

Comprehensive analyses of genetic diversities and population structure of the Guizhou Dong group based on 44 Y-markers

Hongling Zhang*, Xiaolan Huang*, Xiaoye Jin, Zheng Ren, Qiyan Wang, Meiqing Yang, Ronglan Xu, Xiang Yuan, Daiquan Yang, Hongyan Liu, Wanyi Shen, Huiying Zhang, Yangjie Que and Jiang Huang

Department of Forensic Medicine, Guizhou Medical University, Guiyang, Guizhou, China

*These authors contributed equally to this work.

ABSTRACT

Background: The non-recombining region of the human Y chromosome (NRY) is a strictly paternally inherited genetic marker and the best material to trace the paternal lineages of populations. Y chromosomal short tandem repeat (Y-STR) is characterized by high polymorphism and paternal inheritance pattern, so it has been widely used in forensic medicine and population genetic research. This study aims to understand the genetic distribution of Y-STRs in the Guizhou Dong population, provide reference data for forensic application, and explore the phylogenetic relationships between the Guizhou Dong population and other comparison populations.

Methods: Based on the allele profile of 44 Y-markers in the Guizhou Dong group, we estimate their allele frequencies and haplotype frequencies. In addition, we also compare the forensic application efficiency of different Y-STR sets in the Guizhou Dong group. Finally, genetic relationships among Guizhou Dong and other reference populations are dissected by the multi-dimensional scaling and the phylogenetic tree.

Results: A total of 393 alleles are observed in 312 Guizhou Dong individuals for these Y-markers, with allele frequencies ranging from 0.0032 to 0.9679. The haplotype diversity and discriminatory capacity for these Y-markers in the Guizhou Dong population are 0.99984 and 0.97440, respectively. The population genetic analyses of the Guizhou Dong group and other reference populations show that the Guizhou Dong group has the closest genetic relationship with the Hunan Dong population, and followed by the Guizhou Tujia population.

Conclusions: In conclusion, these 44 Y-markers can be used as an effective tool for male differentiation in the Guizhou Dong group. The haplotype data in this study not only enrich the Y-STR data of different ethnic groups in China, but also have important significance for population genetics and forensic research.

Submitted 24 May 2023
Accepted 5 September 2023
Published 25 September 2023

Corresponding authors
Hongling Zhang,
229598103@qq.com
Jiang Huang, mmm_hj@126.com
Academic editor
Luisa Azevedo
Additional Information and
Declarations can be found on
page 9
DOI 10.7717/peerj.16183

© Copyright
2023 Zhang et al.
Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Anthropology, Genetics, Medical Genetics

Keywords Guizhou Dong, Forensic features, Population genetics, STR, Y-chromosome

INTRODUCTION

The non-recombining region of the human Y chromosome (NRY) strictly follows a paternally inherited pattern, implying that it can only be transmitted from father to male descendants. Therefore, genetic markers on the NRY can be viewed as important markers to trace the paternal lineages of populations (Sofie et al., 2019). Since NRY is more susceptible to genetic drift caused by its small effective population size than other autosomal genetic markers (Jobling & Tyler-Smith, 2003), it usually shows diverse genetic distributions among different geographical populations, which makes the Y chromosome a powerful tool for investigating demographic events in human genetic history. Until now, the NRY has been widely applied in the studies of forensics, genealogical reconstruction, human origin, evolution and migration, population history, and medical genetics (Sofie et al., 2019; Jobling & Tyler-Smith, 2003; King & Jobling, 2009; Jobling & Tyler-Smith, 1995). For the NRY, Y chromosomal short tandem repeats (Y-STRs) are commonly used in forensic genetics and population genetics due to their high genetic polymorphism and paternal inheritance pattern. Since Y-STRs have an average mutation rate of 3.35×10^{-3} , they can also be used to study the genetic composition and consanguinity issues of various populations (Ballantyne et al., 2010). Currently, many commercial Y-STR kits have been developed for forensic research. For example, the Power Plex® Y23 System, which is a five-dye system, combines 17 Y-STR loci in the currently commercially available Y-STR kit (Y-filer) with six new high-discrimination Y-STR loci (Thompson et al., 2013). In 2017, He et al. analyzed the genetic polymorphisms of these 23 Y-STRs in the Yi population of Sichuan Province in China. Obtained results revealed that these Y-STRs displayed relatively high genetic diversities and could be used to differentiate males in the Yi population (He et al., 2017). However, as the extant DNA database gathers several individuals, the match chance of the allelic profile of random individuals will increase if only these 23 Y-STRs are tested. Accordingly, more Y-markers should be added to the existing panels to improve their haplotype resolution power. Compared to the Power Plex Y23 system, the Goldeneye DNA Identification System Y Plus Kit (Beijing PeopleSpot Inc.) can simultaneously amplify 41 Y-STRs and 3 Y chromosomal Insertion/Deletion polymorphisms (Y-InDels) in a single well using the six fluorescent dye labeling technology. The kit combines high and low mutation rate loci, especially for three Y-InDels, which can be used as the slow mutation loci for familial search. Besides, they could also enhance the discriminatory capacity of the kit to some degree. We postulate that the panel could achieve higher male differentiation than the Power Plex Y23 system. Therefore, the forensic application value of the panel is further evaluated in the current study.

