



A forensic population database of autosomal STR and X-STR markers in the Qiang ethnic minority of China

Zefei Wang^{a,1}, Mengyuan Song^{b,1}, Qiang Lyu^c, Jun Ying^d, Qian Wu^e, Feng Song^a, Lanrui Jiang^a, Xiaowen Wei^a, Shuangshuang Wang^a, Fei Wang^a, Yuxiang Zhou^a, Xingbo Song^{b,**}, Haibo Luo^{a,*}

^a Department of Forensic Genetics, West China School of Basic Medical Sciences & Forensic Medicine, Sichuan University, Chengdu, 610041, China

^b Department of Laboratory Medicine, West China Hospital, Sichuan University, Chengdu, 610041, China

^c Department of Clinical Laboratory, People's Hospital of Beichuan Qiang Autonomous County, Beichuan, 622750, China

^d Department of Clinical Laboratory, Santai People's Hospital, Santai, 621100, China

^e Department of Clinical Laboratory, Renmin Hospital of Wuhan University, Wuhan, 430060, China

ARTICLE INFO

Keywords:

Population genetics
Short tandem repeat
X chromosome
Autosomal chromosome
Allele frequency
Interpopulation studies

ABSTRACT

The Qiang ethnic group is one of the oldest ethnic groups in China and is the most active ethnic group among all the populations along the Tibetan-Yi corridor. They have had a profound impact nationally and internationally. The paternal and maternal genetic feature of the Qiang ethnic group has been revealed, leaving the question of the genetic characteristics from autosomes and X chromosome not answered. The aim of this study was to explore the potential of 36 A-STR (Microreader™ 36A ID System) and 19 X-STR (Microreader™ 19X System) for application in the Qiang population and to elucidate their genetic diversity in southwest China. The cumulative probability of exclusion (CPE) for autosomal STRs is $1-1.3814 \times 10^{-15}$ and the mean paternity exclusion chance (MEC) for X-STRs is $1-1.7323 \times 10^{-6}$. Forensic parameters suggest that the STRs analyzed here are well-suited for forensic applications. The results of phylogenetic, inter-population differentiation, and principal coordinates analysis (PCoA) indicate that the Qiang people have extensive connections with ethnic minorities in China, supporting the view that the Qiang people are the oldest group in the entire Sino-Tibetan language family. The Qiang appeared genetically more associated with most ethnic groups in China, especially the Han. The calculation of random matching probability (RMP) was improved by F_{st} correction of allele frequencies to make RMP more accurate and reasonable. This study can fill in the gaps in the Qiang STR

Abbreviations: STR, Short tandem repeat; A-STRs, Autosomal short tandem repeats; X-STRs, X chromosome short tandem repeats; Y-STRs, Y chromosome short tandem repeats; RMP, Random matching probability; LR, Likelihood ratio; CODIS, Combined DNA Index System; PD, Power of discrimination; MP, Matching probability; HO, Observed heterozygosities; HE, Expected heterozygosities; PIC, Polymorphism information content; TPI, Typical paternity index; PE, Power of exclusion; HWE, Hardy-Weinberg equilibrium; LD, Linkage disequilibrium; PCoA, Principal Co-ordinates Analysis; CPE, Combined power of exclusion; MDS, Multidimensional scaling; MEC, mean paternity exclusion chance.

* Corresponding author. Department of Forensic Genetics, West China School of Basic Medical Sciences & Forensic Medicine, Sichuan University, No. 3-16 Renmin South Road, Chengdu, China.

** Corresponding author. Department of Laboratory Medicine, West China Hospital, Sichuan University, No. 37 Guoxue Alley, Wuhou District, Chengdu, Sichuan, China.

E-mail addresses: songxingbo1981@163.com (X. Song), luohaibo@263.net (H. Luo).

¹ These authors contributed equally to the study.

<https://doi.org/10.1016/j.heliyon.2023.e21823>

Received 29 May 2023; Received in revised form 28 October 2023; Accepted 30 October 2023

Available online 7 November 2023

2405-8440/© 2023 Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

reference database, providing valuable frequency data for forensic applications and evidence for the Qiang's genetic pattern as an important ancestral position in the Sino-Tibetan populations.

1. Introduction

The Qiang people, an ancient ethnic group, live mainly in the mountains near the Min River. The name "Qiang" first appeared on oracle bones to record the rich interactions and wars between the Han and Qiang people during the Shang Dynasty [1]. The Qiang people have been migrating throughout their long history, ranging from Central Asia in the west to Shandong in the east, and from the Brahmaputra Rivers in the southwest to the Tibetan-Yi Corridor in the south [2]. The Tibetan-Yi corridor is a region rich in cultural heritage and demographic diversity due to a variety of factors, including geographic and historic reasons [3]. The Qiang people are the most active ethnic group among all the populations along the Tibetan-Yi corridor. They have had a profound impact on its neighboring ethnic groups, which is reflected in the production of stone tools, pottery, bronze and other artefactual tools, ornamental features, habitation, clothing, and hair ornaments [4]. Their place of residence is also referred to as the Tibetan-Qiang-Yi corridor [4].

The Qiang have close ties with more than a dozen of ethnic groups in the southwest of China, some of which are believed to share a common ancestor with the Qiang, while others have received a considerable impact from the Qiang. During the Western Jin Dynasty, conflicts between the Qiang, Han, and Hu people were frequent due to the government's weak control over the Min River region [1]. By the Tang dynasty, with the rise of Tubo's power, the Qiang region became a territory long contested by the Tang and Tubo. Alongside the warfare, the Qiang embraced a great deal of Tibetan culture. In the late Ming Dynasty, many Han Chinese began to move into Qiang territory, bringing with them Han production techniques. In the mid-seventeenth century, the Qing court strengthened its centralisation and the Qiang region gradually entered a feudal landlord economy, and exchanges between the Qiang and Han peoples were further strengthened.

Beichuan Qiang Autonomous County locates in the western part of Mianyang City and the northwestern part of the Tibetan-Yi corridor. Beichuan Qiang Autonomous County is the only Qiang autonomous county in China, located in the northwest part of the Sichuan Basin. It went through a big earthquake in 2008, which was known as "the Wenchuan earthquake". According to the 7th official population consensus [5], the total population of Beichuan Qiang Autonomous County is about 230,000, of which 86,194 are Qiang, accounting for 37.3 % of the total population [1].

Current research on the Qiang ethnic group mainly focuses on history and culture, while studies on the genetic characteristics of the Qiang population have mainly focused on the Y chromosome [6]. Short tandem repeat (STR) polymorphisms play an important role in forensic genetic identification and the study of population evolutionary history [7]. Autosomal short tandem repeats (A-STRs) are the most commonly used STR markers in personal forensic identification and kinship testing. China is a multi-ethnic country with a large population. The differences in gene frequencies among the various ethnic groups and the data generated by the large population are a constant challenge for forensic applications. Random matching probability (RMP) is always calculated using the allele frequencies of combinations of multiple populations within a larger group such as a province or even a country. A broad combination may contain multiple populations with a more even distribution of gene frequencies. The calculated RMP and likelihood ratio (LR) using non-matched allele frequency reference may be unfavorable to the suspect.

To further trace the Qiang tribe and provide a reference for forensic applications, we constructed an STR database of the Qiang including A-STRs and X-STRs. It is beneficial for forensic practice, as well as studying the genetic relationship and origin between the Sichuan Qiang population and other Chinese reference populations. X-STRs can complement the analysis of A-STRs, e.g. in cases of paternity where the information obtained from standard autosomal markers is uncertain, and in cases of incest where possible related fathers are distinguished [8].

2. Materials and METHODS

2.1. Ethical requirements and sample collection

Blood samples were collected from 421 healthy, unrelated, and adult voluntary citizens (333 male and 88 female) living in Beichuan Qiang Autonomous County for three generations. Its approximate location on the map and the extent of the Tibetan-Yi Corridor are shown in [Supplementary Fig. 1](#). A-STR and X-STR typing were performed using 205 male samples together, and X-STR typing was carried out in an additional cohort of 128 male and 88 female samples. This study was approved by the Biomedical Research Ethics Committee of West China Hospital, Sichuan University (2022882). All blood samples were collected following the international ethics rules: signed consent from all donors and further securing of anonymity were provided. According to the decision of the local ethics committee, all DNA samples isolated from these blood samples were used for research purposes only.

