

## Forensic characteristics and genetic substructure analysis of the Handan Han population, Northern China

Chi-Zao Wang, Lan-Hai Wei, Jia-Shuo Zhang, Xue-Er Yu, Ru-Feng Bai, Hui Li, Mei-Sen Shi & Shu-Hua Ma

To cite this article: Chi-Zao Wang, Lan-Hai Wei, Jia-Shuo Zhang, Xue-Er Yu, Ru-Feng Bai, Hui Li, Mei-Sen Shi & Shu-Hua Ma (2023) Forensic characteristics and genetic substructure analysis of the Handan Han population, Northern China, *Annals of Human Biology*, 50:1, 123-125, DOI: [10.1080/03014460.2023.2181985](https://doi.org/10.1080/03014460.2023.2181985)

To link to this article: <https://doi.org/10.1080/03014460.2023.2181985>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 21 Mar 2023.



[Submit your article to this journal](#)



Article views: 484



[View related articles](#)




[View Crossmark data](#)

SHORT REPORT



## Forensic characteristics and genetic substructure analysis of the Handan Han population, Northern China

Chi-Zao Wang<sup>a,b,c,\*</sup>, Lan-Hai Wei<sup>d,\*</sup>, Jia-Shuo Zhang<sup>b</sup> , Xue-Er Yu<sup>e</sup>, Ru-Feng Bai<sup>f</sup>, Hui Li<sup>e,g</sup>, Mei-Sen Shi<sup>f</sup> and Shu-Hua Ma<sup>a,b,c</sup>

<sup>a</sup>Department of Radiology, The First Affiliated Hospital of Shantou University Medical College, Shantou, China; <sup>b</sup>Shantou University Medical College, Shantou, China; <sup>c</sup>Laboratory of Medical Molecular Imaging, The First Affiliated Hospital of Shantou University Medical College, Shantou, China; <sup>d</sup>School of Ethnology and Anthropology, Inner Mongolia Normal University, Hohhot, China; <sup>e</sup>State Key Laboratory of Genetic Engineering and Ministry of Education Key Laboratory of Contemporary Anthropology, Fudan University, Shanghai, China; <sup>f</sup>Criminal Justice College of China University of Political Science and Law, Beijing, China; <sup>g</sup>Shanxi Academy of Advanced Research and Innovation, Fudan-Datong Institute of Chinese Origin, Datong, China

### ABSTRACT

We analysed the forensic characteristics and substructure of the Handan Han population based on 36 Y-STR (short tandem repeat) and Y-SNP (single nucleotide polymorphism) markers. The two most dominant haplogroups in Handan Han, O2a2b1a1a1-F8 (17.95%) and O2a2b1a2a1a (21.51%), and their abundant downstream branches, reflected the strong expansion of the precursor of the Hans in Handan. The present results enrich the forensic database and explore the genetic relationships between Handan Han and other neighbouring and/or linguistically close populations, which suggests that the current concise overview of the Han intricate substructure remains oversimplified.

### ARTICLE HISTORY

Received 7 October 2022  
Revised 22 December 2022  
Accepted 27 January 2023

### KEYWORDS

Forensic characteristics;  
genetic substructure;  
Handan Han population

## 1. Introduction





Handan is one of the earliest sites of domesticated common millet in East Asia, where many important events and stories took place in Chinese history (Figure S1). According to the latest census (2021), Han Chinese make up the majority of the nearly 10 million residents of Handan City. Previous studies have sketched out several main clusters in the Han groups, but also pointed out significant differences and intricate substructures among them (Wen et al. 2004; Xu et al. 2009). Therefore, regional lineage accumulations and spatial comparisons of the Han population among groups are still necessary.

## 2. Subjects and methods


In order to investigate the phylogenetic history and forensic characteristics of the Han nationality in Handan, a total of 730 unrelated male individuals were recruited and genotyped by the DNATyper™ 36Y Kit (No.2 Institute of the Ministry of Public Security, Beijing, China). All samples were collected after receiving informed consent, and individuals were considered autochthonous if their ancestors had lived in Handan municipality for at least three generations. The

Ethical Committee of Fudan University, Shanghai, People's Republic of China, approved the study (No. 14012). Experimental and statistics methods are shown in Table S1. Data were submitted to the YHRD (Y chromosomal haplotype reference database, <https://yhrd.org>) under accession number YA004677 for the Handan Han population. First, to outline the genetic relationship between Handan Han and other Asian populations, 27 Y-STR data from 20 typical populations were selected for AMOVA based on linguistic and geographical classification and visualised in a multidimensional scaling (MDS) plot. Subsequently, to enlarge the reference dataset and refine the genetic background understanding of Handan Han, 17 Y-STR data from 38 populations were used for further population comparison. According to genetic distance matrices (Rst matrices), a phylogenetic tree was built with neighbor-joining (N-J) methods by MEGA-X (Kumar et al. 2018) and optimised by the Interactive Tree of Life v5 (Letunic and Bork 2021). The details of the reference dataset are shown in Table S2.

