



## OPEN ACCESS

EDITED AND REVIEWED BY  
Ronny Decorte,  
Faculty of Medicine, KU Leuven, Belgium

## \*CORRESPONDENCE

He Guanglin,  
✉ Guanglinhescu@163.com  
Wang Mengge,  
✉ Menggewang2021@163.com

## SPECIALTY SECTION

This article was submitted to Evolutionary  
and Population Genetics,  
a section of the journal  
Frontiers in Genetics

RECEIVED 12 October 2022  
ACCEPTED 22 December 2022  
PUBLISHED 05 January 2023

## CITATION

Guanglin H, Lan-Hai W and Mengge W  
(2023), Editorial: Forensic investigative  
genetic genealogy and fine-scale structure  
of human populations.  
*Front. Genet.* 13:1067865.  
doi: 10.3389/fgene.2022.1067865

## COPYRIGHT

© 2023 Guanglin, Lan-Hai and Mengge.  
This is an open-access article distributed  
under the terms of the [Creative Commons  
Attribution License \(CC BY\)](#). The use,  
distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Editorial: Forensic investigative genetic genealogy and fine-scale structure of human populations

He Guanglin<sup>1\*</sup>, Wei Lan-Hai<sup>2</sup> and Wang Mengge<sup>3\*</sup>

<sup>1</sup>Institute of Rare Diseases, West China Hospital of Sichuan University, Sichuan University, Chengdu, China, <sup>2</sup>School of Ethnology and Anthropology, Inner Mongolia Normal University, Hohhot, China, <sup>3</sup>Faculty of Forensic Medicine, Zhongshan School of Medicine, Sun Yat-Sen University, Guangzhou, China

## KEYWORDS

population structure, admixture history, genome-wide genetic markers, forensic investigative genetic genealogy, genetic diversity

## Editorial on the Research Topic

[Forensic investigative genetic genealogy and fine-scale structure of human populations](#)

## Introduction

Anatomically modern humans originated in Africa and separated from their most likely recent common ancestor hundreds and thousands of years ago (Bergstrom et al., 2020; Wang et al., 2021a). They followingly migrated out of Africa around 50 thousand years ago and evolved in concert with the complicated interplay of gene flow and adaptive selection during the peopling of Eurasia, Oceania, and America (Bergstrom, et al., 2020; Wang et al., 2021a). Genomic studies have demonstrated the pervasiveness of population differentiation and genetic admixture between long-isolated ethnic groups (Bergstrom et al., 2020; Pan et al., 2022). Extensive population bottleneck, adaptive evolution in changing environments, and introgression from archaic hominins further shaped the complicated patterns of human genetic heritage. In general, complex population divergence, migration, and admixture events extensively shaped the patterns of genetic diversity of ethnolinguistically diverse populations.

There is increasing evidence to suggest that the differences in the susceptibility of many common and rare diseases are primarily attributed to human populations' diverse cultural, environmental, demographic, and genetic histories. The comprehensive understanding of fine-scale population evolutionary history will gradually change our understanding of the genetic architecture of diseases (Timpson et al., 2018; Benton et al., 2021; Pan et al., 2022). Thus, there is an urgent need to expand genetic research to populations with different ancestries. In addition, population genetic studies based on multiple genome-wide genetic markers (short tandem repeat, STR; single nucleotide polymorphism, SNP; Insertion/Deletion, InDel; copy number variation, CNV, and so on) could provide new insights into the detailed process of population admixture and evolutionary history of ethnolinguistically and geographically diverse populations. Understanding the fine-scale population structure also helps the better study design in medical genomics and the comprehensive practice applications in population genetics and forensic science.

With the rapid development of genotyping technologies, sequencing platforms, and computational methods, previous studies have provided the basal framework of the genetic

landscape of worldwide populations from different perspectives (Li et al., 2008; Lippold et al., 2014). However, most genetic studies focused on the relationships and fine-scale population structures were conducted *via* low-density genetic markers. It is also now possible to capture and sequence ancient DNA from ancient samples, which could provide pivotal insights into the formation of spatiotemporally diverse populations with unprecedented resolution. Although huge nationwide biobanks for characterizing the genotypes and phenotypes of millions of people have been established (Barton et al., 2021; Zhang et al., 2021; Chiu et al., 2022), more geographically, linguistically, and culturally diverse populations (especially non-metropolitan populations) are needed to be studied systematically at different spatio-temporal scales.