2.2. DNA extraction, PCR amplification, and genotyping

DNA of all samples was extracted using the Chelex-100 method [9]. The supernatant was aspirated after centrifugation at 12,000 rpm for quantification. DNA was quantified with NanoDrop 2000c (Thermo Fisher Scientific, Waltham, MA, USA) according to the manufacturer's instructions. DNA samples were diluted to the concentration of 1 ng/ μ l for PCR according to the recommended protocol for the Microreader™ 36A ID System (Beijing Micro-reader Genetics, Beijing, China) and Microreader™ 19X System (Beijing

Micro-reader Genetics, Beijing, China) PCR Amplification Kit. The Microreader™ 36A ID System uses a 6-color fluorescent marker with multiplex amplification to detect 37 STR motifs (i.e., TH01, D5S818, D21S11, D18S51, D6S1043, D15S659, D6S477, D3S1358, D13S317, D7S820, D16S539, CSF1PO, Penta D, D8S1132, D7S3048, D2S441, vWA, D8S1179, TPOX, Penta E, D14S608, D4S2366, D3S3045, D19S433, D22S1045, D2S1338, FGA, D5S2500, D10S1435, D18S535, D1S1656, D12S391, D10S1248, SE33, D19S253, D11S2368, DYS391), one sex locus Amelogenin and one Y Indel locus (Rs2032678). The Microreader™ 19X System uses a 5-color fluorescent marker with multiplex amplification to detect Amelogenin and 19 STR motifs (i.e., DXS6795, DXS6803, DXS6807, DXS9907, DXS7423, GATA172D05, DXS101, DXS9902, DXS7133, DXS6810, GATA31E08, DXS6800, DXS981, DXS10162, DXS6809, GATA165B12, DXS10079, DXS10135, HPRTB).

1 ng template was added to 10 µl of Master Mix III and 5 µl Primer Mix, and then nuclease-free water was added up to 25 µl for Microreader™ 36A ID System. Thermal cycler conditions included an initial incubation for 5 min at 95 °C; 28 cycles of denaturation for 20 s at 94 °C, annealing/extension for 90s at 59 °C and final extension for 60 min at 60 °C; and a final hold at 4 °C. Microreader™ 19X System total reaction volume was 25 µl, including 10 µl of Microreader™ 2.5 × Buffer, 5 µl of Microreader™ 19 X-STR 5 × Primer Mix, 0.1 µl Taq, and 0.5 µl ng of template DNA, and then nuclease-free water was added up to 25 µl. The thermal cycling parameters were set as follows: preincubation at 95 °C for 5 min; 29 cycles of denaturation at 94 °C for 30 s, annealing at 59 °C for 60 s, and extension at 72 °C for 60 s; extension at 60 °C for 30 min, finally holding at 25 °C until the amplification products were removed from the PCR machine.

An Applied Biosystems 3500 Genetic Analyzer (Thermo Fisher Scientific, Waltham, MA, USA) was used for all amplification products, using a 36 cm capillary array and POP-4 polymer. 1 µl of PCR products was added to a 9 µl mixture of Hi-Di formamide and a microreader size standard. STR profiles were analyzed by GeneMapper™ ID-X v1.5 software (Applied Biosystems Foster City, CA, USA). Allele lengths were determined by comparing the sample PCR fragments with the allelic ladder provided with the kit.

2.3. Reference database

The STR data of Qiang was compared with the data of ethnic minorities in China derived from published sources as summarized in Table 1. Database I contains autosomal STR data. PCoA, heat map, and structural analysis were performed using 15 overlapping A-STRs. Both Database II and Database III are X-STR data. The difference is that the former has more populations but 7 overlapping X-STRs and the latter consists of fewer populations but uses all 19 overlapping X-STRs. Both were subjected to PCoA and heat map analysis.

Table 1

Datasets for various analyses in this study.

Dataset	Analysis	Minorities	Data
Dataset I	PCoA, heatmap, Structure	Hui, Tibetan and Uyghur [10] Mongol, Kyrgyz and Uzbek [11] Yi [12] Han [13] Du Long and Lisu [14] Hui [15] Xibe [16] Nakhi and Yi [17] Han [18] Uyghur [19] Mongolian [20] Myanmar [21] Han, Nu and Tibetan [22]	A-STRs
Dataset II	PCoA, heatmap	Han and Hui [23] Han [24] Mongolian [25] Han [26] Miao [27] Daur and Oroqen [28] Uyghur and Tibetan [29] Zhuang and Mulao [30] Miao [31] Tibetan [32] Yi [33] Kyrgyz and Han [34]	X-STRs
Dataset III	PCoA, heatmap	Korean [35] Han [36] Gelao [36] Miao [36] Kazakh [37] Tujia [37] Han [38]	X-STRs

2.4. Statistical analysis

Statistical parameters of forensic interest include power of discrimination (PD), matching probability (MP), observed (HO) and expected (HE) heterozygosities, polymorphism information content (PIC), typical paternity index (TPI), power of exclusion (PE) and allele frequencies were determined using modified STRAF 1.0.5 [39] and [ChrX-STR.org](https://www.STR.org) 2.0 [40]. Hardy–Weinberg equilibrium (HWE) and pair-wise linkage disequilibrium (LD) in female samples were calculated by the software of Arlequin 3.5.2 [41]. We used the Principal Coordinate Analysis function in GenALEX 6.503 [42] to perform a Principal Coordinate Analysis (PCoA). The ridge plots were achieved by *ggribdes* and *ggplot2* packages in R. The heat map is implemented by the *readxl*, *gplots*, packages in R. The geographic map was drawn by R packages *ggplot2*, *maps*, *sf*, *rnaturalearth*, and *rnaturalearthdata*.

We used 196 Qiang individuals (9 individuals were removed that contained off-ladder (OL) alleles absent in the Han Chinese database) at 21 autosomal loci (CSF1PO; D10S1248; D12S391; D13S317; D16S539; D18S51; D19S433; D21S11; D2S1338; D2S441; D3S1358; D5S818; D6S1043; D7S820; D8S1179; FGA; PentaD; PentaE; TH01; TPOX; vWA) to analyze the effect of F_{st} correction on RMP. The exact test for population differentiation (F_{st}) was carried out by Arlequin3.5.2 software [41]. The RMP calculation was performed by EuroForMix 3.0.3 [43]. It fills in the F_{st} -correction value in the settings and recalculates the corrected RMP. Afterwards, the individual RMP values were compared and the \log_{10} of the ratio was calculated. The RMP was corrected using the correction formula from David J. Balding and Richard A. Nichols [44] (in formula F represent F_{st}), the derivation of which is described in detail in the text. The two observed alleles were denoted by A and B, and their proportions in the database population were represented by p_A and p_B . Then the probability that a suspect has these alleles given that the offender has them is.

For homozygotes, RMP is defined as:

$$\frac{[2F + (1 - F)p_A][3F + (1 - F)p_A]}{(1 + F)(1 + 2F)} \quad (1)$$

For heterozygotes, RMP is calculated as:

$$\frac{2[F + (1 - F)p_A][F + (1 - F)p_B]}{(1 + F)(1 + 2F)} \quad (2)$$

For structure analysis for autosomal STR, the software Structure v2.3.4 [45] was used with the under the following parameters: The evaluated number of clusters (K) ranged from 2 to 8; for each K, 10 independent runs were performed. We utilized a 100,000-iteration burn-in period followed by 50,000 iterations. Using the Structure Harvester, we assessed the number of genetic groups (K) that best fit the data. Finally, the CLUMPP_Windows.1.1.2 [46] program was used to output a mean of the permuted matrices across replicates for each K. Based on the best K, we estimated and plotted the percentage of admixture in each population with *distruct* 1.1 [47].

3. Results

3.1. Allelic frequencies and forensic parameters of A-STRs

The allele frequencies and forensic parameters of the 36 autosomal STR loci are listed in [Supplementary Table 1](#). A total of 424 alleles were observed in the studied population, with corresponding allele frequencies from 0.0025 to 0.4015. There were no allele frequencies greater than 0.5, indicating that these motifs are well heterozygous in the Qiang population and that the allele frequencies are evenly distributed. There was no significant deviation from HWE after applying the Bonferroni correction, except for locus D6S477. The deviation may be due to undersized populations, inbreeding, population substructure, or selection. Another important factor may be the small sample size, and increasing the sample size may make it no longer deviate from the HWE. P-values for the Linkage disequilibrium test between all A-STRs are shown in [Supplementary Table 2](#), with no significant linkage between any of the A-STRs.

PIC was in the range of 0.584 (TPOX) to 0.948 (SE33), and the most discriminating marker was SE33 with a PD value of 0.989. The observed heterozygosity varied from 0.567 (TPOX) to 0.951 (SE33), and the PE ranged from 0.253 (TPOX) to 0.900 (SE33). The CPD, CPE, and CPM were $1-7.8531 \times 10^{-44}$, $1-1.3814 \times 10^{-15}$, and 7.85314×10^{-44} , respectively. This indicated that the 36 autosomal STR loci were highly polymorphic and appropriate for individual identification and paternity testing in the studied population.

