# STUDY ON FREQUENCY OF MUTATIONS FOR HVII REGION OF MITOCHONDRIAL DNA OF VIETNAMESE POPULATION SAMPLES AND APPLICATIONS IN FORENSIC IDENTIFICATION

**NGUYEN THI NGOC ANH MD., MSc., DO THI XAO MAI MD., MSc., NGUYEN TIEN THANH MD., NGUYEN THI LE DAN Eng., PHAM VAN BINH BSc., NGUYEN TIEN TUNG Tech.**

***Military Institute of Forensic Medicine***

**Scientific reviewer: (1) NGUYEN VAN LOI MD., Ph.D.**

**(2) BUI TIEN SY MD., Ph.D.**

**ABSTRACT:** *There are many mutations in the hypervariable II region in mitochondrial DNA. Determining the frequency of mitochondrial HVII region mutations improves the reliability of identifying human remains using the mitochondrial DNA technique. Our study is mutations in the HVII region in mitochondrial DNA of 300 blood samples that were collected from 300 unrelated Vietnamese individual.* ***Results:*** *Determine 31 mutations in mitochondrial HVII region, in there, frequencies of 24 mutations are low 10% (77.41%), frequencies of mutations are 10-30% (16.14%), frequencies of 2 mutations are high 30% (6.45%).*

**Keywords:** DNA analysis, mitochondrial DNA, frequency.

Responsible for the content: Nguyen Thi Ngoc Anh MD., MSc., Email: Duck19399@gmail.com

Date of receipt: January 15, 2022; invite scientific reviewer: February, 2022; accepted: February 24, 2022.

## 1. INTRODUCTION.

Forensic human identification has been demonstrated in the world as well as in Vietnam. Along with the development of morphological examination techniques, the technique of DNA analysis is an important step in forming a process of human scientific identification with objective, and highly accurate assessment.

The DNA analysis techniques for human identification in forensic medicine include nuclear DNA typing and mitochondrial DNA (mtDNA) typing. Due to the circular nature of the mitochondrial DNA genome, mtDNA typing is attempted on samples for which nuclear DNA typing is not likely to be successful such as bones, teeth and other samples that have been degraded over time. The basis of the application of the mtlDNA typing is to use point mutations in the D-loop region which contains three hypervariable regions: hypervariable region I (HVI), hypervariable region II (HVII), and hypervariable region III (HVIII). Currently, in Vietnam, laboratories are carrying out the identification of two hypervariable regions, HVI, HVII - mitochondrial DNA, because the identification value of point mutations of these two regions is high.

Determining the genetic characteristics of mitochondrial DNA for each population is very important in forensic identification. When analyzing and interpreting the results of mitochondrial DNA sequencing, it is necessary to consider the frequency distribution of point mutations of each population. The point mutations in mitochondrial DNA supervariable regions have different frequencies for each population. When determining the frequency of point mutations in mitochondrial DNA supervariable regions, it will evaluate the level of population specificity and increase the reliability of the results of the identification of remains.

Stemming from the above reasons, we carried out this study to determine the population for HVII region in Vietnamese mitochondrial genome - one of the two hypervariable regions with many point mutations occurring.

## 2. SUBJECTS AND METHODS.

### 2.1. Subject:

- Subject: the point mutations in the HVII region - mitochondrial DNA from 300 sequences of 300 blood samples that were collected from 300 unrelated Vietnamese individual.

- Study time: from 12/2019 to 12/2020.

- Study place: Department of Biological Testing, Military Institute of Forensic medicine.

### 2.2. Methods:

- Study design: cross-sectional description, retrospective combined with a prospective study.

- Method of sample selection: 300 standard sequences were selected from 300 blood collection cards from 300 unrelated Vietnamese individual.

- Content of the study:

+ Extract DNA from 300 blood samples using Chelex.

+ Performed polymerase chain reaction (PCR) set-up for the HVII region - mtDNA. Then, electrophoresis through an agarose gel containing ethidium bromide to evalute the quanlity of the PCR product and ExoSAP-It is used for purification of the PCR product.

+ Performed sequencing reaction. The sequenced product is then purified and the resulting product is loaded on an Applied Biosystems 3500xL Genetic Analyzer.

+ Using sequencing analysis sorfware 6 and Sequencher sorfware 5.4.5 to analysis and interpretation of mtDNA sequencing data,

- Data processing: using Excel 2007 software to calculate the frequency of mutations in the HVII region - Vietnamese mtDNA.

## 3. RESULTS

### 3.1. Results of amplification and sequencing of the HVII region - mtDNA:

- Complete amplification and sequencing of the HVII region - mtDNA of 300 study samples.