A clear understanding of genetic background and diversity of ethnolinguistically diverse populations and decoding their demographic history can provide new medical and forensic application opportunities. Genetic studies have illuminated the population-specific reference database, effective algorithm and the developed panel for the targeted forensic applications were the fundamentals of forensic intelligence inference of external visual appearance, biogeographical ancestry inference and forensic investigative genetic genealogy (FIGG) (He et al., 2018). FIGG, one new and rapidly growing field of forensic genetics since 2018, has attracted the attention of geneticists focused on complex familial search (Phillips 2018). Currently, many projects aim to develop and validate new FIGG panels, construct and complement forensic FIGG databases, and develop new statistical models to promote the practice of FIGG (Kling et al., 2021; Tillmar et al., 2021).

To summarize the new advances in FIGG and fine-scale population structure and illuminate the importance of full-scale genetic structure and diversity as the basis for the FIGG, we organized this Research Topic of the “*Forensic investigative genetic genealogy and fine-scale structure of Human Populations*”. In detail, this Research Topic aimed to characterize the genetic background and demographic history of ethnolinguistically and geographically diverse populations based on different densities of genetic marker. It would be helpful in exploring the long-range familial searches and fine-scale genetic localization within subgroups in other continents. This Research Topic attracted research focused on the basic knowledge exploration of the genetic background of one targeted population and included applied research focused on developing and validating forensic amplification systems.

## Exploration of theoretical knowledge—Fine-scale genetic structure and demographic history reconstruction

Recent studies have demonstrated that population history reconstruction leveraging high-density genetic markers could uncover previously unrecognized population structures at a fine scale compared with forensically relevant loci (Li, et al., 2008; Bergstrom, et al., 2020). Zhou et al. generated and analyzed genome-wide data of Liaoning Han people and found that genetic differences existed in geographically different Sinitic-speaking Han populations, which might result from other migration and admixture events of Hans during the period of “Chuang Guandong”. Hou et al. investigated the population history of Liaoning Mongolians based on ~700,000 SNPs and provided new insights

into the admixture history of Mongolic-speaking Mongolians according to shared allele-based analyses. He et al. explored the demographic history of Qiang people based on Eurasian modern and ancient reference populations. This study revealed that the Tibeto-Burman-speaking Qiang people derived their primary ancestry from Tibetan-related ancestral populations in North China. Wang et al. performed a genome-wide association study on 26,806 Chinese individuals. They identified 21 SNPs associated with widow’s peak, unibrow, double eyelid, earlobe attachment, and freckles. This study may facilitate a better understanding of the genetic basis of facial development in Chinese populations and provide new markers for forensic phenotype predictions. These studied Han, Mongolian and Qiang people are widely distributed in North China, and other populations from southern China and surrounding regions need to be further explored based on the array or sequencing data, such as Austronesian, Austroasiatic, Hmong-Mien and Tai-Kadai people. Generally, population genetic studies showed a strong correlation between Chinese cultural language and geographic patterns and population structure.

The estimated patterns of genetic diversity in China can help accurately biogeographic ancestry inference and FIGG. Similar population stratifications were also identified in Siberian populations. Currently, available tools could effectively distinguish populations with different continental origins, but most of these are not efficient for differentiating people within the same continent. Gorin et al. selected 5,229 AISNPs and tested various mathematical models for biogeographic ancestry inference. The results showed that the accuracy of the prediction of this panel on one of 29 studied ethnic groups reached 71% and the proposed method could be employed to predict ancestries from Russian and neighboring populations. Huang et al. developed a machine learning approach for estimating the relationships with high error SNP profiles and found that this approach was more accurate and robust than the individual measures.