3.2. Improvement of RMP calculation by F_{st} correction

Using a non-cognate database calculation in the RMP calculation can be detrimental to the suspect, so it is more appropriate to use F_{st} correction. We tested the impact of using different reference frequencies on RMP calculations. First, we used the Sichuan Qiang allele frequency database (generated in this study) and Sichuan Han allele frequency databases, for the RMP calculation of the Qiang individuals. The selected Sichuan Han reference dataset was from a study of 2793 unrelated individuals. In China, the Han Chinese are the dominant population and its frequency database is commonly used in forensic practice. We used 196 Qiang individuals (9 individuals were removed that contained off-ladder (OL) alleles absent in the Han Chinese database) of the Qiang ethnic group to calculate the RMP using the allele frequencies of their population and that of the Sichuan Han ethnic group. Calculated using the formula in section 2.4.

Then a parameter d was calculated, corresponding to \log_{10} of the ratio between RMPs obtained from the A-STR profiles of each individual from the Qiang population when using cognate (Qiang) and non-cognate (Han) allele frequencies [48]. The results are

displayed in Table 2 and Fig. 1. Values of $d > 0$ indicate an overestimation of the weight of the evidence when using the non-cognate databases for calculations.

It can be seen that without F_{st} correction (F is set as zero), there is a slight overestimation of the weight of evidence. Applying the F_{st} correction of 0.0016 is more accurate when using the Sichuan Han database for RMP calculations, with an average d of 0.0102. However, the weight of the evidence in more than half of the cases was still overestimated, which would be magnified by the hypothesis that the suspect was related to the offender [44]. It would be more appropriate to use 0.0108, for the risk of non-conservative estimations to become almost negligible, with only 1.02 % of individuals overestimating the weight of the evidence and an average RMP one order of magnitude higher, which is acceptable in practice. For the Qiang, the use of 0.0016 is appropriate in the general case. However, when it comes to family exclusion, it is feasible to add sex chromosome markers or use a larger F_{st} . When a population has more than one recommended F_{st} , it is most appropriate to select the F_{st} that is within the smallest range that can contain cognate and non-cognate databases. The F_{st} used for the correction cannot be applied to all cases, so the use of a cognate database is necessary.

3.3. Population affinity revealed by A-STRs

The pairwise F_{st} values and corresponding p -values between Qiang and other reference populations (listed in Table 1) are shown in Supplementary Table 3 (using overlapping 15 A-STRs). The largest F_{st} value (0.04436) was observed between the Fujian, She and Yunnan, Du long ethnic groups, whereas the smallest F_{st} value (-0.0003) was found between Mongolians from Xinjiang and Inner Mongolia. F_{st} values are larger between ethnic groups in Yunnan and other ethnic groups due to their ethnic diversity and smaller sampling. In general, the F_{st} value increases with geographic distance.

The results of the PCoA are shown in Fig. 2A. It graphically represents the genetic relationships between the studied population and the other 19 reference populations. As can be seen from Fig. 2A, Yunnan minority groups are distributed on the edge of the figure. Most of the ethnic group cluster on the right side of the figure, which suggests a relatively low degree of differentiation. In general, there is a tendency for different ethnic groups to cluster by region and ethnicity. The Sichuan Qiang and Sichuan Tibetans are close together in the upper right quadrant of the figure. In the Y chromosome haplogroup, we studied before [50], both Qiang and Tibetan have a portion of haplogroup D. Both results illustrate the close relationship between the Qiang and Tibetans. It supports the view that the Qiang are the ancestors of both Tibetans and Han Chinese [50]. Similarly, the heatmap (Fig. 2B) was constructed based on the genetic distance of F_{st} . The Qiang are clustered with the ethnic groups in the areas around the Tibetan-Yi corridor and with the Han Chinese in various regions, which illustrates the close relationship between the Qiang and the ethnic groups in the Tibetan-Yi corridor and the extensive contact with the Han Chinese. The ethnic groups of Xinjiang and Mongolia are clustered together, while the rest of the ethnic groups are scattered in other locations, with the Dulong of Yunnan located in a separate group, a small society of only 4000 people with little interaction with modern society [51].

STRUCTURE analysis was performed using Structure v2.3.4 software to genotype 15 STRs from the Sichuan Qiang and other published ethnic groups. Each $K = 2-8$ for 10 runs; then, the optimal K was selected using Structure Harvester and the results of the calculations are displayed in Fig. 3A,B, showing that $K = 2$ is the most appropriate. However, we do not find a significant difference in the structure plot for $K = 2$ (Fig. 3C). In Fig. 3B we can see that $K = 1$ has the highest $L(K)$, which means that $K = 1$ should be more appropriate. Structure Harvester recommends the best $K = 2$ as its algorithm does not generally recommend $K = 1$ [46]. These populations tend to have an identical ancestor, which may be due to the geographical proximity of the selected populations. The populations we have selected are all from the Tibetan-Yi corridor or western China and all belong to the Sino-Tibetan language family, of which Qiang, Tibetan, Yi, Dulong and Burmese all belong to the Tibetan-Burmese branch, so it is reasonable to obtain this result.

3.4. Allelic frequencies and forensic parameters of X-STRs

A total of 333 male and 88 female individuals from the Sichuan Qiang group were analyzed. Allele frequencies for males, females, and pooled are presented in Supplementary Table 4, Supplementary Table 5 and Supplementary Table 6 respectively. The F_{st} and corresponding p values of 19 X-STRs between females and males in Sichuan Qiang were presented in Supplementary Table 7. No gender differentiation is identified after Bonferroni correction ($p = 0.05/19 = 0.0026$). Lower P values were observed for DXS6800, which may be due to the low polymorphism of DXS6800 and the small number of female samples. A total of 180 alleles were observed at the 19 X-STR loci, and the allele numbers ranged from 5 at DXS7133 and DXS7423 to 22 at DXS10135. The most frequent allele observed was allele 16 at locus DXS6800, with a frequency of 0.8490. Bonferroni-corrected HWE exact test (Supplementary Table 8) ($p = 0.05/19 = 0.0026$) showed no significant deviation for the 19 X-STR motifs in the female sample except for DXS7423.

Table 2

The results of d values obtained by applying different F_{st} adjustment when using Sichuan Han allele frequencies for RMP calculation. F_{st} values were recommended by Xuan Dai et al. for the Chinese populations [49].

Table 2. Distribution of values.F	$d > 0$	$d > 0.5$	$d > 1$	$d > 2$	Mean d	\pm SD
0	74.49 %	35.20 %	12.24 %	1.02 %	0.3309	0.5848
0.0016	54.08 %	18.37 %	01.02 %	0.00 %	0.0102	0.5102
0.0108	1.02 %	00.51 %	0.00 %	0.00 %	-1.1970	0.5878
0.0170	00.51 %	00.51 %	0.00 %	0.00 %	-1.8538	0.6630

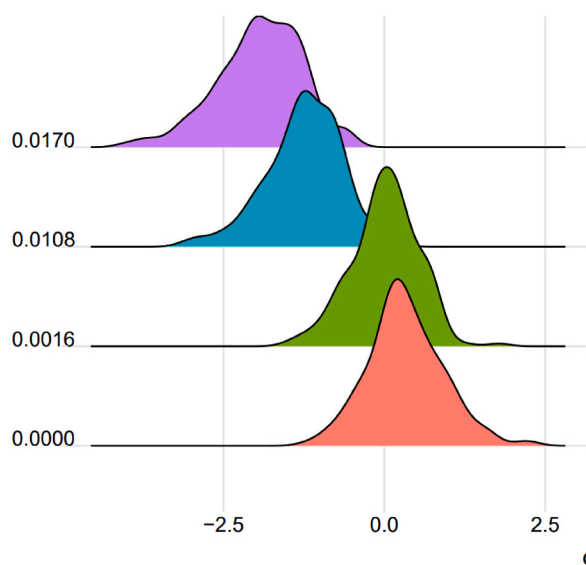


Fig. 1. Distribution of d values after correction using different F_{st} . Colors indicate different F_{st} adjustment.

In the linkage group test (Supplementary Table 9), after Bonferroni correction, no evidence of detectable linkage disequilibrium was found for the same pair of markers except for DXS10079 and DXS6800 ($p = 0.000292$). The haplotype distributions of Sichuan Qiang population are presented in Supplementary Table 10. A total of 28 haplotypes of LG DXS10079-DXS6800 are observed in 333 males, and the most common haplotypes are 16–20(0.2583). The difference is that in a study of Han Chinese in Beijing, China [52] no evidence of detectable linkage disequilibrium was found in DXS10079 and DXS6800. The different results obtained by the Sichuan Qiang and the Beijing Han on the linkage of these two STRs may be due to the small size of our sample or to the unique genetic pattern of the Qiang, which needs to be demonstrated by future experiments with larger samples.