There are more than 40 ethnic minorities in the Guizhou province. The Dong ethnic group, also known as the Kam population, is one of the major ethnic groups living in Guizhou. It is generally believed that the modern Dong population developed from the ancient Baiyue tribe, which migrated to the Guizhou in the Han Dynasty (Wu, 1993). The Dong group has their own language that belongs to a branch of the Tai-Kadai language family. According to the sixth national census statistics, there are 1,431,900 Dong

population in Guizhou, accounting for 4.12% of the province's population and 11.41% of the minority population in Guizhou ([Ma, 2010](#)). These Dongs are mainly distributed in southeastern Guizhou Province and adjacent to Guangxi Zhuang Autonomous Region and Hunan Province (<https://www.britannica.com/topic/Dong>). So far, some genetic data on the Y-STRs of Hunan Dong ([Feng et al., 2020](#)), the autosomal STRs of Guangxi Dong ([Guo et al., 2017a](#)) and the mitochondrial DNA of Guizhou Dong ([Ren et al., 2022](#)) have been reported. However, there are no studies on the Y-STRs of the Dong population in Guizhou. Therefore, we investigate the genetic distributions of 44 Y-markers in the Guizhou Dong group based on the Goldeneye DNA identification system Y Plus kit. Furthermore, Chinese different populations collected in the Y chromosome haplotype reference database (YHRD, <https://yhrd.org/>) are used as reference populations to further explore the genetic background of the Guizhou Dong group. The study is expected to enrich the genetic data of the Guizhou Dong group, which could be viewed as the reference data set for forensic research in the Guizhou Dong group.

MATERIALS AND METHODS

Sample information

In this study, bloodstain samples were collected from 312 unrelated healthy Dong individuals with written informed consent. All participants have lived in the Guizhou Province for at least three generations. This research is performed according to the guidelines of the Ethics Commission of Guizhou Medical University. The ethical permission is issued by the Ethics Committee of Guizhou Medical University (Approval Number: No. 224).

PCR amplification and capillary electrophoresis

According to the instructions of the Goldeneye DNA Identification System Y Plus kit, PCR amplification is performed on the aforementioned samples by using the 9700 PCR instrument (Applied Biosystems, Foster City, CA, USA). The total volume of the PCR reaction system is 10 μ L, including 4 \times premix VII 2.5 μ L, 5 \times Y Plus primer mixture 2.0 μ L, deionized water 5.5 μ L and 1.2 mm bloodstain. The reaction condition is initial denaturation at 95 °C for 2 min; then 27 cycles of 94 °C for 5 s, 60 °C for 90 s, 62 °C for 1 min; and then extension for 5 min at 60 °C, followed by a final hold at 15 °C. Next, we mix 1 μ L amplified product with 10 μ L deionized formamide and 0.4 μ L ORG 500 internal standard; and then the cocktail is separated and detected on the ABI3500 Genetic Analyzer. Finally, allele nomenclature is conducted using GeneMapperID-X1.3 software (Thermo Fisher Scientific, Waltham, MA, USA) by comparing it to the kit's allelic ladder.

Statistical analysis

Allele frequencies and haplotype frequencies of 41 Y-STRs and three Y-InDels in the Guizhou Dong population are calculated by the direct counting method. Gene diversity (GD) and haplotype diversity (HD) are calculated according to the formula 'GD or HD = $n(1 - \sum P_i^2)/(n-1)$ ', where P_i is the frequency of the i -th allele or haplotype and n is the total number of samples ([Nei, 1973](#)). Haplotype match probability (HMP) and

discrimination capacity (DC) values are calculated according to the formulas: $HMP = \sum P_i^2$ (P_i is the frequency of the haplotype), and $DC = N_{diff}/N$ (N_{diff} and N mean the number of different haplotypes and the sample size, respectively). Population genetic analyses of Guizhou Dong and 16 reference populations are conducted by the following methods. Firstly, analysis of molecular variance (AMOVA) is used to estimate genetic distances (R_{ST}) between the Guizhou Dong group and other populations (Guizhou Bouyei (Luo et al., 2019), Yunnan Dai, Hunan Dong (Shu et al., 2015), Guizhou Han (Sun et al., 2019; Zhou et al., 2016), Yunnan Han (Kwak et al., 2005; Yanmei et al., 2010; Yin et al., 2022, 2020), Yunnan Hui (Xie et al., 2019), Guizhou Miao (Tang et al., 2020; Luo et al., 2021), Hulunbuir Mongolian (Wang et al., 2019), Sichuan Qiang, Qinghai Salar (Zhu et al., 2007), Qinghai Tibetan (Zhu et al., 2008; Cao et al., 2018), Guizhou Tujia, Guizhou Yi (Song et al., 2021), Yunnan Yi (Zhu et al., 2006; Fan et al., 2019), Gansu Yugur and Guangxi Zhuang (Wang et al., 2022; Guo et al., 2017b)) by the online tool YHRD. Next, the phylogenetic tree of these populations is constructed by the MEGA v7.0 software (Kumar, Stecher & Tamura, 2016) with the method of Neighbor-Joining based on their pairwise R_{ST} values. In addition, the multi-dimensional scaling (MDS) of the Dong population and other reference populations is conducted by the SPSS v18.0 software (<https://www.ibm.com/products/spss-statistics>).

RESULTS

Haplotype distributions, allele frequencies, and genetic diversities of 44 Y-markers

The different haplotypes and corresponding haplotype frequencies in the Guizhou Dong population are arranged in Table S1. A total of 304 haplotypes are observed in 312 individuals, 296 of which are unique. Allele distributions and GD values of 44 Y-markers in the Dong population of Guizhou are shown in Fig. 1 and Table S2. Two to 46 different alleles are observed per locus, resulting in a total of 393 alleles in 312 Guizhou Dong individuals, with allele frequencies ranging from 0.0032 to 0.9679. The locus with the highest GD value is DYS385 (0.9399), while the locus with the lowest GD value is DYS645 (0.0627). In addition, we find that nine Y-markers (DYS391, DYS388, DYS437, DYS438, DYS596, DYS645, rs199815934, rs759551978 and rs771783753) display relatively low GD values (<0.5) in the studied population. The other loci show relatively high GD values (>0.5) in the Guizhou Dong population, especially for DYS385, DYS527 and DYF387S1 loci whose GD values are higher than 0.9. For four multi-copy Y-STRs, we observe that these four loci (DYS385, DYF387S1, DYS527, and DYF404S1) have higher GD values than most single-copy Y-STRs, with GD values of 0.9399, 0.9363, 0.9170, and 0.8806, respectively. Microvariant alleles are also found at the DYS458 (14.1), DYS518 (37.2), DYS576 (18.1, 19.1 and 20.1) and DYS557 (18.3 and 17.3) loci. Bi-alleles are observed at the single-copy locus DYS557 (13,14), and tri-alleles are detected at two double-copy loci DYS527 (19,21,22) and DYF404S1 (13,14,16) (Table S1).

