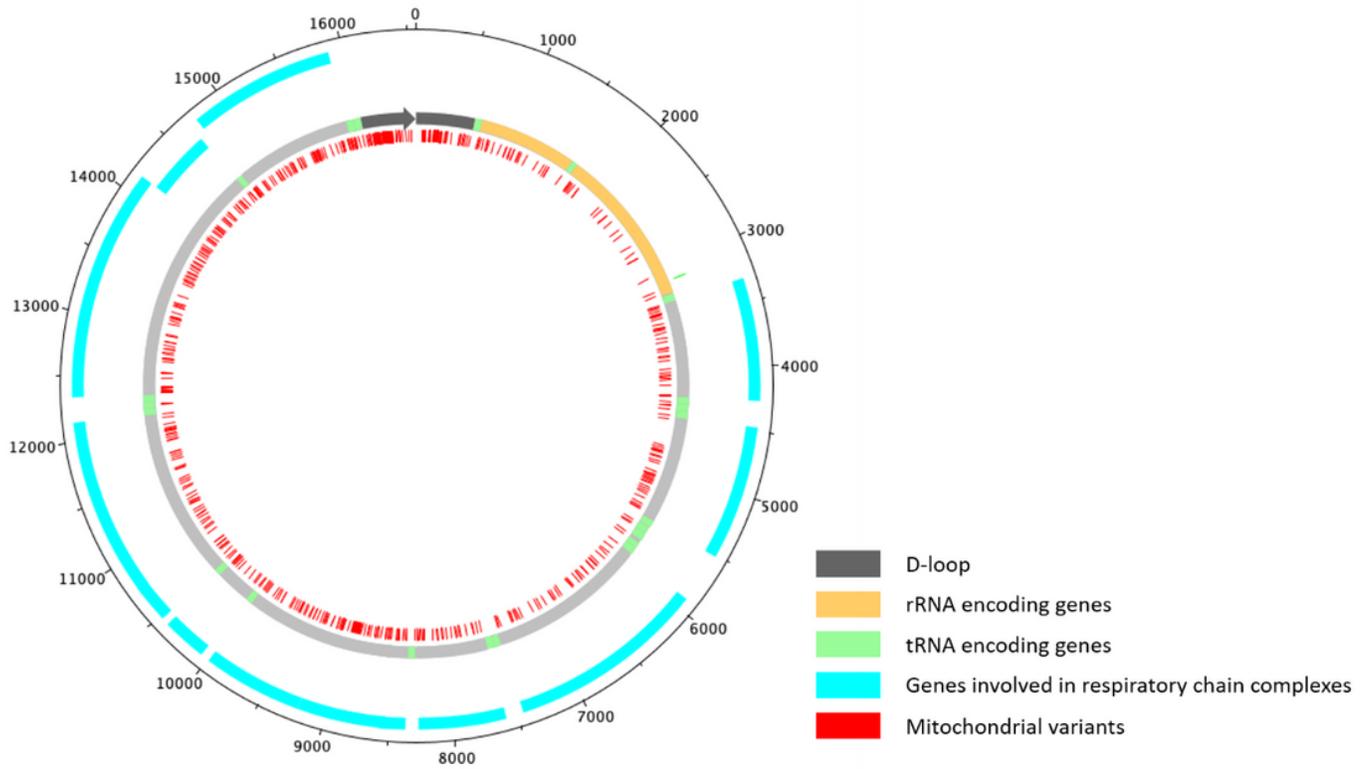


Figure 1

Distribution of variants in mitochondrial genomes in this study. Mitochondrial variants (red) in tRNA genes (green), rRNA genes (orange) and genes involved in mitochondrial respiratory chain complexes (blue), in the Korean population in the present study (n = 118) according to the Revised Cambridge Reference Sequence (NC_012920.1). The mitochondrial genome was generated by using DNAPlotter [40].