## Forensic potential applications—Development and validation of forensic systems focused on personal identification and family research

STR genotyping has been applied in forensic investigations for nearly 30 years (Hagelberg et al., 1991; Kayser and de Knijff 2011). Nowadays, several commercial STR kits have been developed based on the expanded CODIS loci (Oostdik et al., 2014; Wang et al., 2018; Qu et al., 2019; Batham et al., 2020; Green et al., 2021). However, with the rapid increase in the number of STR genotypes in forensic databases, more novel non-CODIS STRs with high genetic polymorphisms are required to minimize the incidence of adventitious matches. Huang et al. validated the forensic performance of a novel multiplex autosomal STR panel (including six CODIS STRs and 20 non-CODIS STRs). They found that this novel kit could be applied as a promising tool for forensic human identification and complex paternity analysis. Than et al. genotyped seven Lao Isan and three Laotian populations using Verifiler plus PCR Amplification kit. The allelic frequency results provided the genetic background of Austroasiatic and Tai-Kadai people from Laos and Thailand.

InDel loci, the second most abundant polymorphism across the human genome, possess low mutation rates and small amplicon lengths compared with STRs (Weber et al., 2002; Mills et al.,

2006), which have been proven to be of value in forensic investigations (Pereira et al., 2009; Zhang et al., 2018). And InDels showing significant allele frequency differences among geographically and linguistically diverse populations can be adopted as ancestry-informative markers (AIMs) (Sun et al., 2016; Inacio et al., 2017). Moreover, previous studies showed that multi-InDel loci behaved well in parentage tests and could be used for forensic applications (Fan et al., 2016; Sun, et al., 2016; Qu et al., 2020). Liu et al. developed and validated a six-color fluorescence multiplex panel including 59 autosomal InDels. Subsequently, Tibetan groups from China have been genotyped using the newly-developed 59-plex InDel panel. The comprehensive population genetic analyses showed that this homemade panel could be used as a powerful tool for individual forensic identification and paternity testing in Chinese Tibetan groups. Jin et al. developed a Next-Generation Sequencing (NGS) InDel panel, including 17 multi-InDels on the X chromosome. They found that the newly-developed panel could be adopted as an effective tool for individual forensic identification, paternity testing, and biogeographical ancestry inference.

Genetic surveys based on uniparentally inherited markers have identified many paternal and maternal founding lineages in regional populations and their corresponding expansion events (Poznik et al., 2016; Li et al., 2019). Jia et al. sequenced complete mitochondrial genomes of 146 Daur individuals in China. The results showed that the Daur ethnic group has high maternal genetic diversity and may have experienced recent population expansion. He et al. developed and validated the AGCU-Y30 Y-STR panel and conducted a Y-STR-based study to explore the paternal history of the Qiang people. The validated results showed that the novel Y-STR kit was sensitive and robust enough for forensic applications. Population genetic analyses revealed that the Qiang people are closely related to lowland Tibetan-Yi corridor populations.

## Prospects and challenges

Human genetics needed the fundamental foundation supporting large-scale population genomic projects to characterize the full landscape of human genetic diversity and population structure, such as NHLBI TOPMed (Taliun et al., 2021), gnomAD (Collins et al., 2020) and UK10K (Wang et al., 2021). These projects made significant advances in European human genetics. Other projects, including the Chinese 10K\_CPGDP (Chinese population genomic diversity project), GSRD-100K<sup>WCH</sup> and ChinaMap, were recently launched to explore the genetic features of under-represented populations. Characterizing the genetic architecture of ethnolinguistically and geographically diverse populations will promote our understanding of population origin, separation, admixture, adaptation, and gene flow from archaic individuals.

Knowledge of the fine-scale genetic structure and population-specific genomic database is the base for FIGG. Synthesizing our growing knowledge of evolutionary history with forensic investigations will help to achieve the promise of long-range familial searches and fine-scale genetic localization. For human population genetics and FIGG, we also had some challenges that needed to be overcome in the next step:

Firstly, European bias in worldwide human genetic studies and Han Chinese bias in Chinese cohort research hinder in health equality of genetic studies and forensic genetics. More population genomic

studies from under-reported populations need to be conducted to characterize their uncharacterized genetic polymorphisms and the genetic spectrum of different genetic markers.

Second, a population-specific genomic database should be constructed to provide a comprehensive landscape of different genetic markers and even include structural variations and mobile elements via the PacBio sequencing plateau.