In accordance with Y-DNA haplogroup Tree 2020 (<https://isogg.org/tree/>), we constructed a simplified phylogenetic tree that showed the overall distribution of Handan Han samples. To better understand the distribution of prevailing haplogroups in the studied population, we integrated data

**CONTACT** Mei-Sen Shi  [shimeisen2000@163.com](mailto:shimeisen2000@163.com)  Criminal Justice College of China University of Political Science and Law, Beijing, China; Shu-Hua Ma  [shuhua6633@163.com](mailto:shuhua6633@163.com)  Department of Radiology, The First Affiliated Hospital of Shantou University Medical College, Shantou, China

\*These authors contributed equally to this work.

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/03014460.2023.2181985>

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

from previously published papers (Shi et al. 2005; Xue et al. 2006; Park et al. 2012; Trejaut et al. 2014; Ning et al. 2016; Li et al. 2020; Zhou et al. 2020) and constructed contour maps by Surfer 15 (<https://www.goldensoftware.com/products/surfer>).

### 3. Results and discussion

The allele frequencies and the gene diversity (GD) values for 36 Y-STR loci are summarised in Table S3. 407 alleles were detected across the 36 Y-STR loci, with frequencies between 0.0014 and 0.7041. Among 30 single-copy loci, DYS449 was the most informative locus with the GD of 0.8743. The three multi-copy Y-STRs, DYS385a/b, DYF387S1 a/b and DYS527a/b were the most diverse markers with GD values of 0.9642, 0.9483 and 0.9437, respectively. The Y-STRs haplotype distributions and Y-SNP of the Handan Han population are listed in Table S4. A total of 729 unique Y-STR haplotypes were found in the 730 Handan Han samples, of which 728(99.73%) were distinct. To evaluate the utility of the new markers for forensic casework, haplotype-based analyses were repeated for various subsets of Y-STRs, namely, Yfiler® marker panel (17 loci), Yfiler® Plus marker panel (27 loci), and DNATyper™ 36Y marker panel (Table S5). Overall, an increase in the number of analysed Y-STR markers decreased the number of shared haplotypes and increased the number of unique haplotypes. These results indicate that the DNATyper™ 36Y kit offers excellent discrimination capabilities and may be useful in forensic investigations and paternal lineage identification in Han populations.

To outline the genetic relationship between Handan Han and other Asian populations, we compared the 27 Yfiler haplotype data of Handan Han with those of 20 neighbouring populations based on linguistic and geographical classification. Table S6 shows that there were significant differences ( $p < 0.0024$  after Bonferroni's correction) between the Handan Han and all the other populations. Among them, Handan Han had the longest genetic distance from populations of Mongolic, Tibeto-Burman, and Turkic groups ( $R_{st} \sim 0.0615-0.3386$ ), followed by Southern Hans (Guangdong, Guizhou, and Hainan) and Hunan-Miao ( $R_{st} \sim 0.0226-0.0350$ ). Low  $R_{st}$  values between Handan Han and the other Northern Hans (Changchun, Shannxi, Beijing, and Shandong) and Baishan-Manchu (a typical Tungusic-speaking population in Northern China) were obtained ( $R_{st} \sim 0.0056-0.0163$ ). To portray the patterns of population genetic relationships, an MDS plot constructed based on linearised  $R_{st}$  values (Figure S2) showed that all populations were generally clustered according to language and geographical factors. Handan Han, all the other Han populations, and Baishan-Manchu gathered in the upper right quadrant. A preliminary analysis indicated that language and geographical factors had a great influence on the genetic composition of Handan Han; for in-depth comparative analysis, we need to expand the range of Chinese-speaking groups and collect more populations from northern China.

On the basis of the above analysis, we further sorted out the 17 Yfiler haplotype data from Chinese-speaking groups and Tungusic-speaking populations (Table S7). According to

the  $R_{st}$  values of these 38 populations, an N-J phylogenetic tree (Figure S3) was performed to depict their forensic genetic landscape. The Handan Han first clustered with the northern Han group (Dalian-Han, Gansu-Han, Beijing-Han, Shandong-Han, Shannxi-Han, Luoyang-Han, and Hulunbuir-Han), followed by the Tungusic-speaking group (Liaoning-Manchu, Chengde-Manchu, Baishan-Manchu, etc.) and then with the Hui group (represented by the Henan-Hui), which together formed a large northern branch in the phylogenetic tree. At the same time, the other populations constituted another large southern group cluster. This north-south differentiation pattern further emphasises the great influence of geographical factors on the genetic background. Although the Northern Han, Southern Han, and Hakka groups share the Han identity, they belong to different branches in the tree. The above groups presented a clear North-South dichotomous pattern, but we noted that Yunnan-Hui from the south was located on the northern branch. The Hui group is unique in that they speak Chinese as the Han group but practice Islam. Yunnan-Hui was mainly shaped by large-scale migration in the Ming Dynasty (Chuang 2018) and there is a lack of fine-scale genetic structure research. The Yunnan-Hui population in the YHRD database (sample size = 43) contains a higher frequency of haplogroup J1 (> 18.60%) than others (Lang et al. 2019). Therefore, the current clustering results are generally acceptable.