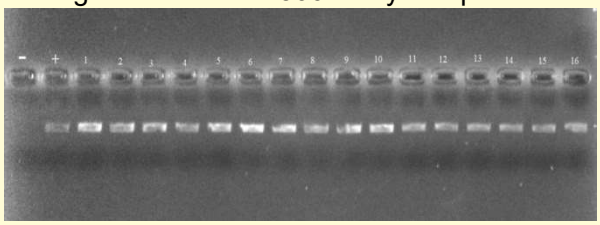
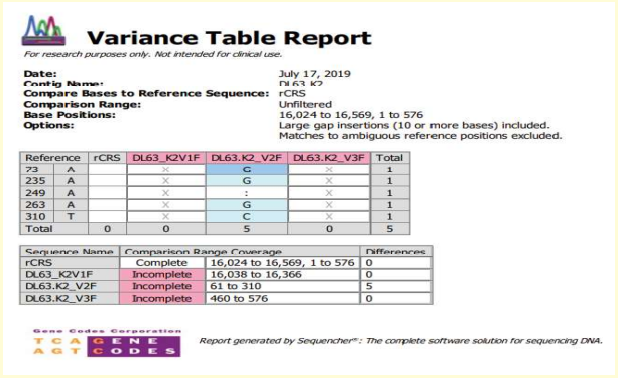


*Figure 1. Results of electrophoresis test of amplification reaction of the HVII region.*

The results of electrophoresis testing of amplified products on 1% agarose gel showed that: negative control samples (symbol -) did not appear on the bands; 16 blood samples (symbols from 1-16) showed clear, dark, and bright bands, the same size as the positive control samples (symbol +).

- The region between positions 73 and 340 in HVII were completely sequenced, analyzed.

*Figure 2. Results of sequencing analysis of HVII region - Mitochondrial DNA by software sequencing analysis software 6.*



Variance Table Report

For research purposes only. Not intended for clinical use.

Date: July 17, 2019
Contig Name: DL63_K2
Compare Bases to Reference Sequence: rCRS
Comparison Range: Unfiltered
Base Positions: 16,024 to 16,569, 1 to 576
Options: Large gap insertions (10 or more bases) included. Matches to ambiguous reference positions excluded.

| Reference | | rCRS | DL63_K2V1F | DL63_K2_V2F | DL63_K2_V3F | Total |
|---|---|---|---|---|---|---|
| 73 | A | | × | G | × | 1 |
| 235 | A | | × | G | × | 1 |
| 249 | A | | × | : | × | 1 |
| 263 | A | | × | G | × | 1 |
| 310 | T | | × | C | × | 1 |
| Total | | 0 | 0 | 5 | 0 | 5 |

| Sequence Name | Comparison Range Coverage | | Differences |
|---|---|---|---|
| rCRS | Complete | 16,024 to 16,569, 1 to 576 | 0 |
| DL63_K2V1F | Incomplete | 16,038 to 16,366 | 0 |
| DL63_K2_V2F | Incomplete | 61 to 310 | 5 |
| DL63_K2_V3F | Incomplete | 460 to 576 | 0 |

Gene Codes Corporation

Report generated by Sequencher™: The complete software solution for sequencing DNA.

*Figure 3. Results of sequencing and analysis of point mutations in HVII region-mitochondrial DNA by software Sequencher 5.4.5.*

The results of sequencing the HVII region - mitochondrial DNA showed that the signal peaks were clear and undisturbed.

### 3.2. Characteristics of nucleotide changes in the HVII region - mtDNA:

Table 1. Characteristics of nucleotide changes in the HVII region - mitochondrial DNA

| Type of changes | | Total change | Rate % |
|---|---|---|---|
| Transition | A→G | 652 | 68.13 |
| | G→A | 19 | 2.00 |
| | C→T | 70 | 7.31 |
| | T→C | 183 | 19.12 |
| | Total | 924 | 96.56 |
| Tranversion | A→C | 01 | 0.10 |
| | C→A | 10 | 1.04 |
| | Total | 11 | 1.15 |
| Deletion of nucleotide | | 11 | 1.15 |
| Total | | 957 | 100 |

There were 957 changes in the HVII region - mtDNA in 300 studied samples. In which, the main changes were substitutions of nucleotides of the same base (purines instead of pyrimidine and vice versa - transition) with the rate of 96.56% (A → G transformation had the highest rate, accounted for 68.13%). Transversion of nucleotide substitutions accounted for 1.15%. Changes in loss nucleotide accounted for 1.15%.