Third, cooperation of national genomic studies should be formed to promote data sharing in different institutions.

Forth, population genomic projects should be conducted among included subjects with deep phenotypes, which would provide values for exploring the genetic basis for physical traits and medical phenotypes.

Fifth, regional population-specific genomic datasets, FIGG panels, algorithms, and forensic databases should be developed and validated. FIGG has promoted successful inspection of criminal suspects among American or European populations, as there are publicly available European/American genomic databases and DNA. Land and GEDmatch servers. Similar databases from other countries or regions should be constructed in the future.

Seventh, IBD-based algorithms need the high-density phased SNPs or whole genome sequencing data. Many forensic samples were highly degraded or had minute amounts of genetic materials. Thus, the primary focus is to develop more FIGG panels based on low-density SNP markers and non-linkage algorithms (allele-sharing status and allele frequency spectrum).

In short, more ethnolinguistically and geographically diverse populations are needed to be studied based on different genetic markers and algorithm models, and population-specific panels based on different genetic variations and corresponding forensic databases need to be developed to achieve long-range familial searches and fine-scale genetic structure reconstruction.

## Author contributions

MW and GH drafted it, and MW, L-HW, and GH proofread it. Before submitting it, all authors reviewed it.

## Funding

This work was funded by the National Natural Science Foundation of China (82202078).