As the simplified phylogenetic tree shows in Figure S4, the majority (69.45%) of Handan Han individuals were classified into subbranches of haplogroup O-M175, which most Hans also belonged to (Shi et al. 2005), followed by haplogroups C (13.29%), N (8.49%) and Q (5.07%). After further subdivision, 38 different terminal haplogroups were predicted (Figure S4). Three Han dominant lineages (Yan et al. 2014) encompass 45.20% of the Handan Han population in total (estimated 17.95% for O2a2b1a1a1-F8, 21.51% for O2a2b1a2a1a-F46, and 5.75% for O2a1b-002611). Among the three super-grandfather types, O2a2b1a1a1-F8 ( $O_{\alpha}$ -F8) and O2a2b1a2a1a-F46 ( $O_{\beta}$ -F46) have particularly high frequencies. In addition, Q1a1a-M120 (31, 4.25%), N1a-F1206 (42, 5.75%) and O2\*-M122 (48, 6.58%) also appeared frequently (Rootsi et al. 2007; Ilumae et al. 2016; Sun et al. 2019; Ma et al. 2021), which is consistent with the geographical background of Handan Han.

To extensively illustrate the frequency distribution of the two prevailing haplogroups in this study,  $O_{\alpha}$ -F8 and  $O_{\beta}$ -F46, contour maps were drawn (Figure S5). Although both of them are the most abundant paternal lineages in East Asian populations,  $O_{\beta}$ -F46 was mainly centred in northern China, while  $O_{\alpha}$ -F8, in addition to northern China, is also very popular among many groups in Southwest China. It is generally believed that these two haplogroups began to spread in the Upper and Central Yellow River Basin approximately 6,000–7,000 years ago and were related to the shift from hunter-gather subsistence to intensive agriculture in North China (Xue et al. 2006; Yan et al. 2014; Wang et al. 2018). Geographically speaking, the Handan City is located at or very close to the diffusion centres of these two haplogroups, so it is rational that  $O_{\alpha}$ -F8 and  $O_{\beta}$ -F46 appear frequently in

Handan Hans. Furthermore, we supposed that Handan Han samples also include abundant downstream branches of these two haplogroups, especially some unanalysed paralogues. From the view of network structure (Figure S6), in Handan Hans, O $\beta$ -F46 seems to contain two large secondary clades, while the downstream structure of O $\alpha$ -F8 is more dispersive.

#### 4. Conclusions

This study attempts to provide a clearer scenario of the Han genetic background in Handan, which is also an important complement to the current oversimplified Han substructure. We analysed the forensic characteristics and genetic substructure of the Handan Han population based on 36 Y-STR and predicted Y-SNP markers. The DNATyper<sup>TM</sup> 36Y set showed excellent discrimination capabilities in Handan Han and should be reliable for database and casework samples. Stratified population comparisons showed that Handan had close relationships with Chinese-speaking groups in northern China which is consistent with the linguistic and geographic classifications. The combination of Y-STR and Y-SNP markers reflected the strong expansion of the precursor of the Hans in Handan. With the development of sequencing technology, the Han group may exhibit a deeper and more complex evolutionary structure. These data obtained in this study could potentially be useful for regional-specific and prerequisite references to forensic, genealogical, and evolutionary researches.

#### Acknowledgments

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

#### Research involving human participants

The Ethical Committee of Fudan University, Shanghai, People's Republic of China approved the study (No. 14012). All samples were collected after receiving informed consent.

#### Disclosure statement

No potential conflict of interest was reported by the author(s).

#### Funding

This study was funded by the National Natural Science Foundation of China [31900406,81774395]; Qian Duansheng distinguished scholar support program of China University of Political Science and Law [01140065140]; Grant for Key Disciplinary Project of Clinical Medicine under the Guangdong High-level University Development Program [Grant No. 002-18120302]; Cross disciplinary construction project of evidence investigation [No: 10322308].

