The Database System of Mitochondrial Haplogroups in Populations Living in Northern Thailand

Athicha Chumphukhiao[1], Supakit Khachaananda[2], Hataichanoke Niamsup[1], Phatcharin Mahawong[2], Pimpinun Punyathi[2], Nitaya Khamkomkai[2]

[1] Division of Biochemistry and Biochemical Innovation, Department of Chemistry, Faculty of Science, Chiang Mai University
[2] Department of Forensic Medicide, Faculty of Medicine, Chiang Mai University

Email address: Athicha_C@cmu.ac.th

ABSTRACT

This study aimed to develop a database and search system for mitochondrial haplogroups in the Thai population living in Northern Thailand. Retrospective mitochondrial DNA (mtDNA) profiles from 200 individuals were collected from medical records. Detected polymorphisms were classified into haplogroups using the EMPOP database. The data were subsequently developed into a polymorphism-search platform using Microsoft Excel. The results showed a haplotype diversity of 0.99206, and mtDNA was classified into 11 haplogroups, predominantly F (28.5%), M (27.0%), and C (12.0%). A Microsoft Excel-based comparative search system was developed, allowing users to input polymorphism of mtDNA of new samples to identify matches within a reference database of 200 individuals and automatically report the matching frequency. Analysis identified 143 unique haplotypes, demonstrating the database's effectiveness in individual discrimination. The findings indicate that the generated data and platform are suitable for further development into a database system. The results can be utilized to establish a mitochondrial DNA database for the laboratory, which would support the analysis of maternal lineage relationships and enhance the reliability of genetic testing results.

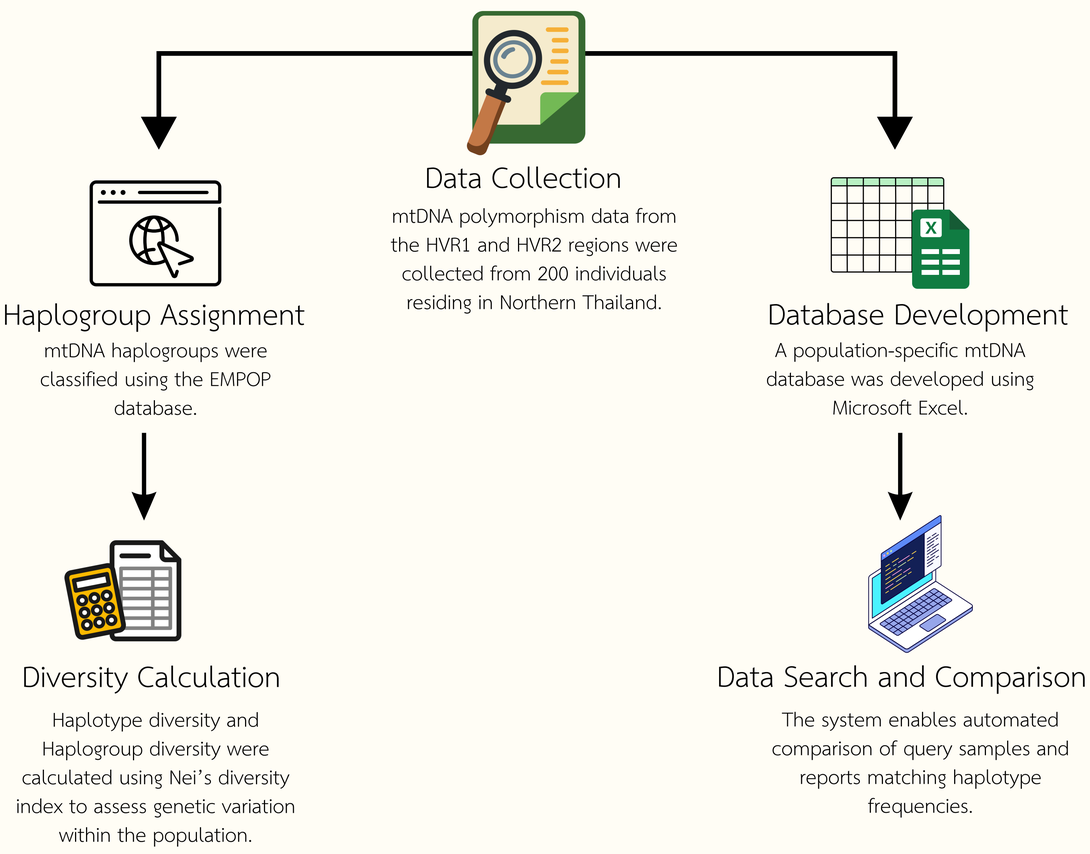
INTRODUCTION

Due to a high copy number of DNA per cell and maternal inheritance pattern, mtDNA is a valuable tool for maternal lineage analysis. Also, it plays an essential role in forensic genetics particularly in case with highly degraded samples or insufficient nuclear DNA. Northern Thailand exhibits substantial ethnic diversity and a complex settlement history; however, population-specific reference databases remain limited. This poses a challenge for forensic laboratories in interpreting newly obtained genetic profiles and assessing evidential weight due to the lack of a local system capable of rapidly retrieving and reporting profile frequencies. To address this limitation, this study established a mitochondrial DNA (mtDNA) database from 200 individuals representing the Northern Thai population and developed a comparative search platform using Microsoft Excel. The system enables systematic data management and automatically reports matching frequencies, facilitating estimation of haplotype occurrence in the population. This platform is expected to enhance the reliability of forensic genetic interpretation in Northern Thailand.