## Acknowledgments

Thanks to all the brilliant Frontiers team whose invaluable assistance at each step made the Research Topic successful. We also thank Ronny Decorte from the Faculty of Medicine, KU Leuven, who gave the essential revised advice for our work.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Barton, A. R., Sherman, M. A., Mukamel, R. E., and Loh, P. R. (2021). Whole-exome imputation within UK Biobank powers rare coding variant association and fine-mapping analyses. *Nat. Genet.* Aug 53, 1260–1269. Epub 2021/07/07. doi:10.1038/s41588-021-00892-1
- Batham, M. S., Kushwaha, K. P. S., Chauhan, T., Kumawat, R. K., and Shrivastava, P. (2020). Autosomal STR allele frequencies in Kahars of Uttar Pradesh, India, drawn with PowerPlex® 21 multiplex system. *Int. J. Leg. Med. Mar.* 134, 517–519. Epub 2019/03/29. doi:10.1007/s00414-019-02046-9
- Benton, M. L., Abraham, A., LaBella, A. L., Abbot, P., Rokas, A., and Capra, J. A. (2021). The influence of evolutionary history on human health and disease. *Nat. Rev. Genet. May* 22, 269–283. Epub 20210106. doi:10.1038/s41576-020-00305-9
- Bergstrom, A., McCarthy, S. A., Hui, R., Almarri, M. A., Ayub, Q., Danecek, P., et al. (2020). Insights into human genetic variation and population history from 929 diverse genomes. *Sci. Mar* 20, eaay5012. Epub 2020/03/21. doi:10.1126/science.aay5012
- Chiu, A. M., Molloy, E. K., Tan, Z., Talwalkar, A., and Sankararaman, S. (2022). Inferring population structure in biobank-scale genomic data. *Am. J. Hum. Genet. Apr* 7 (109), 727–737. Epub 2022/03/18. doi:10.1016/j.ajhg.2022.02.015
- Collins, R. L., Brand, H., Karczewski, K. J., Zhao, X., Alfoldi, J., Francioli, L. C., et al. (2020). A structural variation reference for medical and population genetics. *Nat. May* 581, 444–451. Epub 20200527. doi:10.1038/s41586-020-2287-8
- Fan, G. Y., Ye, Y., and Hou, Y. P. (2016). Detecting a hierarchical genetic population structure via Multi-InDel markers on the X chromosome. *Sci. Rep.* 6 (6), 32178. Epub 2016/08/19. doi:10.1038/srep32178
- Green, R., Elliott, J. L., Norona, W., Go, F., Nguyen, V. T., Ge, J., et al. (2021). Developmental validation of VeriFiler™ plus PCR amplification kit: A 6-dye multiplex assay designed for casework samples. *Forensic Sci. Int. Genet. Jul* 53, 102494. Epub 2021/03/20. doi:10.1016/j.fsigen.2021.102494
- Hagelberg, E., Gray, I. C., and Jeffreys, A. J. (1991). Identification of the skeletal remains of a murder victim by DNA analysis. *Nature* 352, 427–429. doi:10.1038/352427a0
- He, G., Wang, Z., Wang, M., Luo, T., Liu, J., Zhou, Y., et al. (2018). Forensic ancestry analysis in two Chinese minority populations using massively parallel sequencing of 165 ancestry-informative SNPs. *Electrophor. Nov.* 39, 2732–2742. Epub 20180628. doi:10.1002/elps.201800019
- Inacio, A., Costa, H. A., da Silva, C. V., Ribeiro, T., Porto, M. J., Santos, J. C., et al. (2017). Study of InDel genetic markers with forensic and ancestry informative interest in PALOP's immigrant populations in Lisboa. *Int. J. Leg. Med. May* 131, 657–660. Epub 2016/11/01. doi:10.1007/s00414-016-1484-3
- Kayser, M., and de Knijff, P. (2011). Improving human forensics through advances in genetics, genomics and molecular biology. *Nat. Rev. Genet. Mar.* 12, 179–192. Epub 2011/02/19. doi:10.1038/nrg2952
- Kling, D., Phillips, C., Kennett, D., and Tillmar, A. (2021). Investigative genetic genealogy: Current methods, knowledge and practice. *Forensic Sci. Int. Genet. May* 52, 113–124. Epub 2021/02/17. doi:10.1016/j.fsigen.2019.06.019
- Li, J. Z., Absher, D. M., Tang, H., Southwick, A. M., Casto, A. M., Ramachandran, S., et al. (2008). Worldwide human relationships inferred from genome-wide patterns of variation. *Sci. Feb* 22, 3191100–3191104. Epub 2008/02/23. doi:10.1126/science.1153717
- Li, Y. C., Ye, W. J., Jiang, C. G., Zeng, Z., Tian, J. Y., Yang, L. Q., et al. (2019). River valleys shaped the maternal genetic landscape of han Chinese. *Mol. Biol. Evol.* 36, 1643–1652. Epub 2019/05/22. doi:10.1093/molbev/msz072
- Lippold, S., Xu, H., Ko, A., Li, M., Renaud, G., Butthof, A., et al. (2014). Human paternal and maternal demographic histories: Insights from high-resolution Y chromosome and mtDNA sequences. *Investig. Genet.* 5, 13. Epub 20140924. doi:10.1186/2041-2223-5-13