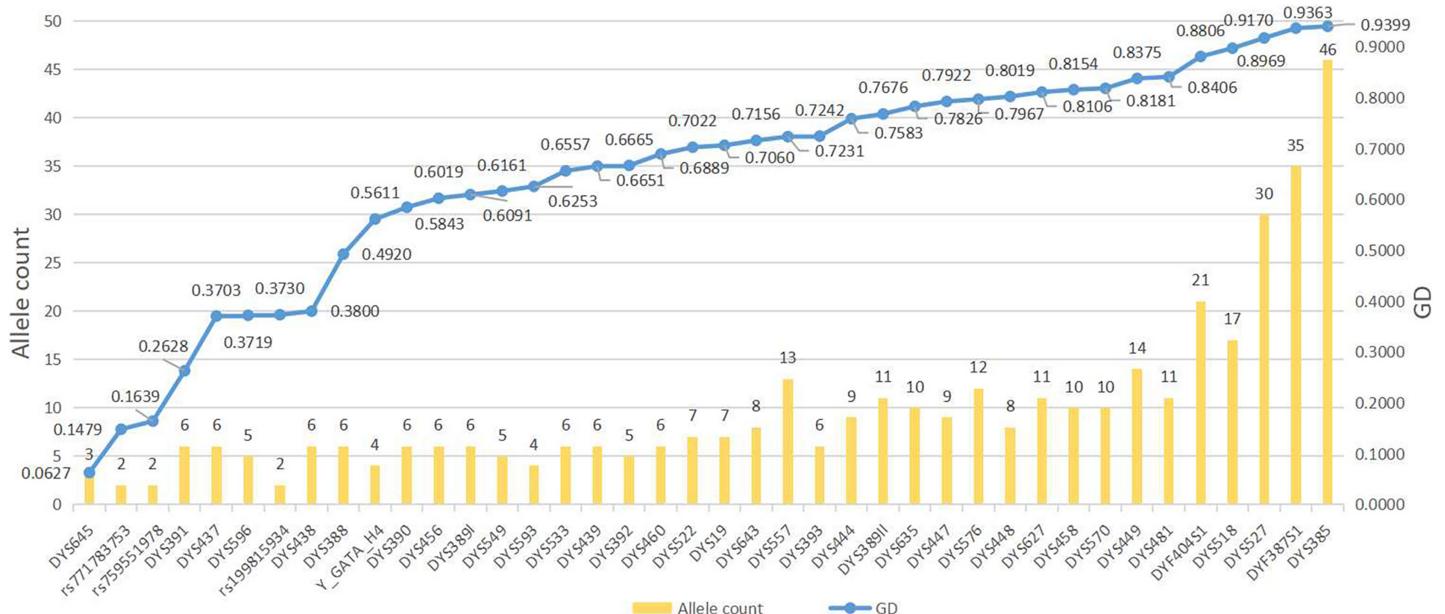


Figure 1 The numbers of alleles and gene diversities of 44 Y-markers in the Guizhou Dong population. The blue line on behalf of GD values, and the yellow bar on behalf of alleles counts. Y axis on the left indicates allele count, and Y axis on the right indicates GD values.

Full-size DOI: 10.7717/peerj.16183/fig-1

Comparison of forensic efficacy of different Y sets in the Guizhou Dong population

The number of loci included by different Y kits and the number of overlapping loci between these kits are shown in [Table S3](#). The DC, HD, HMP values and haplotype numbers for different Y-marker sets in the Guizhou Dong population are shown in [Fig. 2](#) and [Table S4](#). The HMP, HD and DC values of the 44 Y-markers are 0.00337, 0.99984 and 0.97440, respectively. The DC value increases from 0.72115 for 9 Y-STRs of the minimal set to 0.97440 for 44 Y-markers (the kit in this study). The HD value increases from 0.99143 for 9 Y-STRs to 0.99984 for 44 Y-markers. Furthermore, in [Fig. 2](#), we find that with the increase of the number of Y-markers, DC and HD values gradually increase, while HMP gradually decreases. More importantly, the Goldeneye DNA identification system Y Plus kit in this study has higher DC values and higher HD values than other kits except for the Y filer Plus kit.

Genetic relationships between the Guizhou Dong population and other reference populations

Firstly, the genetic relationships between the Dong population in Guizhou and 16 other reported populations are evaluated using the R_{ST} values, which represent genetic distances between populations. [Table S5](#) summarizes the pairwise R_{ST} values in the Dong population of Guizhou and 16 other reference populations from China. Heatmap of the R_{ST} values is revealed in [Fig. 3](#). The genetic distance between the Dong population in Guizhou and the Yi population in Yunnan is the largest ($R_{ST} = 0.1573$), followed by the Tibetan population in Qinghai ($R_{ST} = 0.1550$). This indicates that the genetic distances between the Dong