Forensic statistical parameters are listed in Supplementary Table 7. We consider these STRs as 18 loci to calculate the parameters. DXS10135, with the highest PIC (0.9097), HD (0.9176), PDF (0.9868) and PDM (0.9158), had the highest polymorphism. The lowest PIC (0.3052), HD (0.3533), PDF (0.5335) and PDM (0.3527) were observed at DXS7133. Interestingly, we found that the polymorphism of DXS6800 was low in many ethnic groups in China, but not in some other countries. For example, the polymorphism information content (PIC) is only 0.2633 and 0.2651 [38] in the Qiang of Sichuan and the Han of Yunnan, but is high in Germany (0.7314) [53] and Austria (0.7322) [54]. The combined PDF and PDM are 2.0622×10^{-17} and $1-9.9450 \times 10^{-11}$, respectively. The combined mean paternity exclusion chance calculated by the formula of $MEC_{Krüger}$ (mean paternity exclusion chance), $MEC_{Kishida}$, $MEC_{Desmarais}$, and $MEC_{Desmarais Duo}$ are $1-4.7542 \times 10^{-7}$, $1-7.2687 \times 10^{-11}$, $1-7.3127 \times 10^{-11}$, and $1-9.2682 \times 10^{-8}$, respectively. The forensic parameters described above indicate that the 19 X-STRs in the Microreader™ 19X PCR amplification kit are highly polymorphic and informative in the Sichuan Qiang population and can be used as a powerful tool for forensic complex kinship identification.

3.5. Population affinity revealed by X-STRs

Current research on X-STRs has focused on the analysis of linkage groups, and such linkage-free X-STR kits are less commonly used. Therefore, to explore the genetic homogeneity and heterozygosity of different ethnic groups in X-STR, two databases were chosen for comparison, the first using the Sichuan Qiang and 15 other ethnic groups using overlapping 7 X-STRs, the second group using the Sichuan Qiang and 7 other ethnic groups using overlapping 19 X-STRs.

In the first group we calculated the F_{st} genetic distance between the Sichuan Qiang and 15 other ethnic groups using overlapping 7 X-STRs. The pairwise F_{st} values are listed in Supplementary Table 11. The closest genetic distance (F_{st} genetic distance: 0.002) appeared between Sichuan Qiang and Guizhou Miao and between Sichuan Qiang and Southern Han. Meanwhile, the Qiang also possessed a close genetic distance ($F_{st} < 0.05$) from the Heilongjiang Daur, Guizhou Han, Sichuan Yi, and Sichuan Han.

PCoA is shown in Fig. 4A, with most ethnic groups showing clustering by geographic location. Guangxi Mulao and Guangxi Zhuang appear in the lower right corner, which is far from Sichuan Qiang population. The Xinjiang minority appears in the middle and lower part of the figure, of which Kyrgyz ethnic group has the most genetic affinity with Sichuan Qiang. Sichuan and Guizhou minorities are clustered surrounding Sichuan Qiang population. Tibetans in Tibet are alone in the lower left, quite separate from Qiang population. The Qiang and the southern Han Chinese are located in the center of the cluster above. Mansha Jia et al. [55] revealed the genetic homology of Oroqen, Mongolian, Kirk, and Uyghur, as evidenced by our study. Heatmap (Fig. 4B) were constructed according to F_{st} genetic distance. The different ethnic groups are distributed approximately according to geographic location, with the Qiang clustering with the south-western Chinese ethnic groups. The Heilongjiang Daur are closer to the south-western ethnic groups. A similar structure was found in previous studies of the autosomes of the Daur ethnic groups, suggesting that the Daur are genetically more closely related

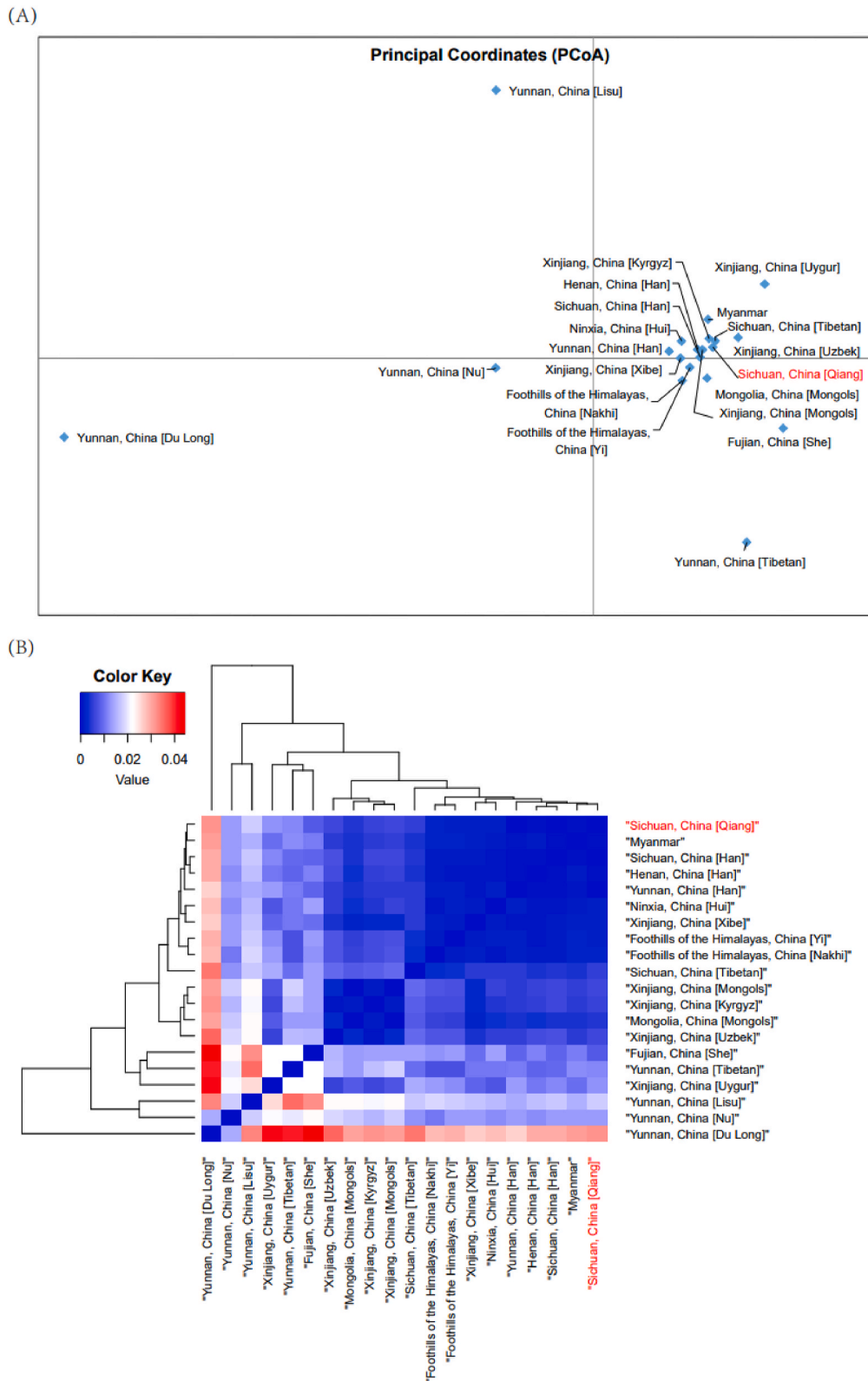


Fig. 2. Genetic relationships estimated by genetic distances of autosomes (A-STRs) between Qiang (red) and the reference database are represented in principal coordinates analysis (PCoA) (A) and heatmap (B). In general, there is a tendency for different ethnic groups to cluster by region and ethnicity.

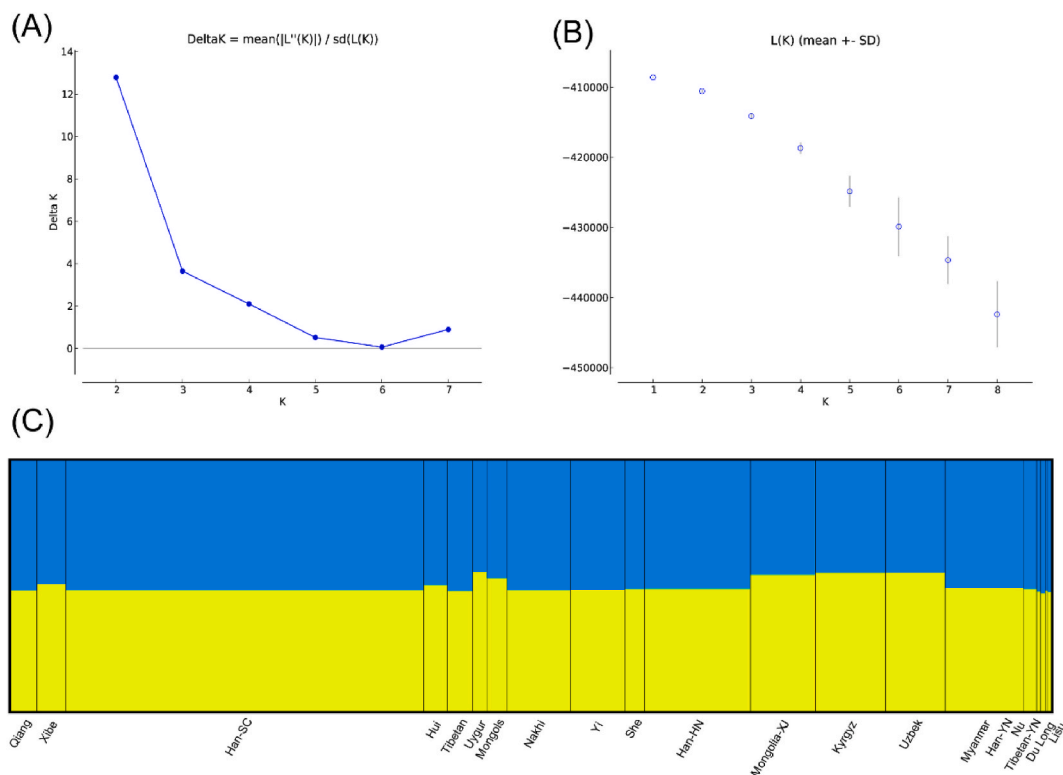


Fig. 3. Results of STRUCTURE analysis. (A) Illustrates maximum of delta K (B) Illustrates maximum of L(K) (C) Bar plot representing structure analysis of Qiang in comparison to 19 other populations based on 15 autosomal STRs when $k = 2$.