- Mills, R. E., Luttig, C. T., Larkins, C. E., Beauchamp, A., Tsui, C., Pittard, W. S., et al. (2006). An initial map of insertion and deletion (INDEL) variation in the human genome. *Genome Res. Sep.* 16, 1182–1190. Epub 20060810. doi:10.1101/gr.4565806
- Oostdik, K., Lenz, K., Nye, J., Schelling, K., Yet, D., et al. (2014). Developmental validation of the PowerPlex® fusion system for analysis of casework and reference samples: A 24-locus multiplex for new database standards. *Forensic Sci. Int. Genet. Sep.* 12, 69–76. Epub 2014/06/07. doi:10.1016/j.fsigen.2014.04.013
- Pan, Y., Zhang, C., Lu, Y., Ning, Z., Lu, D., Gao, Y., et al. (2022). Genomic diversity and post-admixture adaptation in the Uyghurs. *Natl. Sci. Rev. Mar.* 9, nwab124. nwab124. Epub 20210911. doi:10.1093/nsr/nwab124
- Pereira, R., Phillips, C., Alves, C., Amorim, A., Carracedo, A., and Gusmao, L. (2009). A new multiplex for human identification using insertion/deletion polymorphisms. *Electrophor. Nov.* 30, 3682–3690. Epub 2009/10/29. doi:10.1002/elps.200900274
- Phillips, C. (2018). The Golden State Killer investigation and the nascent field of forensic genealogy. *Forensic Sci. Int. Genet. Sep.* 36, 186–188. Epub 2018/07/25. doi:10.1016/j.fsigen.2018.07.010
- Poznik, G. D., Xue, Y., Mendez, F. L., Willems, T. F., Massaia, A., Wilson Sayres, M. A., et al. (2016). Punctuated bursts in human male demography inferred from 1, 244 worldwide Y-chromosome sequences. *Nat. Genet. Jun* 48, 593–599. Epub 2016/04/26. doi:10.1038/ng.3559
- Qu, S., Li, H., Li, Y., Lv, M., Yang, F., Zhu, J., et al. (2019). Developmental validation of the Microreader™ 20A ID system. *Electrophoresis* 40, 3099–3107. Epub 2019/10/10. doi:10.1002/elps.201900221
- Qu, S., Lv, M., Xue, J., Zhu, J., Wang, L., Jian, H., et al. (2020). Multi-indel: A microhaplotype marker can be typed using capillary electrophoresis platforms. *Front. Genet.* 11, 567082. Epub 2020/11/17. doi:10.3389/fgene.2020.567082
- Sun, K., Ye, Y., Luo, T., and Hou, Y. (2016). Multi-InDel analysis for ancestry inference of sub-populations in China. *Sci. Rep.* 6 (6), 39797. Epub 2016/12/23. doi:10.1038/srep39797
- Taliun, D., Harris, D. N., Kessler, M. D., Carlson, J., Szpiech, Z. A., Torres, R., et al. (2021). Sequencing of 53, 831 diverse genomes from the NHLBI TOPMed Program. *Nat. Feb* 590, 290–299. Epub 20210210. doi:10.1038/s41586-021-03205-y
- Tillmar, A., Sturk-Andreaggi, K., Daniels-Higginbotham, J., Thomas, J. T., and Marshall, C. (2021). The FORCE panel: An all-in-one SNP marker set for confirming investigative genetic genealogy leads and for general forensic applications. *Genes. (Basel)* 12, 1968. Epub 20211210. doi:10.3390/genes12121968
- Timpson, N. J., Greenwood, C. M. T., Soranzo, N., Lawson, D. J., and Richards, J. B. (2018). Genetic architecture: The shape of the genetic contribution to human traits and disease. *Nat. Rev. Genet. Feb* 19, 110–124. Epub 20171211. doi:10.1038/nrg.2017.101
- Wang, C. C., Yeh, H. Y., Popov, A. N., Zhang, H. Q., Matsumura, H., Sirak, K., et al. (2021a). Genomic insights into the formation of human populations in East Asia. *Nat. Mar.* 591, 413–419. Epub 2021/02/23. doi:10.1038/s41586-021-03336-2
- Wang, M., Wang, Z., He, G., Jia, Z., Liu, J., and Hou, Y. (2018). Genetic characteristics and phylogenetic analysis of three Chinese ethnic groups using the Huaxia Platinum System. *Sci. Rep. Feb* 5 (8), 2429. Epub 2018/02/07. doi:10.1038/s41598-018-20871-7
- Wang, Q., Dhindsa, R. S., Carss, K., Harper, A. R., Nag, A., Tachmazidou, I., et al. (2021b). Rare variant contribution to human disease in 281, 104 UK Biobank exomes. *Nat. Sep.* 597, 527–532. Epub 20210810. doi:10.1038/s41586-021-03855-y
- Weber, J. L., David, D., Heil, J., Fan, Y., Zhao, C., and Marth, G. (2002). Human diallelic insertion/deletion polymorphisms. *Am. J. Hum. Genet. Oct.* 71, 854–862. Epub 20020904. doi:10.1086/342727
- Zhang, P., Luo, H., Li, Y., Wang, Y., Wang, J., Zheng, Y., et al. (2021). NyuWa genome resource: A deep whole-genome sequencing-based variation profile and reference panel for the Chinese population. *Cell. Rep. Nov.* 16, 37110017. Epub 2021/11/18. doi:10.1016/j.celrep.2021.110017
- Zhang, S., Zhu, Q., Chen, X., Zhao, Y., Zhao, X., Yang, Y., et al. (2018). Forensic applicability of multi-allelic InDels with mononucleotide homopolymer structures. *Electrophoresis* 39, 2136–2143. Epub 20180710. doi:10.1002/elps.201700468