Genetic polymorphism analysis of mitochondrial DNA from Chinese Guangdong Liannan Yao ethnic group

Shiyu Sun[1], Jinpeng Chen[2], Mingqi Zhang[2], Huang Huang[2], Xihong Lin[2], Weian Du[3], Huilin Sun[2*]

[1] School of the Third Clinical Medicine, Guangzhou University of Chinese Medicine, Guangzhou, 510006, China.

[2]Department of Endocrinology, The First Affiliated Hospital of Guangdong Pharmaceutical University, Guangzhou, 510515, China.

[3]HOMY GeneTech Incorporation, Foshan, 518000, China.

*Corresponding author:
Dr Huilin Sun, E-mail address:sun-hui-lin@126.com.

Abstract

Background Genetic polymorphism and haplotype distribution characteristics analysis of mitochondrial DNA in Chinese Guangdong Liannan Yao group was conducted in this study, to provide genetic basis for tracing the origin and historical migration of Liannan Yao people.
Results 46 mutation sites were found, and among which single nucleotide transition was the most commonly observed variant (86.17%). Multiple (sub)haplogroups were detected in Liannan Yao ethnic group, among which haplogroup D was the most common haplogroup (29.80%), and the least were C and Y(0.48% respectively). **Conclusions** The Liannan Yao population had the commonalities of the ethnic groups in southern China, but it was significantly different from other Chinese ethnic populations. The present results revealed that Liannan Yao ethnic group was genetically closer related to Fujian She ethnic population, Yunnan Yao population, and Hunnan Miao population. The data enriched the Chinese mtDNA database and provided a reference for forensic identification and screening for potential pathogenic mutations.
Keywords Forensic genetics; Genetic polymorphism; Haplogroups; mtDNA; Liannan Yao

Background

The genetic analysis of human mitochondrial DNA(mtDNA) is a very important part of the genetic research of the global population. Due to its molecular characteristics, genetic patterns, specific polymorphisms, rapid evolution speed and low recombination rate, highly variable mtDNA fragments have become a research hotspot in related genetics.[1-5]Genetic analysis of mtDNA is not only widely used in forensic medicine [6-8], but also in anthropological research because it could indicate the genetic migration and evolution of the maternal population[9-10]. The researchers draft a phylogenetic tree of human mtDNA sequences referring to mutations of related

genomes in mtDNA, and classified them into different haplotypes and haplogroups. Migration history of maternal ancestors of modern populations can be deduced by exploring the distribution of mtDNA haplogroups of different nationalities in different regions of the world[11-13]. The haplogroups named after letters composed the human mtDNA phylogenetic tree, the root of the tree represented the common maternal ancestry of all humans, which was recorded as "*". Haplogroup “L” was unique to Africa, which indicated that modern humans originated in Africa. Haplogroup “L3” branched into three large groups “M”, “N”, and “R” (a branch of “N”), which contained all the variations observed outside Africa.

The Yao nationality, one of the 56 ethnic groups in China, distributed most widely in southern China, including Guizhou, Hunan, Yunnan, Guangdong and Guangxi provinces. According to the 6th population census in China (www.stats.gov.cn), Yao ethnic group has a population of 2,796,003. The Yao Autonomous County of Liannan, a small county town in northwest Guangdong province, has its own language, culture and genetic characteristics, because of less intermarriage with other groups.[14] Therefore, the Liannan Yao nationality’s genetic analysis should be the most valuable data among the previous Yao nationality’s data, but its genetic polymorphism has not been reported.

Here we used an Expressmarker mtDNA-SNP 60 kit ((AGCU ScienTech Incorporation, Jiangsu, Wuxi) to detect 209 samples collected from unrelated Liannan Yao individuals, analyze the genetic polymorphism of 60 mtDNA loci, reveal the genetic relationships between Liannan Yao group and the reference populations, as well as provide useful information for the maternal lineage study and migration history study of Chinese Liannan Yao group.

Results

Allele frequencies of the 60 mtDNA loci detected in Chinese Liannan Yao ethnic group were listed in Table 1. As for the 58 selected mtDNA SNP loci of Liannan Yao ethnic group, the most common polymorphism was single nucleotide transition (79.31%), followed by single nucleotide transversion (nt5178, nt7196, nt9824, nt13928 variants, 6.89%). At nt9824 locus, single nucleotide transition and transversion were simultaneously observed (A/T/C). While at loci nt1719, nt2706, nt3348, nt4491, nt7600, nt9123, nt9477, nt11251, nt15784, no polymorphisms were detected.