to the western Chinese groups. The ethnic groups of Xinjiang cluster together with the Tibetans of Sichuan, and the ethnic groups of Guangxi cluster with each other. The genetic relationship between Tibetans and Han Chinese is different from A-STR, where Qiang are closely genetically related to Tibetans, but in X-STR the Qiang are distantly genetically related to both Tibetans in Sichuan and Tibetans in Tibet.

The second group is similar to the first group, the pairwise F_{st} values are listed in [Supplementary Table 12](#). The closest genetic distance (F_{st} genetic distance: 0.002) appeared between Sichuan Qiang and Guizhou Han and between Sichuan Qiang and Hubei Tujia. In the heatmap ([Fig. 5B](#)) the Qiang, Tujia and Han form a cluster, the two Guizhou minorities are clustered together and the Kazakhs are in a separate group. PCoA ([Fig. 5A](#)) also shows similar results to heatmap. Consistent with previous analyses, the Qiang are closely related to the Han. The Tujia are closely related to the Han in many studies, so the Tujia and Qiang are also clustered together.

In general, Sichuan Qiang showed a genetic affinity with Sichuan Han and extensive genetic heterogeneity with Mongols and Uyghur. At the same time, the Sichuan Qiang are genetically homogeneous with most ethnic groups in southwestern China. This is consistent with our expectations and demonstrates that the Qiang has been continuously developing and differentiating since about 1000 BCE [1], with extensive and profound influence on the history and ethnic formation of the surrounding region. Interestingly, Qiang showed a genetic affinity with Tibetans in A-STRs but was not seen in X-STRs. Different results based on the analysis of different chromosomes are also found in other ethnic groups [56,57]. However, it is also possible that this is due to the low number of overlapping X-STRs used, and more Tibetan data are needed to confirm this phenomenon.

4. Discussion

We have typed A-STR for 205 samples and X-STR for 421 samples from the Sichuan Qiang ethnic group and provided allele frequencies for the Sichuan Qiang ethnic group, which will provide a basis for application and ethnic analysis in forensic cases. Of all STRs, we found two (D6S477, DXS6800) that failed HWE after Bonferroni correction, in all cases because the homozygote exceeded expectations. DXS6800 also has a low PIC (0.2634) and a gene frequency of 0.8529 for allele 16. A similar situation is seen in many Chinese ethnic groups, but has a high PIC in Europe and Oceania, suggesting that this locus may not be suitable for forensic applications in the Chinese region.

In practice, DNA analysts generally place more emphasis on genetic profile matches, and the introduction of RMP and LR allows forensic scientists to assess the strength of evidence. When a suspect is from a small and isolated subpopulation, the population allele frequency should be used preferentially to calculate the RMP. If appropriate data are lacking, then the use of F_{st} values corrected for other populations or allele frequencies containing combinations of multiple populations should be applied. If not calibrated, the weight

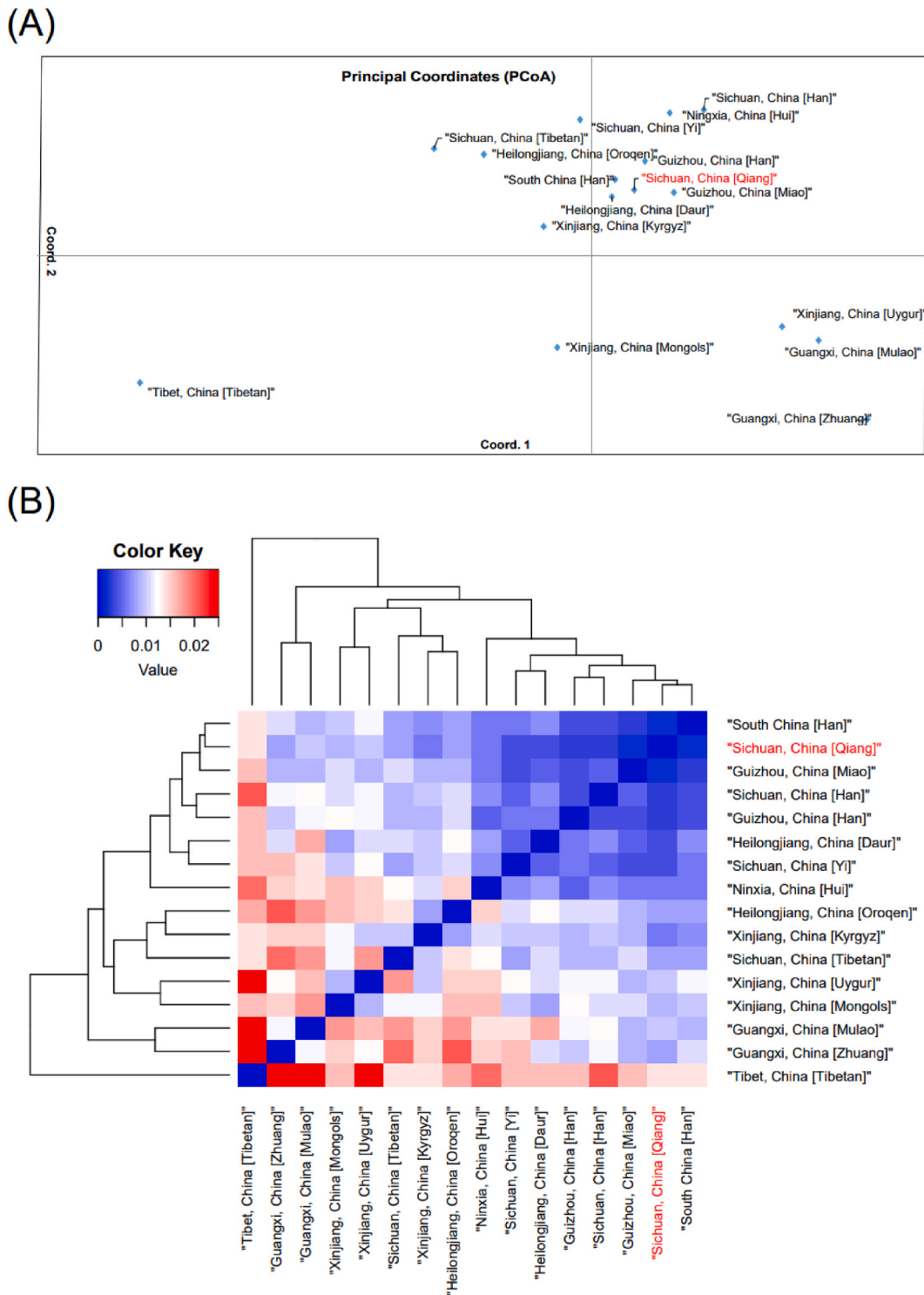


Fig. 4. Genetic relationships estimated by genetic distances of X-STRs between Qiang (red) and the reference database containing 15 ethnic groups are represented in principal coordinates analysis (PCoA) (A) and heatmap (B).

of evidence will be overestimated. For Sichuan Qiang, the use of 0.0016 for F_{st} adjustment is closest to the homology database results, but does not mitigate the overestimation of the evidence in some special cases such as when the suspect is a sibling of the offender. Using 0.0108 mitigates the overestimation of the evidence, but is somewhat overly conservative and underestimates the strength of the