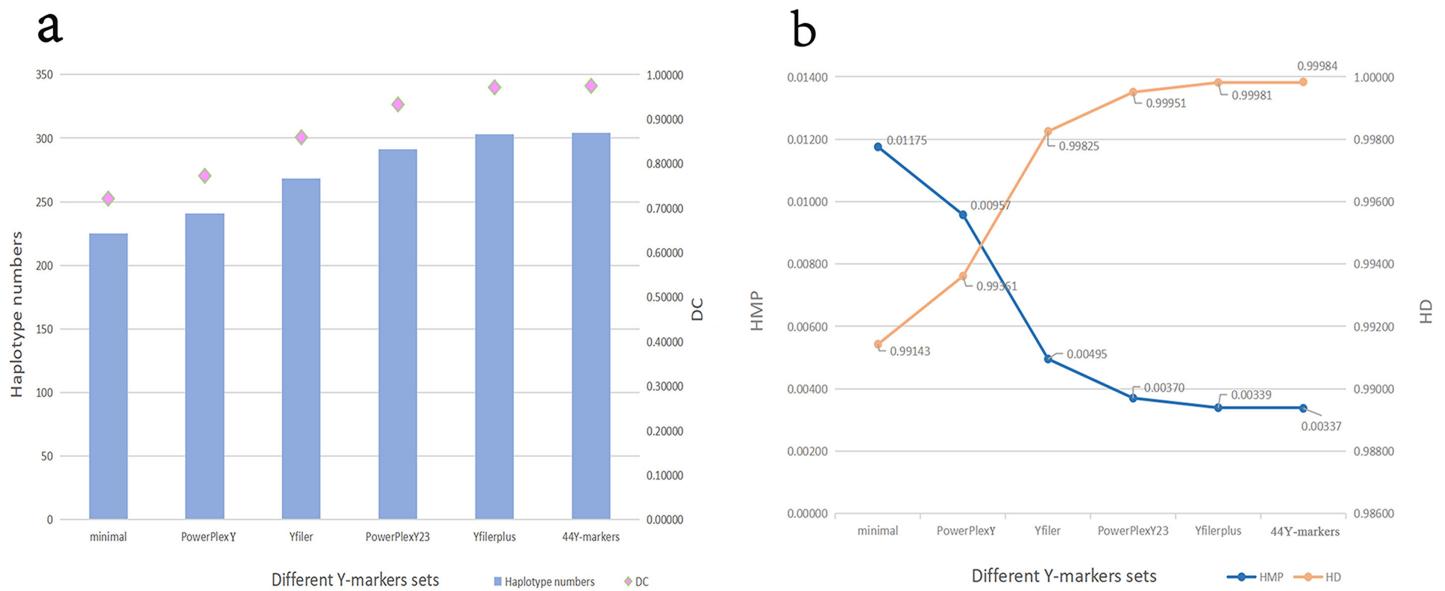


Figure 2 Forensic efficiency comparisons of different Y-STR sets in the Guizhou Dong population. (A) Haplotype numbers and DC values of different Y-STRs kits in Guizhou Dong population. Y axis on the left indicates haplotype numbers, and Y axis on the right indicates discriminatory capacity. (B) HMP and HD values of different Y-STRs kits in Guizhou Dong population. Y axis on the left indicates haplotype match probability, and Y axis on the right indicates haplotype diversity.

[Full-size](#) DOI: 10.7717/peerj.16183/fig-2

population in Guizhou and the above two groups are relatively far. Genetically close to the Dong population in Guizhou are the Hunan Dong population ($R_{ST} = 0.0162$), Guizhou Tujia population ($R_{ST} = 0.0211$), Guizhou Yi population ($R_{ST} = 0.0261$) and Guizhou Han population ($R_{ST} = 0.0283$).

Next, based on the paired R_{ST} values, we conducted the MDS of the Dong group in Guizhou and other reference groups, as shown in Fig. 4A. The results show that the 17 groups can be roughly divided into three parts: the Qinghai Tibetan is in the upper right corner; the Dai and Yi nationalities of Yunnan are located directly below; the remaining groups are on the left. As for the Dong ethnic group studied, we find that this group is close to the Dong ethnic group in Hunan and the Tujia ethnic group in Guizhou. In addition, a phylogenetic tree is constructed using the Neighbor-Joining method based on the paired R_{ST} values of the 17 populations, as shown in Fig. 4B. The results show that the studied Dong ethnic group first gathers in a branch with the Hunan Dong ethnic group, and then with the Guizhou Tujia ethnic group, indicating that these groups have relatively close genetic relationships. These findings are consistent with the above MDS results.

DISCUSSION

The results of this study indicate that the 44 Y-markers can be used for forensic applications and population genetic studies in the Dong population of Guizhou. Most of these 44 Y-markers display relatively high genetic polymorphisms ($GD > 0.5$) in the Guizhou Dong population. Even so, we find that DYS438, DYS391, DYS437 and DYS645 loci exhibit relatively low genetic diversities in the studied population. Similar results are also observed in Hunan Zhuang and Guangxi Zhuang populations (Feng et al., 2020; Guo

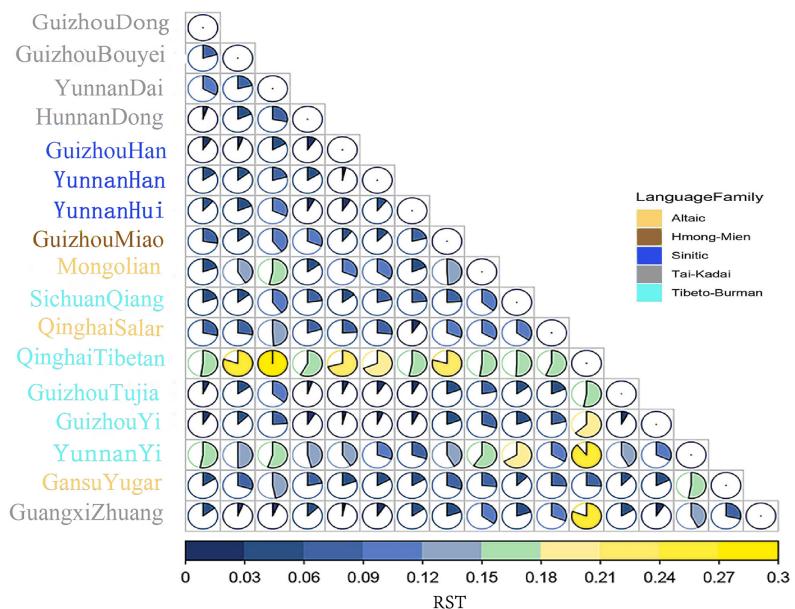


Figure 3 The heatmap based on the paired-*Rst* values for Guizhou Dong population and other reference populations. The squares of different colors on the right represent different language families, and populations of the same language family show the same color. Besides, the larger area of the color blocks in the circular chart indicates a more distant genetic distance.