### 3.3. The frequency of appearing of mutations in the HVII region - Vietnamese mtDNA:

There were 31 positions with changes in the HVII region - mtDNA compared with revised Cambridge Reference Sequence (rCRS) and divided into three groups: the group with low frequency of changes less than 10% (24 positions, accounted for 77.41%), the group with a high frequency of changes from 10-30% (5 sites, accounted for 16.14%), the group

with very high frequency over 30% (2 locations, accounted for 6.45%).

Table 2. The frequency of mutations in the HVII region - mtDNA.

| Frequency of changes | Point mutations | Total | Rate % |
|---|---|---|---|
| Less than 10% | 249Del, A93G, T125C, T127C, C128T, G143A, T146A, A153G, A183G, G185A, T195C, T196C, C198T, A200G, T204C, G207A, A234G, A235G, A237G, A281G, T282C, A302C, T310C, T318C | 24 | 77.41 |
| From 10-30% | T146C, C150T, T152C, T199C, A210G, | 05 | 16.14 |
| Over 30% | A73G, A263G | 02 | 6.45 |
| Total | | 31 | 100 |

**4. DISCUSSION.**

**4.1. Regarding the complete technique of analyzing genetic polymorphism in the HVII region - mitochondrial DNA:**

The electrophoresis image of the PCR product of the HVII region (Figure 1) showed that the quality of the DNA extracted from the blood samples was very well, ensuring that the implementing techniques in the HVII region - mtDNA able to conduct. The image of the result of sequencing in the HVII region-mitochondrial DNA (Figure 2, Figure 3) showed that the gene sequencing analysis technique ensured quality.

Thus, we can analyze the nucleotide sequencing in the HVII region of the mitochondrial genome used in forensic identification generally, and the missing martyr remains by Project 515 particularly. The analysis of DNA from mitochondria opened up the possibility of identification much easier because the number of copies of mitochondrial DNA was hundreds or even thousands of times more than DNA in the nucleus. Another advantage was that the DNA in the nucleus in the children received 50% from the father, 50% from the mother, so when analyzing the generation of children, it was more likely due to different combinations during the formation of the zygote. On the other hand, there were cases of identification assessment, the generation of parents who need to be identified (especially for cases of identifying martyrs' remains) was no more, the comparison sample was only collected from brothers/sisters of martyrs and later generations.

Analysis of nuclear DNA was difficult if a sample was taken from siblings or cousins to identify someone. With mitochondrial DNA, all siblings from the same mother had the same DNA and were similar to their mother. These were features that made it easier for us to identify. In the past, to identify someone, people relied mainly on morphology, ethnicity, and some physical and chemical characteristics. With the development of genetics, the DNA test used in identification helped this work more exact. With traces, the remaining specimens were few and were affected by external factors such as too long time, impacts of the environment were too strong (fire, earthquake...), the remaining traces were limited and destroyed by environmental impact. Analysing based on the mitochondrial DNA with a high accuracy rate because the DNA in people of different origins was different.

**4.2. Number of mutation positions in the HVII region - mtDNA:**

This study was carried out in 300 Vietnamese people and found 31 mutations in the HVII region - mtDNA. While the study of Phan Van Chi (2006) in 44 Vietnamese people, there were 12 mutation positions in the HVII region-mtDNA [1]. The difference may be because the sample size in our study was more than that of author Phan Van Chi. Therefore, the number of variable positions encountered more than Phan Van Chi was appropriate. The number of mutations that were common in the HVII region - Vietnamese mtDNA showed that, using DNA analysis of the HVII region - mtDNA was valuable in forensic identification for Vietnamese people, especially in the identification of perennial remains.

**4.3. Genetic variation type:**

- Compare types of genetic variation in the HVII region:

Table 3. Comparison of genetic variation type in HVII region - mtDNA of Vietnamese (300 samples) with Japanese (162 samples) [2] and Malaysian (195 samples) [3].

| Variation type | Japan | Malaysia | Vietnam |
|---|---|---|---|
| Transition | 92.92% | 92.16% | 96.56% |
| Transversion | 7.08% | 7.84% | 1.15%. |
| Deletion | 13.58 | 14.36 | 1.15% |

The study by Budowle et al (1999) in Japanese [2] and Lian Lay Hoong et al (2005) in Malaysian people [3] showed that the type of variation in the HVII region - mtDNA was mainly transition. Thus, the substitution of nucleotides of the same origin in our study accounted for a high rate; similar to the Japanese and Malaysians. Changes in nucleotide loss in this study encountered 11 cases in the HVII

region (accounted for 1.15%); lower than studies in Japan and Malaysia.