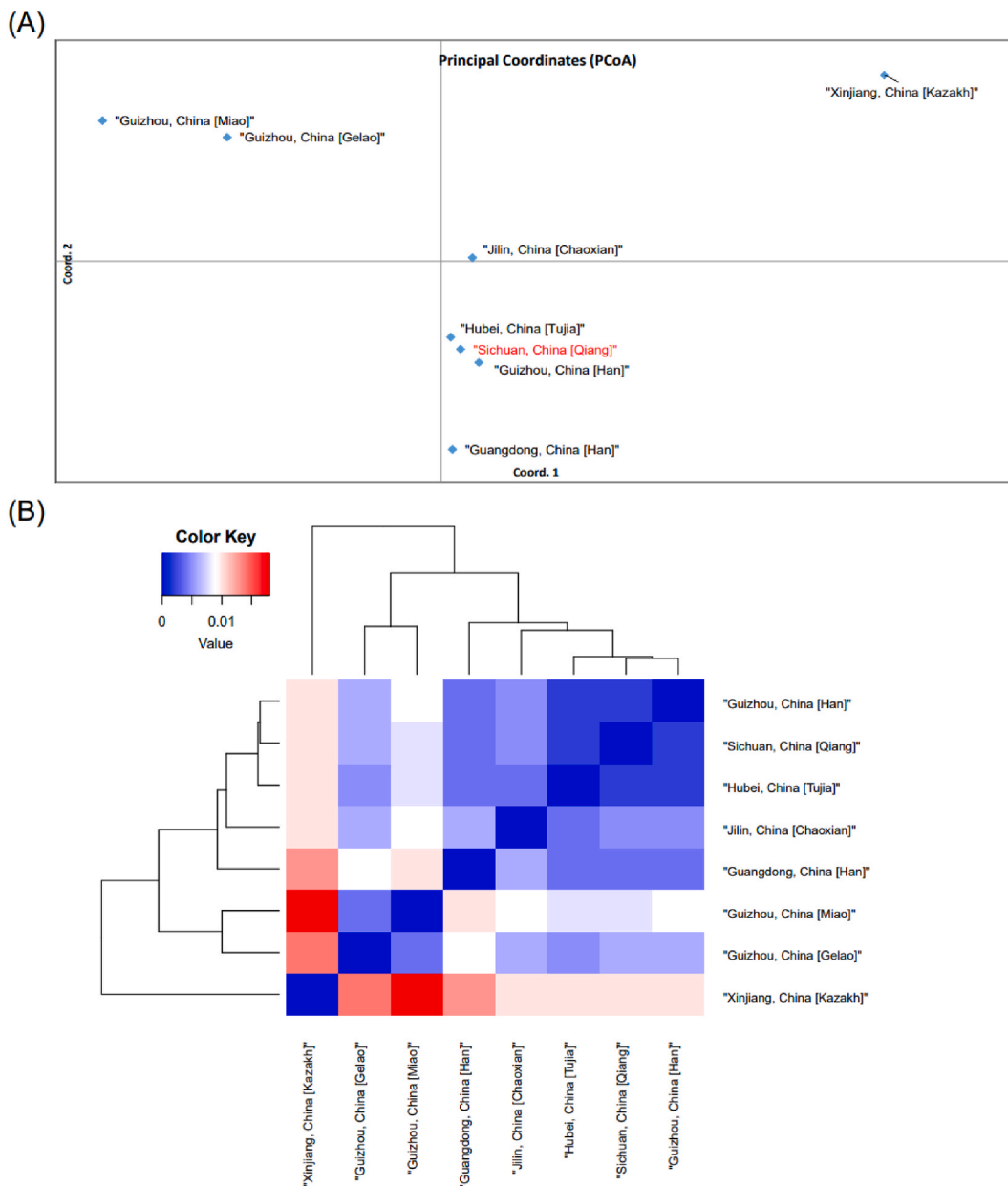


Fig. 5. Genetic relationships estimated by genetic distances of X-STRs between Qiang (red) and the reference database containing 7 ethnic groups are represented in and principal coordinates analysis (PCoA) (A) and heatmap (B).

evidence across the board. Therefore, the use of a cognate database is more accurate and comprehensive for the estimation of the weight of evidence.

For population genetic comparisons of Qiang and other ethnic groups, our genetic results show its historical complexity, which has a non-negligible role in the evolution of many ethnic groups. In the A-STRs, the Qiang show a wide range of close ties with the populations of the Tibetan-Yi corridor. For X-STRs, the Sichuan Qiang are genetically homogeneous with most ethnic groups in southwestern China. All STR markers showed gene flow between the Qiang and Han individuals, which is consistent with previous studies of the Y chromosome. Song et al. [50] and Wu et al. [6] analyzed the structure of the Qiang on the Y chromosome in terms of the Y-STR and Y haplogroup of the Qiang, both indicating clear gene flow between the Han and the Qiang. The Di-Qiang tribe is an ancient name for both the Di and Qiang tribes, at present the Di are integrated into the Han. A number of studies of ancient DNA of the Di-Qiang people have found that the Di-Qiang contributed to the Han STR gene pool [58], as well as gene flow to the Xinjiang region and the Hexi Corridor [59]. In maternal inheritance, the results of the Qiang mtDNA study show similar results to those of the paternal lineage [60], and the diversity of the Qiang population is also mainly contributed by haplogroups prevalent in North Asia. The difference is

that genetic differentiation with Tibetans is shown on the X-STR, which is different from the A-STR and Y-STR, but there are fewer overlapping loci for comparison, so using the same kit for Tibetans would help to further investigate this phenomenon.

5. Conclusion

We have typed the 55 STRs contained in the Microreader™ 36A ID System and Microreader™ 19X System in 481 samples from Sichuan Qiang and provide allele frequencies and forensic parameters for the Qiang population in Sichuan, which will provide valuable assistance in the personal identification and paternity testing. Improving the calculation of RMP by F_{st} correction of allele frequencies can make RMP more accurate and reasonable, if non-cognate database was used for RMP calculation. To better understand the genetic structure and inter-population relationships, different STR combinations were applied in the genetic distance calculations to estimate the distant relationships between ethnic groups. Taken together, these results suggest that the Qiang are closely related to the ethnic groups in the Tibetan-Yi corridor and have few genetic differences from the Han Chinese. Our genetic results show its historical complexity, which has a non-negligible role in the evolution of many ethnic groups. The cultural and biological characteristics of this ethnic group are important for population diversity estimation, and further research on it and surrounding population is needed.

Data availability statement

Data included in article/supplementary material/referenced in article.

Fundings

This study was supported by grants from the National Natural Science Foundation of China (81772030, 81672096, 82202614). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

CRediT authorship contribution statement

Zefei Wang: Data curation, Formal analysis, Writing – original draft. **Mengyuan Song:** Data curation, Investigation, Writing – review & editing. **Qiang Lyu:** Data curation, Formal analysis. **Jun Ying:** Data curation, Formal analysis. **Qian Wu:** Data curation, Formal analysis. **Feng Song:** Conceptualization, Supervision. **Lanrui Jiang:** Data curation, Formal analysis. **Xiaowen Wei:** Software. **Shuangshuang Wang:** Conceptualization, Supervision. **Fei Wang:** Software. **Yuxiang Zhou:** Software. **Xingbo Song:** Conceptualization, Funding acquisition, Supervision, Writing – review & editing. **Haibo Luo:** Conceptualization, Funding acquisition, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Haibo Luo reports financial support was provided by National Natural Science Foundation of China. Xingbo Song reports financial support was provided by National Natural Science Foundation of China. Mengyuan Song reports financial support was provided by National Natural Science Foundation of China.

Acknowledgements

Ethics statement: This study was approved by the Biomedical Research Ethics Committee of West China Hospital, Sichuan University (2022882).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.heliyon.2023.e21823>.

References

- [1] L. Xue, · Xi, X. Wang, H. Zhou, Traditional Chinese Villages, (n.d).
- [2] Hongyun Duan, Erbo Luo, On the migration and ethnic fusion of qiang people from pre- qin to HanDynasties, *Journal of Sichuan Minzu College* 23 (2014) 42–46.
- [3] Z. Zhang, Y. Zhang, Y. Wang, Z. Zhao, M. Yang, L. Zhang, B. Zhou, B. Xu, H. Zhang, T. Chen, W. Dai, Y. Zhou, S. Shi, R. Nielsen, S.C. Li, S. Li, The Tibetan-Yi region is both a corridor and a barrier for human gene flow, *Cell Rep.* 39 (2022), <https://doi.org/10.1016/j.celrep.2022.110720>.
- [4] Jian Ye, From “Tibetan and Yi corridor” to “Tibetan-Qiang-Yi corridor” :TheHistorical role of the ancient qiang people, *Journal of Yuxi Normal University* (2014) 1–8.
- [5] <http://www.beichuan.gov.cn/public/17471/42089661.html>, (2023.7).