[Full-size](#) DOI: [10.7717/peerj.16183/fig-3](https://doi.org/10.7717/peerj.16183/fig-3)

et al., 2017a). We infer that these four Y-STRs might be selected for forensic application in European populations because these loci exhibit relatively high GD values in European populations, especially for DYS391, DYS437 and DYS438 loci ([Hallenbergs et al., 2005](#); [Pickrahn et al., 2016](#)). In addition, we observe that the GD values of DYS645, rs759551978 and rs771783753 loci are lower than 0.2, indicating that these three loci possess poor genetic diversities in the Guizhou Dong. Accordingly, we need to select more highly polymorphic Y-STRs to replace these loci to obtain better haplotype resolution. At the same time, trialleles are found at two double-copy loci (DYS527 and DYS404S1) in the Guizhou Dong population. Similar results are found in individuals of the Shandong Han population, which may be related to the extra copies of these loci on the Y chromosome ([Chen et al., 2022](#)). This needs to be verified by the Sanger sequencing. For these 44 Y-markers, we find that they show higher DC values than other Y-marker sets. However, no obvious difference between 44 Y-markers and 27 Y-STRs (the Y filer Plus kit) is seen in the Guizhou Dong population, which may be related to the relatively small sample size (312) of the studied population. As the increasing of the sample size, we postulate that these 44 Y-markers will display higher DC and HD values than the Y filer Plus kit. Anyway, we think that these 44 Y-markers could be viewed as a valuable tool for distinguishing unrelated male individuals in the Guizhou Dong population based on their HD and DC values.

In Fig. 3, the genetic distance between Guizhou Dong and Hunan Dong is the closest, followed by the Guizhou Tujia population. In Figs. 4A and 4B, the Guizhou Dong group and Hunan Dong group are also clustered together, which indicates that the same

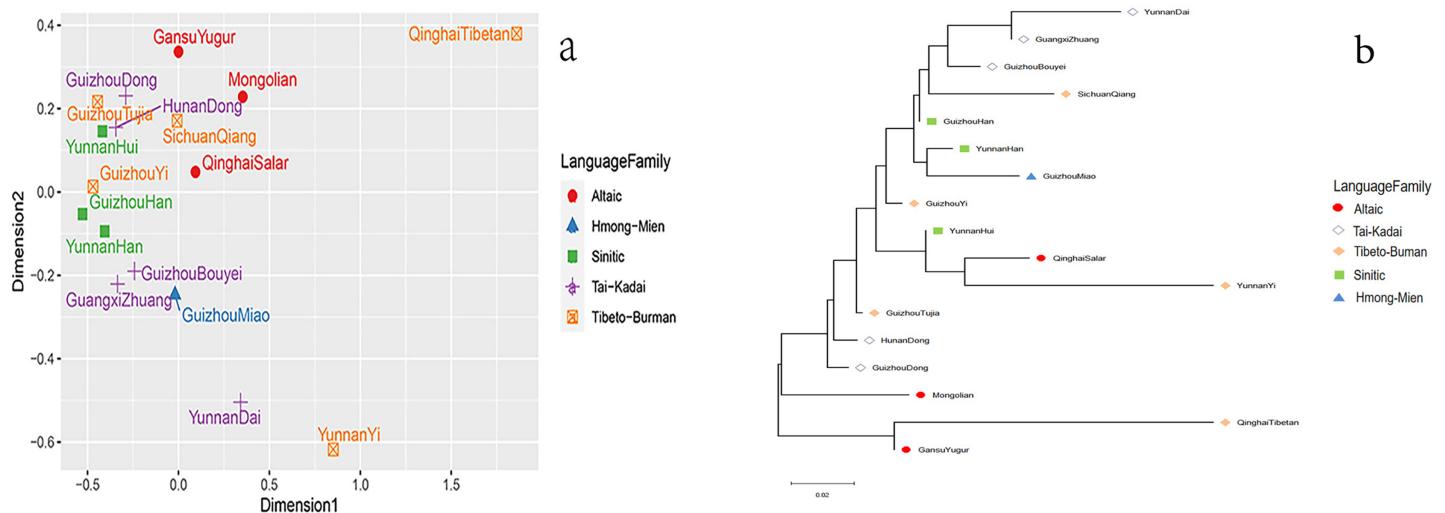


Figure 4 Population genetic analyses of Guizhou Dong population and other reference populations. (A) The MDS of Guizhou Dong and other populations; (B) the phylogenetic tree of Guizhou Dong and other populations.

[Full-size](#) DOI: 10.7717/peerj.16183/fig-4

population in different regions is genetically closer. Conversely, the genetic distance from the Guizhou Dong is the farthest from the Qinghai Tibetan population, followed by the Gansu Yugur, Mongolian and Yunnan Dai populations. The distant genetic relationship between the Dong population in Guizhou and the Tibetan population in Qinghai may be due to the fact that most Tibetans live at high altitudes and have little contact with other populations, which makes them genetically distant from other populations (Lu *et al.*, 2016). In the previous study on 30 InDel loci of the Dong population, Liu *et al.* (2020) also found that the Dong population in Guizhou and the Tibetan population in Qinghai had high genetic differentiation. Nonetheless, the exploration of the genetic structure of the Dong population in Guizhou is not enough, and it is necessary to further study the genetic composition of the Dong population in Guizhou by using more genetic markers, such as whole genome sequencing, to clarify the genetic background of the Dong population in Guizhou.