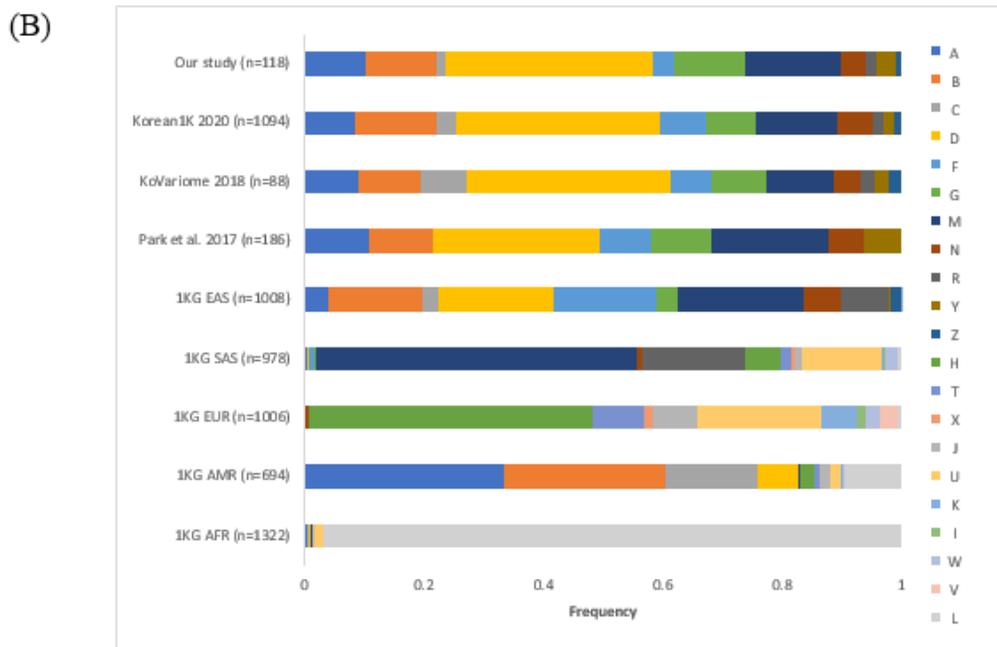
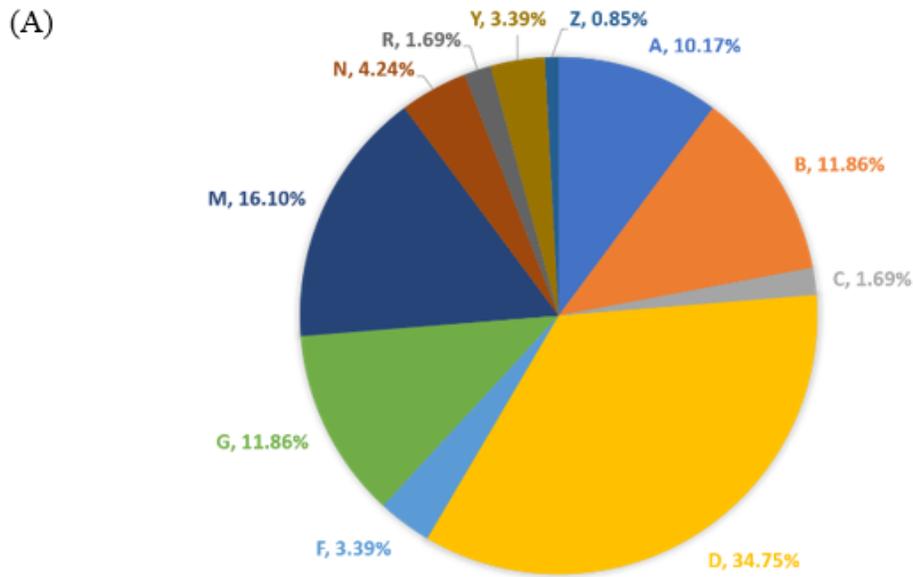


Figure 2

The haplogroup of mitochondrial genomes for Korean population (A) Distribution of haplogroups in the Korean population in the present study. (B) Stacked bar plots of the frequencies of haplogroups in the Korean population and five super-populations from the 1000 Genomes Project (1KG).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AdditionalfilesVinhHoaetal.docx](#)