Analysis of the HVII region - Vietnamese mtDNA showed that there were many changes compared to the original sequence first published by Anderson et al in 1981, at the Cambridge University and the additional sequence by Andrews and colleagues edited the CRS sequence (now known as the Revised Cambridge reference sequence: rCRS). These specific transformations were the basis for forensic identification.

**4.4. Frequency of mutations in the HVII region - mtDNA of Vietnamese population:**

The study by Phan Van Chi in 2006 in 44 Vietnamese people, found that there were 12 point mutations in the HVII region. Some new polymorphisms appearing in the HVII region were G94A, T216G, C269G, T279G, C330G. The common mutations in the region were: 73G (20.45%), C269G (13.63%) [1]. The results of this study found 31 mutation locations in the HVII region (from positions 73 to 340). The changes in the HVII region found in the study of Phan Van Chi, but not in our study, were G94A, T216G, C269G, T279G, and C330G. These were the point mutations that Phan Van Chi mentioned as newly discovered mutations. The common point mutations in our study and also mentioned by Phan Van Chi as common mutations were: A73G (96.67%).

In addition, this study also encountered other point mutations with a high rate: T146C (12.67%), T152C (12.67%), T199C (16.67%), A210G (13.67%)., A263G (97.67%); low frequency point mutations: G93A (0.67%), T125C (1.0%), T127C (1.0%), C128T (0.33%), G143A (2.0%), T146A (3.33%, A153G (1.67%), A183G (3.33%), T195C (4.0%), T196C (0.33%), C198T (0.33%), A200G (1, 0%), T204C (5.033%), G207A (0.033%), A234G (1.0%), A235G (0.67%), A237G (0.67%), A281G (0.33%), T282C (0.67%), A302C (0.33%), T310C (4.67%), T318C (2.0%). The location with 2 types of variation in the HVII region was: T146 (A = 3.33%, C = 12.67%). There was one location where nucleotide loss changed: 249Del (3.67%).

**4.5. Some practical results:**

We used the results of the study to increase the reliability in forensic identification and identification of martyrs' remains at the request of the Department of National Devotees (Ministry of Labor - Invalids and Social Affairs) in 2019. For example:

- Case 1: mutations in the HVII region - mtDNA in the case of forensic identification with code HT69.19. According to our study results, in the above case, the point mutations A73G, A263G were the mutations with very high frequency (A73G - 96.67%; A263G - 97.67%) in the Vietnamese population. The nucleotide loss mutation at positions A249Del was a very low-frequency mutation (3.67%) in the Vietnamese population.

| Mutation location | Sequence rCRS | Sample of remains (HT69.19_1) | Sample of relatives (HT69.19_2) |
|---|---|---|---|
| 73 | A | G | G |
| 150 | C | T | T |
| 249 | A | Del | Del |
| 263 | A | G | G |

- Case 2:mutations in the HVII region - mtDNA in the case of identification of martyrs' remains at Lot E, Martyrs cemetery in Dak Lak province. According to our study results, in this case, the point mutations A73G, A263G were the mutations with very high frequency (A73G - 96.67%; A263G - 97.67%); T310C was a very low-frequency mutation (4.67%) in the HVII region of the mitochondrial genome of the Vietnamese population.

| Mutation location | Sequence rCRS | Sample of remains (DL23_Q) | Sample of relatives (DL09_K) |
|---|---|---|---|
| 73 | A | G | G |
| 146 | T | C | C |
| 263 | A | G | G |
| 310 | T | C | C |

**5. CONCLUSION.**

Data are provided on HVII – mtNDA sequences of 300 Vietnamese individuals, found that there were 31 mutations: 24 mutations with low frequency (<10%); 5 mutations with a frequency of 10-30%; 2 mutations with a variable frequency of over 30%.

Through this study, we recommended: continuing the study with many different populations to include the Vietnamese mitochondrial genome data; carrying out research with the analytical method to evaluate the stability of D-loop region t mutations across generations in a family.

**REFERENCES:**

1. Phan Van Chi et al. (2006), *Study on decoding mitochondrial genomes of Vietnamese ethnic groups and application orientation*, State-level project code KC 04-25.

2. Budowle B, Wilson M.R, DiZinno et al (1999), *Mitochondrial DNA regions HV1 and HV1I population data*, Forensic Science International l103: 23-35.

3. Lian Lay Hoong, Koh Chong Lek (2005), "Mitochondrial DNA types in Malaysia Asia Pacific Journal of Molecular Biology and Biotechnology", *AsPac J. Mol. Biol. Biotechnol.* Vol.13(2) 79-85. ❑