CONCLUSION

We first assess the genetic distributions of 44 Y-markers in the Guizhou Dong population based on the Goldeneye DNA Identification System Y Plus kit. Among these loci, most markers exhibit high genetic polymorphisms in the Guizhou Dong population. The HD and DC values of 44 Y-markers show that the kit is capable of providing relatively high-resolution haplotypes, and these 44 loci can be well applied to forensic applications in the Dong population of Guizhou. At the same time, the haplotype data in this study can also provide valuable reference data for population genetics studies in the future.

ACKNOWLEDGEMENTS

We are grateful to all volunteers for providing their blood samples.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This study was supported by the Guizhou Province Education Department, Characteristic Region Project, Qian Education KY No. (2021) 065; the Guizhou Innovation training program for college students (202110660018); the Guizhou Provincial Science and Technology Projects (ZK[2022] General 355 the Guizhou Education Department Young Scientific and Technical Talents Project, Qian Education KY NO. (2022) 215; the Guizhou Scientific Support Project, Qian Science Support (2021) General 448; the Guizhou "Hundred" High-level Innovative Talent Project, Qian Science Platform Talents (2020) 6012; the Guizhou Scientific Support Project, Qian Science Support (2019) 2825; the National Natural Science Foundation (No. 82160324); and the Guizhou Science Project, Qian Science Foundation (2020) 1Y353. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

Guizhou Province Education Department, Characteristic Region Project, Qian Education: (2021) 065.

Guizhou Innovation training program: (202110660018).

Guizhou Provincial Science and Technology: (ZK(2022) General 355).

Guizhou Education Department Young Scientific and Technical: (2022) 215.

Guizhou Scientific: (2021) General 448.

Qian Science Platform Talents: (2020) 6012.

Qian Science Support: (2019) 2825.

National Natural Science Foundation: (No. 82160324).

Guizhou Science Project, Qian Science Foundation: (2020) 1Y353.

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Hongling Zhang conceived and designed the experiments, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Xiaolan Huang conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Xiaoye Jin analyzed the data, prepared figures and/or tables, and approved the final draft.
- Zheng Ren analyzed the data, prepared figures and/or tables, and approved the final draft.
- Qiyan Wang analyzed the data, prepared figures and/or tables, and approved the final draft.

- Meiqing Yang analyzed the data, prepared figures and/or tables, and approved the final draft.
- Ronglan Xu performed the experiments, prepared figures and/or tables, and approved the final draft.
- Xiang Yuan performed the experiments, prepared figures and/or tables, and approved the final draft.
- Daiquan Yang performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Hongyan Liu performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Wanyi Shen performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Huiying Zhang performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Yangjie Que performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Jiang Huang conceived and designed the experiments, prepared figures and/or tables, and approved the final draft.

Human Ethics

The following information was supplied relating to ethical approvals (*i.e.*, approving body and any reference numbers):

The Ethics Committee of Guizhou Medical University.

Data Availability

The following information was supplied regarding data availability:

The raw data is available in the [Supplemental Files](#).

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.16183#supplemental-information>.

REFERENCES

- Ballantyne KN, Goedbloed M, Fang R, Schaap O, Lao O, Wollstein A, Choi Y, van Duijn K, Vermeulen M, Brauer S, Decorte R, Poetsch M, von Wurmb-Schwarz N, de Knijff P, Labuda D, Vezina H, Knoblauch H, Lessig R, Roewer L, Ploski R, Dobosz T, Henke L, Henke J, Furtado MR, Kayser M. 2010. Mutability of Y-chromosomal microsatellites: rates, characteristics, molecular bases, and forensic implications. *American Journal of Human Genetics* 87(3):341–353 DOI [10.1016/j.ajhg.2010.08.006](https://doi.org/10.1016/j.ajhg.2010.08.006).
- Cao S, Bai P, Zhu W, Chen D, Wang H, Jin B, Zhang L, Liang W. 2018. Genetic portrait of 27 Y-STR loci in the Tibetan ethnic population of the Qinghai province of China. *Forensic Science International: Genetics* 34:e18–e19 DOI [10.1016/j.fsigen.2018.02.005](https://doi.org/10.1016/j.fsigen.2018.02.005).
- Chen A, Luo L, Tao R, Zhang S, Li C. 2022. Forensic parameters of 41 Y-STR loci in Shandong Han individuals and comparison with 42 other populations. *Forensic Sciences Research* 7(4):823–825 DOI [10.1080/20961790.2021.1963397](https://doi.org/10.1080/20961790.2021.1963397).