- [6] Q. Wu, J. Ying, C. Xue, Q. Lv, F. Song, J. Wang, X. Song, J. Chen, Genetic analysis of 29 Y-STR loci in qiang population from beichuan qiang autonomous county in China, *Leg. Med.* 45 (2020), <https://doi.org/10.1016/j.legalmed.2020.101714>.
- [7] C. Phillips, L. Devesse, D. Ballard, L. van Weert, M. de la Puente, S. Melis, V. Álvarez Iglesias, A. Freire-Aradas, N. Oldroyd, C. Holt, D. Syndercombe Court, Á. Carracedo, M.V. Lareu, Global patterns of STR sequence variation: sequencing the CEPH human genome diversity panel for 58 forensic STRs using the Illumina ForenSeq DNA Signature Prep Kit, *Electrophoresis* 39 (2018), <https://doi.org/10.1002/elps.201800117>.
- [8] A.O. Tillmar, D. Kling, J.M. Butler, W. Parson, M. Prinz, P.M. Schneider, T. Egeland, L. Gusmão, DNA commission of the international society for forensic genetics (ISFG): guidelines on the use of X-STRs in kinship analysis, *Forensic Sci Int Genet* 29 (2017), <https://doi.org/10.1016/j.fsigen.2017.05.005>.
- [9] P.S. Walsh, D.A. Metzger, R. Higuchi, Chelex® 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material, *Biotechniques* 10 (1991), <https://doi.org/10.2144/000114018>.
- [10] J. Liu, Z. Wang, G. He, M. Wang, Y. Hou, Genetic polymorphism and phylogenetic differentiation of the Huaxia Platinum System in three Chinese minority ethnicities, *Sci. Rep.* 9 (2019), <https://doi.org/10.1038/s41598-019-39794-y>.
- [11] X. Zhan, A. Adnan, Y. Zhou, A. Khan, K. Kasim, D. McNeven, Forensic characterization of 15 autosomal STRs in four populations from Xinjiang, China, and genetic relationships with neighboring populations, *Sci. Rep.* 8 (2018), <https://doi.org/10.1038/s41598-018-22975-6>.
- [12] J. Cheng, B. Song, J. Fu, X. Zheng, T. He, J. Fu, Genetic polymorphism of 19 autosomal STR loci in the Yi ethnic minority of Liangshan Yi autonomous prefecture from Sichuan province in China, *Sci. Rep.* 11 (2021), <https://doi.org/10.1038/s41598-021-95883-x>.
- [13] Z. Zhou, C. Shao, J. Xie, H. Xu, Y. Liu, Y. Zhou, Z. Liu, Z. Zhao, Q. Tang, K. Sun, Genetic polymorphism and phylogenetic analyses of 21 non-CODIS STR loci in a Chinese Han population from Shanghai, *Mol Genet Genomic Med* 8 (2020), <https://doi.org/10.1002/mgg3.1083>.
- [14] L. Yuan, Y. Ou, Q. Liao, J. Gui, X. Bai, J. Ge, J. Ye, L. Zhang, Population genetics analysis of 38 STR loci in the She population from Fujian Province of China, *Leg. Med.* 16 (2014), <https://doi.org/10.1016/j.legalmed.2014.05.008>.
- [15] X. Zhang, Y. Yang, H. Liu, W. Zuo, L. Huang, Population data and phylogenetic structure of 19 STR loci in Chinese Hui ethnic group residing in Yunnan province of China, *Leg. Med.* 56 (2022), <https://doi.org/10.1016/j.legalmed.2022.102044>.
- [16] H. Meng, Y. Guo, Q. Dong, G. Yang, J. Yan, J. Shi, B. Zhu, Autosomal-STR based genetic structure of Chinese Xibe ethnic group and its relationships to various groups, *Int. J. Leg. Med.* 130 (2016), <https://doi.org/10.1007/s00414-016-1369-5>.
- [17] G. He, Y. Su, X. Zou, M. Wang, J. Liu, S. Wang, Y. Hou, Z. Wang, Allele frequencies of 15 autosomal STRs in Chinese Nakhi and Yi populations, *Int. J. Leg. Med.* 133 (2019), <https://doi.org/10.1007/s00414-018-1931-4>.
- [18] W. Hongdan, K. Bing, S. Ning, H. Miao, Z. Bo, G. Yuxin, Z. Bofeng, L. Shixiu, Z. Zhaoshu, Evaluation of the genetic parameters and mutation analysis of 22 STR loci in the central Chinese Han population, *Int. J. Leg. Med.* 131 (2017), <https://doi.org/10.1007/s00414-016-1389-1>.
- [19] L. Yuan, H. Liu, Q. Liao, X. Xu, W. Chen, S. Hao, Genetics analysis of 38 STR loci in uygur population from southern Xinjiang of China, *Int. J. Leg. Med.* 130 (2016), <https://doi.org/10.1007/s00414-015-1270-7>.
- [20] R. Wu, R. Li, N. Wang, D. Peng, H. Li, Y. Zhang, C. Zheng, H. Sun, Genetic polymorphism and population structure of Torghut Mongols and comparison with a Mongolian population 3000 kilometers away, *Forensic Sci Int Genet* 42 (2019), <https://doi.org/10.1016/j.fsigen.2019.07.017>.
- [21] F. Guo, Genetic polymorphism of 17 autosomal STR loci in the Lahu ethnic minority from Yunnan Province, Southwest China, *Forensic Sci Int Genet* 31 (2017), <https://doi.org/10.1016/j.fsigen.2017.08.002>.
- [22] T. Kraaijenbrink, S. Zuniga, B. Su, H. Shi, C.J. Xiao, W.R. Tang, P. de Knijff, Allele frequency distribution of 21 forensic autosomal STRs in 7 populations from Yunnan, China, *Forensic Sci Int Genet* 3 (2008), <https://doi.org/10.1016/j.fsigen.2008.02.007>.
- [23] X. Yang, X. Zhang, J. Zhu, L. Chen, C. Liu, X. Feng, L. Chen, H. Wang, C. Liu, Genetic analysis of 19 X chromosome STR loci for forensic purposes in four Chinese ethnic groups, *Sci. Rep.* 7 (2017), <https://doi.org/10.1038/srep42782>.
- [24] P. Chen, G. He, X. Zou, M. Wang, F. Jia, H. Bai, J. Li, J. Yu, Y. Han, Forensic characterization and genetic polymorphisms of 19 X-chromosomal STRs in 1344 Han Chinese individuals and comprehensive population relationship analyses among 20 Chinese groups, *PLoS One* 13 (2018), <https://doi.org/10.1371/journal.pone.0204286>.
- [25] L. Chen, Y. Guo, C. Xiao, W. Wu, Q. Lan, Y. Fang, J. Chen, B. Zhu, Genetic polymorphisms and forensic efficiency of 19 X-chromosomal STR loci for Xinjiang Mongolian population, *PeerJ* 2018 (2018), <https://doi.org/10.7717/peerj.5117>.
- [26] G. He, Y. Li, X. Zou, M. Wang, P. Chen, M. Liao, J. Wu, Genetic polymorphisms for 19 X-STR loci of Sichuan Han ethnicity and its comparison with Chinese populations, *Leg. Med.* 29 (2017), <https://doi.org/10.1016/j.legalmed.2017.09.001>.
- [27] X. Zhang, X. Yuan, Y. Huang, J. Yao, J. Zhang, J. Dai, L. Liu, S. Nie, L. Hu, Forensic genetic polymorphisms of 16 X-STR loci in the Yunnan Miao population and their relationship to other Chinese groups, *Leg. Med.* 53 (2021), <https://doi.org/10.1016/j.legalmed.2021.101961>.
- [28] Y. Liu, Y. Ye, X.Q. Yi, J.H. Zhang, G.Y. Fan, D.Y. Hao, Phylogenetic analyses of the 19 STR loci in X chromosome revealed discriminant forensic characteristics for Chinese Daur and Oroqen minorities, *Int. J. Leg. Med.* 136 (2022), <https://doi.org/10.1007/s00414-021-02697-7>.
- [29] C. Deng, F. Song, J. Li, Y. Ye, L. Zhang, W. Liang, H. Luo, Y. Li, Forensic parameters of 19 X-STR polymorphisms in two Chinese populations, *Int. J. Leg. Med.* 131 (2017), <https://doi.org/10.1007/s00414-017-1538-1>.
- [30] C. Xiao, S. Li, X. Zhang, X. Yang, C. Liu, L. Chen, Population genetic analysis of Chinese Zhuang and Mulao minorities using AGCU-X19 STR kit, *Int. J. Leg. Med.* 134 (2020), <https://doi.org/10.1007/s00414-019-02004-5>.
- [31] Y. Han, G. He, S. Gong, J. Chen, Z. Jiang, P. Chen, Genetic diversity and haplotype analysis of Guizhou Miao identified with 19 X-chromosomal short tandem repeats, *Int. J. Leg. Med.* 133 (2019), <https://doi.org/10.1007/s00414-018-1871-z>.
- [32] G. He, Y. Li, X. Zou, Y. Zhang, H. Li, M. Wang, J. Wu, X-chromosomal STR-based genetic structure of sichuan Tibetan minority ethnicity group and its relationships to various groups, *Int. J. Leg. Med.* 132 (2018), <https://doi.org/10.1007/s00414-017-1672-9>.
- [33] G.L. He, Y. Li, X. Zou, P. Li, P.Y. Chen, F. Song, T. Gao, M. Liao, J. Yan, J. Wu, Forensic characteristics and phylogenetic analyses of the Chinese Yi population via 19 X-chromosomal STR loci, *Int. J. Leg. Med.* 131 (2017), <https://doi.org/10.1007/s00414-017-1563-0>.
- [34] Y.S. Liu, H.T. Meng, T. Mei, L.P. Zhang, J.G. Chen, Y.D. Zhang, J. Chen, Y.X. Guo, Q. Dong, J.W. Yan, B.F. Zhu, Genetic diversity and haplotypic structure of Chinese Kazak ethnic group revealed by 19 STRs on the X chromosome, *Gene* 600 (2017), <https://doi.org/10.1016/j.gene.2016.11.018>.
- [35] Daorina Han, Study on Genetic Polymorphism of 19 X-STR in Korean People in Yanbian China, Yanbian University, 2023.
- [36] Yaju Liu, Long Fei, Li Jin, Juntao Yue, Meisen Shi, Genetic polymorphisms of 19 -STR loci for forensic application in China's three ethnicities, *Basic & Clin. Med.* 38 (2018) 913–921.
- [37] Yaju Liu, Juntao Yue, Li Jin, Meisen Shi, Xuebo Li, Genetic polymorphism of 19 X-STR loci for forensic application in Kazakh and Tujia population, *Laboratory Medicine and Clinic* 17 (2020) 1500–1508.
- [38] C. Xiao, X. Yang, H. Liu, C. Liu, Z. Yu, L. Chen, C. Liu, Validation and forensic application of a new 19 X-STR loci multiplex system, *Leg. Med.* 53 (2021), <https://doi.org/10.1016/j.legalmed.2021.101957>.
- [39] A. Gouy, M. Zieger, STRAF—a convenient online tool for STR data evaluation in forensic genetics, *Forensic Sci Int Genet* 30 (2017), <https://doi.org/10.1016/j.fsigen.2017.07.007>.
- [40] J.F. Ferragut, N. Pinto, A. Amorim, A. Picornell, Improving publication quality and the importance of Post Publication Peer Review: the illustrating example of X chromosome analysis and calculation of forensic parameters, *Forensic Sci Int Genet* 38 (2019), <https://doi.org/10.1016/j.fsigen.2018.11.006>.
- [41] L. Excoffier, H.E.L. Lischer, Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows, *Molecular Ecology Resources* 10 (2010) 564–567. *Evol Bioinform Online*.
- [42] P.E. Smouse, S.C. Banks, R. Peakall, Converting quadratic entropy to diversity: both animals and alleles are diverse, but some are more diverse than others, *PLoS One* 12 (2017), <https://doi.org/10.1371/journal.pone.0185499>.
- [43] Ø. Bleka, G. Storvik, P. Gill, EuroForMix: an open source software based on a continuous model to evaluate STR DNA profiles from a mixture of contributors with artefacts, *Forensic Sci Int Genet* 21 (2016), <https://doi.org/10.1016/j.fsigen.2015.11.008>.
- [44] D.J. Balding, R.A. Nichols, DNA profile match probability calculation: how to allow for population stratification, relatedness, database selection and single bands, *Forensic Sci. Int.* 64 (1994), [https://doi.org/10.1016/0379-0738\(94\)90222-4](https://doi.org/10.1016/0379-0738(94)90222-4).