#### ORCID

Jia-Shuo Zhang  <http://orcid.org/0000-0002-0738-7653>

#### References

- Chuang MA. 2018. A study on migration of Yunnan Hui people and intercultural dialogue since the Ming and Qing dynasties. *Journal of Yunnan Open University*. 20(1):45–48.
- Ilumae AM, Reidla M, Chukhryaeva M, Jarve M, Post H, Karmin M, Saag L, Agdzhoyan A, Kushniarevich A, Litvinov S, et al. 2016. Human Y chromosome haplogroup N: A non-trivial time-resolved phylogeography that cuts across language families. *Am J Hum Genet*. 99(1):163–173.
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol*. 35(6):1547–1549.
- Lang M, Liu H, Song F, Qiao X, Ye Y, Ren H, Li J, Huang J, Xie M, Chen S, et al. 2019. Forensic characteristics and genetic analysis of both 27 Y-STRs and 143 Y-SNPs in Eastern Han Chinese population. *Forensic Sci Int Genet*. 42:e13–e20.
- Letunic I, Bork P. 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res*. 49(W1):W293–W296.
- Li M, Zhang Y, Luo L, Bian Y, Li C. 2020. Development and validation of a custom panel including 183 Y-SNPs for Chinese Y-chromosomal haplogroups dissection using a MALDI-TOF MS system. *Electrophoresis*. 41(23):2047–2054.
- Ma P, Yang X, Yan S, Li C, Gao S, Han B, Hou K, Robbeets M, Wei LH, Cui Y. 2021. Ancient Y-DNA with reconstructed phylogeny provide insights into the demographic history of paternal haplogroup N1a2-F1360. *J Genet Genomics*. 48(12):1130–1133.
- Ning C, Yan S, Hu K, Cui YQ, Jin L. 2016. Refined phylogenetic structure of an abundant East Asian Y-chromosomal haplogroup O\*-M134. *Eur J Hum Genet*. 24(2):307–309.
- Park MJ, Ee HYL, Yang WI, Shin KJ. 2012. Understanding the Y chromosome variation in Korea—relevance of combined haplogroup and haplotype analyses. *Int J Legal Med*. 126(4):589–599.
- Rootsi S, Zhivotovskiy LA, Baldovic M, Kayser M, Kutuev IA, Khusainova R, Bermisheva MA, Gubina M, Fedorova SA, Ilumae AM, et al. 2007. A counter-clockwise northern route of the Y-chromosome haplogroup N from Southeast Asia towards Europe. *Eur J Hum Genet*. 15(2):204–211.
- Shi H, Dong YL, Wen B, Xiao CJ, Underhill PA, Shen PD, Chakraborty R, Jin L, Su B. 2005. Y-chromosome evidence of southern origin of the East Asian-specific haplogroup O3-M122. *Am J Hum Genet*. 77(3):408–419.
- Xu SH, Yin XY, Li SL, Jin WF, Lou HY, Yang L, Gong XH, Wang HY, Shen YP, Pan XD, et al. 2009. Genomic dissection of population substructure of Han Chinese and its implication in association studies. *Am J Hum Genet*. 85(6):762–774.
- Sun N, Ma PC, Yan S, Wen SQ, Sun C, Du PX, Cheng HZ, Deng XH, Wang CC, Wei LH. 2019. Phylogeography of Y-chromosome haplogroup Q1a1a-M120, a paternal lineage connecting populations in Siberia and East Asia. *Ann Hum Biol*. 46(3):261–266.
- Trejaut JA, Poloni ES, Yen JC, Lai YH, Loo JH, Lee CL, He CL, Lin M. 2014. Taiwan Y-chromosomal DNA variation and its relationship with Island Southeast Asia. *BMC Genet*. 15:77.
- Wang LX, Lu Y, Zhang C, Wei LH, Yan S, Huang YZ, Wang CC, Mallick S, Wen SQ, Jin L, et al. 2018. Reconstruction of Y-chromosome phylogeny reveals two neolithic expansions of Tibeto-Burman populations. *Mol Genet Genomics*. 293(5):1293–1300.
- Wen B, Li H, Lu D, Song X, Zhang F, He Y, Li F, Gao Y, Mao X, Zhang L, et al. 2004. Genetic evidence supports demic diffusion of Han culture. *Nature*. 431(7006):302–305.
- Xue Y, Zerjal T, Bao W, Zhu S, Shu Q, Xu J, Du R, Fu S, Li P, Hurler ME, et al. 2006. Male demography in East Asia: a north-south contrast in human population expansion times. *Genetics*. 172(4):2431–2439.
- Yan S, Wang CC, Zheng HX, Wang W, Qin ZD, Wei LH, Wang Y, Pan XD, Fu WQ, He YG, et al. 2014. Y chromosomes of 40% Chinese descend from three Neolithic super-grandfathers. *Plos One*. 9(8):e105691.
- Zhou Z, Zhou Y, Yao Y, Qian J, Liu B, Yang Q, Shao C, Li H, Sun K, Tang Q, et al. 2020. A 16-plex Y-SNP typing system based on allele-specific PCR for the genotyping of Chinese Y-chromosomal haplogroups. *Leg Med (Tokyo)*. 46:101720.