- Fan GY, An YR, Peng CX, Deng JL, Pan LP, Ye Y.** 2019. Forensic and phylogenetic analyses among three Yi populations in Southwest China with 27 Y chromosomal STR loci. *International Journal of Legal Medicine* **133**(3):795–797 DOI [10.1007/s00414-018-1984-4](https://doi.org/10.1007/s00414-018-1984-4).
- Feng R, Zhao Y, Chen S, Li Q, Fu Y, Zhao L, Zhou Y, Zhang L, Mei X, Shi M.** 2020. Genetic analysis of 50 Y-STR loci in Dong, Miao, Tujia, and Yao populations from Hunan. *International Journal of Legal Medicine* **134**(3):981–983 DOI [10.1007/s00414-019-02115-z](https://doi.org/10.1007/s00414-019-02115-z).
- Guo F, Li J, Chen K, Tang R, Zhou L.** 2017b. Population genetic data for 27 Y-STR loci in the Zhuang ethnic minority from Guangxi Zhuang autonomous region in the south of China. *Forensic Science International: Genetics* **27**:182–183 DOI [10.1016/j.fsigen.2016.11.009](https://doi.org/10.1016/j.fsigen.2016.11.009).
- Guo F, Li J, Wei T, Ye Q, Chen Z.** 2017a. Genetic variation of 17 autosomal STR loci in the Zhuang ethnic minority from Guangxi Zhuang Autonomous Region in the south of China. *Forensic Science International: Genetics* **28**:e51–e52 DOI [10.1016/j.fsigen.2017.03.015](https://doi.org/10.1016/j.fsigen.2017.03.015).
- Hallenberg C, Nielsen K, Simonsen B, Sanchez J, Morling N.** 2005. Y-chromosome STR haplotypes in Danes. *Forensic Science International* **155**(2–3):205–210 DOI [10.1016/j.forsciint.2004.12.019](https://doi.org/10.1016/j.forsciint.2004.12.019).
- He G, Chen P, Zou X, Chen X, Song F, Yan J, Hou Y.** 2017. Genetic polymorphism investigation of the Chinese Yi minority using PowerPlex® Y23 STR amplification system. *International Journal of Legal Medicine* **131**(3):663–666 DOI [10.1007/s00414-017-1537-2](https://doi.org/10.1007/s00414-017-1537-2).
- Jobling MA, Tyler-Smith C.** 1995. Fathers and sons: the Y chromosome and human evolution. *Trends in Genetics* **11**(11):449–456 DOI [10.1016/S0168-9525\(00\)89144-1](https://doi.org/10.1016/S0168-9525(00)89144-1).
- Jobling MA, Tyler-Smith C.** 2003. The human Y chromosome: an evolutionary marker comes of age. *Nature Reviews Genetics* **4**(8):598–612 DOI [10.1038/nrg1124](https://doi.org/10.1038/nrg1124).
- King TE, Jobling MA.** 2009. Founders, drift, and infidelity: the relationship between Y chromosome diversity and patrilineal surnames. *Molecular Biology and Evolution* **26**(5):1093–1102 DOI [10.1093/molbev/msw022](https://doi.org/10.1093/molbev/msw022).
- Kumar S, Stecher G, Tamura K.** 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33**(7):1870–1874 DOI [10.1093/molbev/msw054](https://doi.org/10.1093/molbev/msw054).
- Kwak KD, Jin HJ, Shin DJ, Kim JM, Roewer L, Krawczak M, Tyler-Smith C, Kim W.** 2005. Y-chromosomal STR haplotypes and their applications to forensic and population studies in east Asia. *International Journal of Legal Medicine* **119**:195–201 DOI [10.1007/s00414-004-0518-4](https://doi.org/10.1007/s00414-004-0518-4).
- Liu Y, Zhang H, He G, Ren Z, Zhang H, Wang Q, Ji J, Yang M, Guo J, Yang X.** 2020. Forensic features and population genetic structure of Dong, Yi, Han, and Chuanqing human populations in Southwest China inferred from insertion/deletion markers. *Frontiers in Genetics* **11**:360 DOI [10.3389/fgene.2020.00360](https://doi.org/10.3389/fgene.2020.00360).
- Lu D, Lou H, Yuan K, Wang X, Wang Y, Zhang C, Lu Y, Yang X, Deng L, Zhou Y, Feng Q, Hu Y, Ding Q, Yang Y, Li S, Jin L, Guan Y, Su B, Kang L, Xu S.** 2016. Ancestral origins and genetic history of tibetan highlanders. *American Journal of Human Genetics* **99**(3):580–594 DOI [10.1016/j.ajhg.2016.07.002](https://doi.org/10.1016/j.ajhg.2016.07.002).
- Luo Y, Wu Y, Qian E, Wang Q, Wang Q, Zhang H, Wang X, Zhang H, Yang M, Ji J.** 2019. Population genetic analysis of 36 Y-chromosomal STRs yields comprehensive insights into the forensic features and phylogenetic relationship of Chinese Tai-Kadai-speaking Bouyei. *PLOS ONE* **14**:e0224601 DOI [10.1371/journal.pone.0224601](https://doi.org/10.1371/journal.pone.0224601).
- Luo L, Yao L, Chai S, Zhang H, Li M, Yu J, Hu X, Li C, Bian Y, Chen P.** 2021. Forensic characteristics and population construction of two major minorities from southwest China revealed by a novel 37 Y-STR loci system. *Royal Society Open Science* **8**(7):210447 DOI [10.1098/rsos.210447](https://doi.org/10.1098/rsos.210447).