MtDNA haplogroups and haplotypes based on polymorphisms of the 60 mtDNA loci in Liannan Yao ethnic group were presented in Table 2 and Fig. 1. Fifty-one polymorphic loci (excluding nt1719, nt2706, nt3348, nt4491, nt7600, nt9123, nt9477, nt11251 and nt15784 loci) defined 13 haplogroups and 50 haplotypes. Among them, Haplogroup M was the most common haplogroup (55.02%) in Liannan Yao group, while the most common haplotype was D4e(19.14%).

Among the total 50 haplotypes, 22 of them were observed for only once, 11 for twice, and 17 for three times or more. The frequency of each haplotype observed in Liannan Yao ethic group was not high but the diversity of haplogroups was rich.

The values of the random match probability(RMP), the discrimination power(DP), and variability of the mtDNA haplotype in the Liannan Yao population was 0.074334, 0.9225665, and 0.9301159 respectively. Genetic distance Fst values of mtDNA haplogroup polymorphism in Liannan Yao ethnic group together with that of 12 other ethnic groups were presented in Table 3. Moreover, Liannan Yao nationality and Fujian She nationality belong to the same group, while Yunnan Yao nationality, Guizhou Dong nationality, Hunan Miao nationality, and Hunan Tujia nationality belong to another group, presented in Fig. 2. And the genetic distances of Liannan Yao, Yunnan

Yao, Hunan Miao and Fujian She are relatively close, presented in Fig. 3.

Discussion

There are many advantages to using mtDNA for the study of human evolution, expansion, and migration. In the process of sperm-egg fusion, the egg cytoplasm is mainly retained, so mtDNA is transmitted exclusively by maternal inheritance. MtDNA exists in the haploid form and lacks germline recombination, so it can retain the complete genetic information of the ancestors in the process of transmission. The mutation rate of mtDNA is 10-fold that of chromosomal DNA, which allows mtDNA to accumulate stable mutations more quickly and form genetic markers for specific populations, thereby increasing the resolution and information amount in mtDNA detection. There are hundreds to thousands of copies of mtDNA in each cell, making it more convenient to detect mtDNA than chromosomal DNA. These advantages make mtDNA analysis widely used in the studies of population evolution and migration.

Thanks to the rapid development of sequencing technology, the genetics community has conducted a large number of assessments and analysis of mtDNA haplogroup diversity in the past few decades in different regions and ethnic groups in China.[18-26] MtDNA plays an important role in studying the matrilineal origins of various ethnic groups in the contemporary East Asian population. Although many problems remain to be solved and the related database is not perfect, the historical migration route of the population is still relatively clear. Previous mtDNA haplogroup polymorphism studies showed that in East Asia, the mitochondrial characteristic haplogroups of southern China were B, M, F and R haplotypes, and the haplotype A, C, D and G appeared frequently in the northern area were also likely coming from the south.This also verified the conclusion in archaeology that modern people gradually migrated to the north and expanded there after entering Southeast Asia.[27]

According to the results in this study, Chinese Liannan Yao group's RMP value of 58 mtDNA SNP sites was lower, which was consistent with that in other Chinese reports. The most common haplogroups of Liannan Yao ethnic group were M and N, belonged to the Asian population, which was consistent with our previous research on the Y chromosome gene haplotype.[28] Haplogroups M and N were two large groups based on mtDNA, haplotypes A, B, R9, and N9 in haplogroup N and haplotypes C, D, G, M7, M8, and M9 in haplogroup M were specific for the East Asian population, which was verified in this study.

There were large differences in genetic markers in Chinese different ethnic groups, Liannan Yao ethnic group is a relatively closed group, and its genetic characteristics are also different from other ethnic groups. Liannan Yao group had more close genetic distance with Guizhou Miao, Yunnan Yao and Fujian She group, which supported the statement "three 'miao' groups derived from a same group" in anthropology, linguistics and archaeology.

Conclusions

This study is the first to investigate 60 specific mtDNA loci and test the genetic mtDNA polymorphisms in the Liannan Yao nationality, which provides support for Chinese mtDNA genome databases and further research on population genetics. It can be concluded from this study that the mtDNA haplogroup of Liannan Yao ethnic group has obvious genetic characteristics of the southern Chinese population, and closer genetic distance to Fujian She, Yunnan Yao, and Hunan Miao ethnic groups. However, in order to further determine the genetic background of Chinese

Liannan Yao group, more referenced populations and genetic markers need to be collected in our future research.

Methods

Blood samples from 209 unrelated healthy volunteers were randomly collected from Chinese Liannan Yao ethnic group in Guangdong province. The written informed consent was acquired from each of them and migration events did not exist in their family history of the participants for at least three generations. Blood samples were collected respectively in terms of the standard procedure. MtDNA were extracted to build up gene database. 60 mtDNA loci were sequenced, compared with the revised Cambridge Reference Sequence(rCRS) to determine the haplogroups, then haplotype frequency of mtDNA was calculated and Principal Component Analysis(PCA) was conducted between Liannan Yao nationality and other ethnic groups.