- [45] M.J. Hubisz, D. Falush, M. Stephens, J.K. Pritchard, Inferring weak population structure with the assistance of sample group information, *Mol Ecol Resour* 9 (2009), <https://doi.org/10.1111/j.1755-0998.2009.02591.x>.
- [46] N.M. Kopelman, J. Mayzel, M. Jakobsson, N.A. Rosenberg, I. Mayrose, Clumpak, A program for identifying clustering modes and packaging population structure inferences across K, *Mol Ecol Resour* 15 (2015), <https://doi.org/10.1111/1755-0998.12387>.
- [47] N.A. Rosenberg, J.T.L. Kang, Genetic diversity and societally important disparities, *Genetics* 201 (2015), <https://doi.org/10.1534/genetics.115.176750>.
- [48] K. Haddish, E. Chierito, G. Di Vella, D. Lacerenza, S. Raddi, S. Aneli, A.L. Bogale, E. Kidane, A. Yizengaw, Y. Getaneh, G. Teweledmedhin, H.R.S. Kumar, C. Robino, A reference database of forensic autosomal and gonosomal STR markers in the Tigray population of Ethiopia, *Forensic Sci Int Genet* 56 (2022), <https://doi.org/10.1016/j.fsigen.2021.102618>.
- [49] X. Dai, Q. Zhu, C. Wang, A. Rukeye, Z. Cao, T. Shan, Y. Wang, J. Zhang, FST estimates of 94 populations in China based on STR markers, *Forensic Sci Int Genet* 64 (2023), <https://doi.org/10.1016/j.fsigen.2023.102854>.
- [50] M. Song, Z. Wang, Q. Lyu, J. Ying, Q. Wu, L. Jiang, F. Wang, Y. Zhou, F. Song, H. Luo, Y. Hou, X. Song, B. Ying, Paternal genetic structure of the Qiang ethnic group in China revealed by high-resolution Y-chromosome STRs and SNPs, *Forensic Sci Int Genet* 61 (2022), <https://doi.org/10.1016/j.fsigen.2022.102774>.
- [51] L. Tan, S. Li, X. Zhang, Risk attitude in the DuLong minority ethnicity of China, *Front. Psychol.* 12 (2021), <https://doi.org/10.3389/fpsyg.2021.596745>.
- [52] J. Jia, X. Liu, Q. Fan, C. Fang, M. Wang, J. Zhang, W. Li, L. Shi, X. Zhang, C. Chen, Z. Yu, C. Li, K. Yun, J. Yan, Development and validation of a multiplex 19 X-chromosomal short tandem repeats typing system for forensic purposes, *Sci. Rep.* 11 (2021), <https://doi.org/10.1038/s41598-020-80414-x>.
- [53] J. Edelman, S. Hering, M. Michael, R. Lessig, D. Deischel, G. Meier-Sundhausen, L. Roewer, I. Plate, R. Szibor, 16 X-chromosome STR loci frequency data from a German population, *Forensic Sci. Int.* 124 (2001), [https://doi.org/10.1016/S0379-0738\(01\)00565-5](https://doi.org/10.1016/S0379-0738(01)00565-5).
- [54] S. Turrina, D. De Leo, Population genetic comparisons of three X-chromosomal STRs (DXS7132, DXS7133 and GATA172D05) in North and South Italy, *Int Congr Ser* 1261 (2004), [https://doi.org/10.1016/S0531-5131\(03\)01626-1](https://doi.org/10.1016/S0531-5131(03)01626-1).
- [55] M. Jia, Q. Li, T. Zhang, B. Dong, X. Liang, S. Fu, J. Yu, Genetic diversity analysis of the Chinese Daur ethnic group in Heilongjiang province by complete mitochondrial genome sequencing, *Front. Genet.* 13 (2022), <https://doi.org/10.3389/fgene.2022.919063>.
- [56] A. Rubab, M. Shafique, F. Javed, S. Saleem, F.T. Zahra, D. McNevin, A.A. Shahid, Population genetic portrait of Pakistani Lahore-Christians based on 32 STR loci, *Sci. Rep.* 10 (2020), <https://doi.org/10.1038/s41598-020-76016-2>.
- [57] T.I. Huszar, W.F. Bodmer, K. Hutnik, J.H. Wetton, M.A. Jobling, Sequencing of autosomal, mitochondrial and Y-chromosomal forensic markers in the People of the British Isles cohort detects population structure dominated by patrilineages, *Forensic Sci Int Genet* 59 (2022), <https://doi.org/10.1016/j.fsigen.2022.102725>.
- [58] J. Li, W. Zeng, Y. Zhang, A.M.S. Ko, C. Li, H. Zhu, Q. Fu, H. Zhou, Ancient DNA reveals genetic connections between early Di-Qiang and Han Chinese, *BMC Evol. Biol.* 17 (2017), <https://doi.org/10.1186/s12862-017-1082-0>.
- [59] S.Z. Gao, Y. Zhang, D. Wei, H.J. Li, Y. Bin Zhao, Y.Q. Cui, H. Zhou, Ancient DNA reveals a migration of the ancient Di-qiang populations into Xinjiang as early as the early Bronze Age, *Am. J. Phys. Anthropol.* 157 (2015), <https://doi.org/10.1002/ajpa.22690>.
- [60] C.C. Wang, L.X. Wang, R. Shrestha, M. Zhang, X.Y. Huang, K. Hu, L. Jin, H. Li, Genetic structure of Qiangic populations residing in the Western Sichuan corridor, *PLoS One* 9 (2014), <https://doi.org/10.1371/journal.pone.0103772>.