- Ma J.** 2010. Population census office under the state council department of population and employment statistics national bureau of statistics. In: *Tabulation on the 2010 Population Census of People's Republic of China*. Beijing, China: China Statistics Press.
- 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 DOI [10.1073/pnas.70.12.3321](https://doi.org/10.1073/pnas.70.12.3321).
- Pickrahn I, Muller E, Zahrer W, Dunkelmann B, Cemper-Kiesslich J, Kreindl G, Neuhuber F.** 2016. Yfiler plus amplification kit validation and calculation of forensic parameters for two Austrian populations. *Forensic Science International: Genetics* **21**(8):90–94 DOI [10.1016/j.fsigen.2015.12.014](https://doi.org/10.1016/j.fsigen.2015.12.014).
- Ren Z, Feng Y, Zhang H, Wang Q, Yang M, Liu Y, Le C, Wang J, Huang J.** 2022. Genetic analysis of the mitochondrial DNA control region in Tai-Kadai-speaking Dong population in southwest China. *Annals of Human Biology* **49**:1–7 DOI [10.1080/03014460.2022.2131334](https://doi.org/10.1080/03014460.2022.2131334).
- Shu L, Li L, Yu G, Yu B, Liu Y, Li S, Jin L, Yan S.** 2015. Genetic analysis of 17 Y-STR loci in Han, Dong, Miao and Tujia populations from Hunan province, central-southern China. *Forensic Science International: Genetics* **19**:250–251 DOI [10.1016/j.fsigen.2015.07.007](https://doi.org/10.1016/j.fsigen.2015.07.007).
- Sofie C, Michiel VdH, Lisa V, Larmuseau MHD, Ronny D.** 2019. A game of hide and seq: identification of parallel Y-STR evolution in deep-rooting pedigrees. *European Journal of Human Genetics: EJHG* **27**(4):637–646 DOI [10.1038/s41431-018-0312-2](https://doi.org/10.1038/s41431-018-0312-2).
- Song Z, Wang Q, Zhang H, Tang J, Wang Q, Zhang H, Yang M, Ji J, Ren Z, Wu Y, Huang J.** 2021. Genetic structure and forensic characterization of 36 Y-chromosomal STR loci in Tibeto-Burman-speaking Yi population. *Molecular Genetics & Genomic Medicine* **9**(2):e1572 DOI [10.1002/mgg3.1572](https://doi.org/10.1002/mgg3.1572).
- Sun H, Su K, Fan C, Long F, Liu Y, Sun J, Mo X, Ge Y, Zhang L, Zhai L, Li W, Yin C, Li S.** 2019. Y-STRs' genetic profiling of 1953 individuals from two Chinese Han populations (Guizhou and Shanxi). *Forensic Science International: Genetics* **38**(6942):e8–e10 DOI [10.1016/j.fsigen.2018.10.011](https://doi.org/10.1016/j.fsigen.2018.10.011).
- Tang J, Yang M, Wang X, Wang Q, Wang Q, Zhang H, Qian E, Zhang H, Ji J, Ren Z.** 2020. Genetic structure and forensic characterisation of 36 Y-chromosomal STR loci in Hmong-Mien-speaking Miao population. *Annals of Human Biology* **47**(6):541–548 DOI [10.1080/03014460.2020.1788159](https://doi.org/10.1080/03014460.2020.1788159).
- Thompson JM, Ewing MM, Frank WE, Pogemiller JJ, Nolde CA, Koehler DJ, Shaffer AM, Rabbach DR, Fulmer PM, Sprecher CJ, Storts DR.** 2013. Developmental validation of the PowerPlex® Y23 system: a single multiplex Y-STR analysis system for casework and database samples. *Forensic Science International: Genetics* **7**(2):240–250 DOI [10.1016/j.fsigen.2012.10.013](https://doi.org/10.1016/j.fsigen.2012.10.013).
- Wang F, Song F, Song M, Luo H, Hou Y.** 2022. Genetic structure and paternal admixture of the modern Chinese Zhuang population based on 37 Y-STRs and 233 Y-SNPs. *Forensic Science International: Genetics* **58**(43):102681 DOI [10.1016/j.fsigen.2022.102681](https://doi.org/10.1016/j.fsigen.2022.102681).
- Wang C-Z, Su M-J, Li Y, Chen L, Jin X, Wen S-Q, Tan J, Shi M-S, Li H.** 2019. Genetic polymorphisms of 27 Yfiler® Plus loci in the Daur and Mongolian ethnic minorities from Hulunbuir of Inner Mongolia Autonomous region, China. *Forensic Science International: Genetics* **40**(4):e252–e255 DOI [10.1016/j.fsigen.2019.02.003](https://doi.org/10.1016/j.fsigen.2019.02.003).
- Wu TD.** 1993. Dong nationality is a developed indigenous nationality in Baiyue. *Guizhou Ethnic Studies* **1**:66–74.

- Xie M, Song F, Li J, Lang M, Luo H, Wang Z, Wu J, Li C, Tian C, Wang W.** 2019. Genetic substructure and forensic characteristics of Chinese Hui populations using 157 Y-SNPs and 27 Y-STRs. *Forensic Science International: Genetics* 41(2):11–18 DOI 10.1016/j.fsigen.2019.03.022.
- Yanmei Y, Tao G, Yubao Z, Chunjie X, Bifeng C, Shi L, Bingying X, Qiang J, Qinyong Z, Wen Z, Shengjun L, Shengjie N.** 2010. Genetic polymorphism of 11 Y-chromosomal STR loci in Yunnan Han Chinese. *Forensic Science International: Genetics* 4(2):e6769 DOI 10.1016/j.fsigen.2009.06.002.
- Yin C, He Z, Wang Y, He X, Zhang X, Xia M, Zhai D, Chang K, Chen X, Chen X, Chen F, Jin L, Li S.** 2022. Improving the regional Y-STR haplotype resolution utilizing haplogroup-determining Y-SNPs and the application of machine learning in Y-STR haplogroup prediction in a forensic Y-STR database: a pilot study on male Chinese Yunnan Zhaoyang Han population. *Forensic Science International: Genetics* 57(5):102659 DOI 10.1016/j.fsigen.2021.102659.
- Yin C, Su K, He Z, Zhai D, Guo K, Chen X, Jin L, Li S.** 2020. Genetic reconstruction and forensic analysis of Chinese Shandong and Yunnan Han populations by co-analyzing Y chromosomal STRs and SNPs. *Genes (Basel)* 11(7):743 DOI 10.3390/genes11070743.
- Zhou H, Ren Z, Zhang H, Wang J, Huang J.** 2016. Genetic profile of 17 Y chromosome STRs in the Guizhou Han population of southwestern China. *Forensic Science International: Genetics* 25: e6–e7 DOI 10.1016/j.fsigen.2016.05.010.
- Zhu B, Shen C, Qian G, Shi R, Dang Y, Zhu J, Huang P, Xu Y, Zhao Q, Ma J, Liu Y.** 2006. Genetic polymorphisms for 11 Y-STRs haplotypes of Chinese Yi ethnic minority group. *Forensic Science International* 158(2–3):229–233 DOI 10.1016/j.forsciint.2005.05.031.
- Zhu B, Shen C, Xun X, Yan J, Deng Y, Zhu J, Liu Y.** 2007. Population genetic polymorphisms for 17 Y-chromosomal STRs haplotypes of Chinese Salar ethnic minority group. *Legal Medicine* 9(4):203–209 DOI 10.1016/j.legalmed.2007.01.003.
- Zhu B, Wu Y, Shen C, Yang T, Deng Y, Xun X, Tian Y, Yan J, Li T.** 2008. Genetic analysis of 17 Y-chromosomal STRs haplotypes of Chinese Tibetan ethnic group residing in Qinghai province of China. *Forensic Science International* 175(2–3):238–243 DOI 10.1016/j.forsciint.2007.06.012.