MtDNA was extracted according to the previous protocol [15]. The Expressmarker mtDNA-SNP 60 kit (AGCU ScienTech Incorporation, Jiangsu, Wuxi) was used for direct polymerase chain reaction amplification of 60 mtDNA loci(709，3010，3970，5178，6446，7196，8414，10310，10398，10873，12705，13104，13928，15043，16311，8794/8793，16126/16129，9bp，1541，1719，2706，4216，4883，5460，8020，8584，9123，1541，9698，9824，10400，11215，11251，11719，12372，12811，16362，8701/8697，152，1811，3348，4491，4833，5417，5442，7600，8684，8964，9477，9545，10397，11944，12007，14569，15784，16316/16319). 25ul PCR reaction system included: 10ul reaction mix, 1ul hot start C-Taq enzyme, 5ul amplification primer, and 9ul water. PCR amplification cycle parameters were as follows: 95℃ for 3 min, 94℃ for 30 s, 60℃ for 45 s, 10 cycles of 72℃ for 1 min, 94℃ for 10 s, 59℃ for 1 min, 15 cycles of 65℃ for 80 s, 60℃ for 20 min, finally stored at 4℃. Electrophoresis was performed by 3130XL Genetic Analyzer, analyzed by GeneMapper ID-X software.

The study was conducted following the recommendations of the DNA Committee of the International Society for Forensic Genetics (ISFG) as described by Carracedo, et al.[16]

The results of genotyping were referred to the rCRS [17]for subsequent statistical analysis. Haplotype determination were referred to the HaploGrep 2.0 platform (https://haplogrep.i-med.ac.at), Haplogroups were classified by phylotree (http://www.phylotree.org). Genetics parameter statistics include RMP, DP, and genetic distance. RMP = $\sum Xi^2$(Xi represents the frequency of the i-th mtDNA haplotype), DP=1- $\sum X2$, variability=[n(1-X2)]/(n-1)(n represents number of samples, x represents the frequency of the haplogroups). Genetic differentiation index (Fst) between groups were analyzed by PHYLIP, and phylogenetic tree was reconstructed by POPTREEW based on the pairwise population genetic distance values to infer the genetic background of the Liannan Yao nationality.

List of abbreviations

mtDNA	mitochondrial DNA
SPN	Single nucleotide polymorphism
rCRS	revised Cambridge Reference Sequence
PCR	Polymerase chain reaction
RMP	random match probability
DP	discrimination power
PCA	Principal Component Analysis

Declarations

Ethics approval and consent to participate

This study was approved by the Ethical Committee of the First Affiliated Hospital of Guangdong Pharmaceutical University, Medical Ethics Review [2019] (No. 143). All participants were interviewed to confirm their ethnic origins and sign the informed consents.

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Compeieting interests

The authors declares that they have no conflict of interest.

Funding

Not applicable.

Authors' contributions

SS analyzed and interpreted the data, and was a major contributor in writing the manuscript. JC and MZ were secondary contributors in analyzing data, HH and XL were secondary contributors in writing the manuscript, WD performed the examination of the blood sample, HS provided the blood sample of Liannan Yao population. All authors read and approved the final manuscript.

Acknowledgments

The authors are very grateful to all sample donors for their contributions to this work and all those who helped with sample collection.

References

[1]Michaels G S , Hauswirth W W , Laipis P J . Mitochondrial DNA copy number in bovine oocytes and somatic cells. Developmental biology. 1982;94:246-51.

[2]Robin E D , Wong R . Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells. Journal of Cellular Physiology. 1988;136:507-13.

[3]Moore M K , Frazier K . Humans Are Animals, Too: Critical Commonalities and Differences Between Human and Wildlife Forensic Genetics. Journal of Forensic Sciences. 2019;64:3.

[4]Zavala E I , Rajagopal S , Perry G H , et al. Impact of DNA degradation on massively parallel sequencing-based autosomal STR, iiSNP, and mitochondrial DNA typing systems. International Journal of Legal Medicine. 2019; doi:10.1007/s00414019021104.

[5] Zheng J H , Nie H T , Yang F , et al. Genetic variation and population STRucture of different geographical populations of Meretrix petechialis based on mitochondrial gene COI. Journal of Genetics. 2019;98:3.

[6]Ingman M , Kaessmann H , S Pääbo, et al. Mitochondrial genome variation and the origin of modern humans. Nature. 2000;6813:708-13.