OBJECTIVES

1. To characterize mitochondrial DNA (mtDNA) genetic diversity and haplogroup distribution in the northern Thai population.
2. To establish a reference database of HVR1/HVR2 sequences (n = 200) for efficient genetic profile comparison.
3. To enhance forensic maternal lineage identification through mtDNA haplotype frequency analysis.

METHODOLOGY



ACKNOWLEDGEMENTS

- The author would like to express sincere gratitude to the Department of Forensic Medicine, Faculty of Medicine, Chiang Mai University, for providing the facilities, equipment, and data for this project. The author also gratefully acknowledges the Division of Biochemistry and Biochemical Innovation, Department of Chemistry, Faculty of Science, Chiang Mai University, for providing knowledge and valuable skill development throughout this study.

RESULTS

Analysis of mtDNA sequences from the hypervariable regions HVR1and HVR2 was performed on 200 individuals residing in NorthernThailand.

Table 1. Example dataset showing HVR1 and HVR2 polymorphisms and haplogroup assignment.

Sample	HVR1 1	HVR1 2	HVR1 3	...	HVR2 1	HVR2 2	HVR2 3	...	Haplogroup	
S1	16092C	16223T	16362C	...	73G	94A	146C	...	D	D4e1a
S2	16223T	16288C	16298C	...	73G	249DEL	263G	...	C	C5
S3	16173T	16209C	16223T	...	73G	146C	198T	...	M	M11a2
S4	16108T	16129A	16162G	...	73G	249DEL	263G	...	F	F1a1a

Haplogroups were classified using the EMPOP database. A total of 11 mtDNA haplogroups were identified. The three most frequent haplogroups were F (28.5%), M (27.0%), and C (12.0%). The remaining haplogroups (D, A, R, N, B, L, G, and U) collectively accounted for 32.5% of the population.

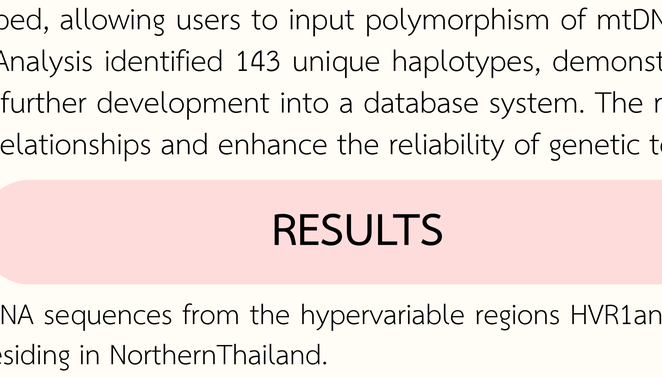


Figure 1. Radiation of human mitochondrial DNA (mtDNA) as humans migrated out of Africa to Eurasia, Australia, and the Americas (adapted from Wallace, 2005).

The haplotype diversity was 0.99206, indicating high genetic diversity, with 142 unique haplotypes identified among 200 samples. The haplogroup diversity was 0.8171, demonstrating substantial variation in haplogroup distribution. A Microsoft Excel–based database was developed to enable automated comparison of mtDNA polymorphisms and reporting of haplotype match frequencies.

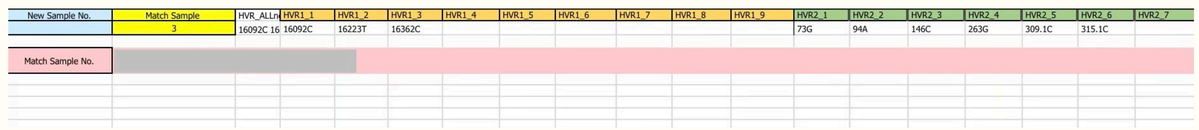


Figure 2. Example of the developed Microsoft Excel–based database used for comparing HVR1 and HVR2 polymorphisms and identifying matched samples.

CONCLUSION

Overall, the high haplotype diversity and the large number of unique haplotypes demonstrate substantial genetic variation among individuals in Northern Thailand and indicate strong discriminatory power of mtDNA markers. The developed database and automated comparison platform provide a practical tool for rapid identification of haplotype matches and frequency estimation. This system establishes an important population-specific reference resource and strengthens the reliability of mtDNA analysis for forensic identification and maternal lineage investigations in Northern Thailand.

REFERENCES

1. Nei, M. (1973). Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences of the United States of America, 70(12), 3321–3323.
2. Pakendorf, B., & Stoneking, M. (2005). Mitochondrial DNA and human evolution. Annual Review of Genomics and Human Genetics, 6, 165–183.
3. Wallace, D. C. (2005). A mitochondrial bioenergetic etiology of disease. Journal of Clinical Investigation, 115(9), 2319–2325.
4. Parson, W., & Dür, A. (2007). EMPOP — A forensic mtDNA database. Forensic Science International: Genetics, 1(2), 88–92.
5. Kutanan, W., et al. (2018). Reconstructing the population history of Thailand using mitochondrial DNA. Scientific Reports, 8, 16354.