[7]He G , Wang Z , Wang M , et al. Forensic anceSTRy analysis in two Chinese minority populations using massively parallel sequencing of 165 anceSTRy nformative SNPs. Electrophoresis. 2018;39:21.

[8]Zhang C , Li H , Zhao X , et al. Validation of Expressmarker mtDNA-SNP60: A mitochondrial-SNP kit for forensic application. ELECTROPHORESIS. 2016;3721:2848-61.

[9]Dennis McNevin. Forensic inference of biogeographical anceSTRy from genotype: The Genetic AnceSTRy Lab. Wiley Interdisciplinary Reviews: Forensic Science. 2019;doi:10.1002/wfs2.1356.

[10]Gregoire Leroy, Elise Gicquel, Paul Boettcher,et al. CoanceSTRy rate's estimate of effective population size for genetic variability monitoring. Conservation Genetics Resources. 2019; doi:10.1007/s12686019010920.

[11] Zhao Shilei, Zhang Tao, Liu Qi,et al. Identifying Lineage-Specific Targets of Natural Selection by a Bayesian Analysis of Genomic Polymorphisms and Divergence from Multiple Species. Molecular Biology & Evolution. 2019;6:6.

[12] Mara Cerezo, Leonor Gusmo, ViktorCerny,et al. Comprehensive Analysis of Pan-African Mitochondrial DNA Variation Provides New Insights into Continental Variation and Demography. Journal of Genetics and Genomics. 2016.

[13] Fourdrilis S , Backeljau T . Highly polymorphic mitochondrial DNA and deceiving haplotypic differentiation: implications for assessing population genetic differentiation and connectivity. Bmc Evolutionary Biology. 2019;19:1.

[14] Zheng, S. A Study of Yao's Architecture in Liannan: from a Humanity Perspective. Journal of Huizhou University. 2018;38:90–3.

[15]Walsh PS, Metzger DA, Higuchi R. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques. 2013;54:134–9.

[16]Carracedo, A. et al. Update of the guidelines for the publication of genetic population data. Forensic science international. Genetics. 2014;doi:10.1016/j.fsigen.2014.01.004.

[17]Andrews R, Kubacka I, Chinnery P, et al.Reanalysis and revision of the Cambridge reference

sequence for human mitochondrial DNA. Nat Genet. 1999;23:147.
[18]Anderson S , Bankier A T , Barrell B G , et al. Anderson, S. et al. Sequence and organization of the mitochondrial human genome. Nature. 1981;5806:457-65.
[19]Andrews, R., Kubacka, I., Chinnery, P. et al. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat Genet. 1999;23:147.
[20]Kivisild T , Tolk H V , Parik J , et al. The Emerging Limbs and Twigs of the East Asian mtDNA Tree. Molecular Biology & Evolution. 2002;19:1737-51.
[21]Yong-Gang Yao, Qing-Peng Kong, Hans-Jürgen Bandelt,et al. Phylogeographic Differentiation of Mitochondrial DNA in Han Chinese. American Journal of Human Genetics. 2002;doi:10.1086/338999.
[22] Wen B, Li H, Lu D et al. Genetic evidence supports demic diffusion of Han culture. Nature. 2004;7006: 302-5.
[23] Qiong Lan, Tong Xie, Xiaoye Jin,et al. MtDNA polymorphism analyses in the Chinese Mongolian group: Efficiency evaluation and further matrilineal genetic STRucture exploration. Molecular Genetics & Genomic Medicine. 2019;doi:10.1002/mgg3.934.
[24] Yao Y G , Nie L , Harpending H , et al. Genetic relationship of Chinese ethnic populations revealed by mtDNA sequence diversity. American Journal of Physical Anthropology. 2002;118:63-76.
[25]Xu, Kaikai,Hu, Shengping.Population data of mitochondrial DNA HVS-I and HVS-II sequences for 208 Henan Han Chinese. Legal medicine. 2015;17:287-94.
[26] Kakuda T , Shojo H , Tanaka M , et al. Multiplex APLP System for High-Resolution Haplogrouping of Extremely Degraded East-Asian Mitochondrial DNAs. Plos One. 2016;11:6.
[27]Beny F , Toucanne S , Skonieczny C , et al. Geochemical provenance of sediments from the northern East China Sea document a gradual migration of the Asian Monsoon belt over the past 400,000 years. Quaternary Science Reviews. 2018;190:161-75.
[28]Yaoqi Liao, Ling Chen, Runze Huang, et al. Genomic Portrait of Guangdong Liannan Yao Population Based on 15 Autosomal STRs and 19 Y-STRs. Scientific reports. 2019;doi:10.1038/s4159801836262x.