Heliyon 9 (2023) e23005

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

A Tibetan group from Ngawa Tibetan and Qiang Autonomous Prefecture, southwest China, is rich in genetic polymorphisms at 36 autosomal STR loci and shares a complex genetic structure with other Chinese populations

Binghui Song [a,b,1], Jiewen Fu [a,b,c,1], Kan Guo [a,1], Jie Qian [a], Ting He [a], Lisha Yang [a,d], Jingliang Cheng [a,b,c,**], Junjiang Fu [a,b,c,*]

[a] Key Laboratory of Epigenetics and Oncology, the Research Center for Preclinical Medicine, Southwest Medical University, Luzhou, 646000, Sichuan, China
[b] Laboratory of Precision Medicine and DNA Forensic Medicine, the Research Center for Preclinical Medicine, Southwest Medical University, Luzhou, 646000, Sichuan, China
[c] Laboratory of Forensic DNA, the Judicial Authentication Center, Southwest Medical University, Luzhou, 646000, Sichuan, China
[d] Department of Obstetrics and Center for Prenatal Diagnosis, the Affiliated Hospital of Southwest Medical University, Luzhou, China

ARTICLE INFO

Keywords:
Short tandem repeats (STR)
Genetic polymorphism
Forensic genetics
Ngawa Tibetan population
Population data

ABSTRACT

The Tibetan people are ancient and populous, constituting the seventh-largest of the fifty-five ethnic minority groups in China. The Ngawa Tibetan and Qiang Autonomous Prefecture (NTQAP), situated on the border of northwest and southwest China, has its distinct group relationships. Short tandem repeat (STR) is extremely polymorphic and extensively used in the application of forensic medicine and population genetics. However, it is not clear the genetic information including linkage disequilibrium (LD) by 36 autosomal STR (A-STR) markers in the Tibetan group from NTQAP. The Tibetan population from NTQAP of southwest China was examined for 36 A-STR loci in the research. Every marker across the 36 A-STR loci was consistent with Hardy–Weinberg equilibrium (HWE). The results of the calculation revealed that the total discrimination power (TDP) is 1–2.2552 $\times$ 10^{-42} and the cumulative probability of exclusion (CPE) is 1–1.3031 $\times$ 10^{-16}. Subsequently, a total of 345 alleles with allelic frequencies ranging from 0.00382 to 0.55343 were identified, and the allelic numbers varied from 5 in both the TH01 and TPOX markers to 28 in the SE33 locus. The Ngawa Tibetan population, along with other Chinese populations, exhibited influences from historical factors and regional distribution, as indicated by the results of population genetics analysis. We thus first explored the genetic characteristics and correlated forensic parameters of the 36 A-STR markers in NTQAP to fill the gap in the Tibetan population. It was discovered that these 36 autosomal STR markers supplemented forensic STR databases and offered extremely valuable polymorphisms for Chinese forensic applications, such as parentage testing and personal identification. Moreover, the study

* Corresponding author. Key Laboratory of Epigenetics and Oncology, the Research Center for Preclinical Medicine, Southwest Medical University, Luzhou, 646000, Sichuan, China
** Corresponding author. Key Laboratory of Epigenetics and Oncology, the Research Center for Preclinical Medicine, Southwest Medical University, Luzhou 646000, Sichuan, China.
E-mail addresses: jingliangc@swmu.edu.cn (J. Cheng), fujunjiang@hotmail.com, fujunjiang@swmu.edu.cn (J. Fu).
[1] Equal contributions.

https://doi.org/10.1016/j.heliyon.2023.e23005
Received 8 July 2023; Received in revised form 14 November 2023; Accepted 23 November 2023
Available online 29 November 2023

would contribute additional information regarding the substructure and diversity in the Chinese population.

1. Introduction

Short tandem repeat (STR) as a DNA genetic marker is highly polymorphic and accurate and is now the most widely used genetic marker in the identification of forensic medicine, and STR typing has become a routine technique in the identification of forensic genetics [1]. Forensic genetics and population genetics based on STR data have also been extensively studied, and they provide a powerful tool for researching ethnic origin, formation and development, population migration, population structure, and genetic differentiation. Additionally, they offer data support for the identification of forensic genetics and other judicial practice, ensuring the scientificity and accuracy of identification results [2]. The commonly used STR genetic markers are autosomal STR (A-STR), Y chromosome STR (Y-STR), and X chromosome STR (X-STR). The A-STR system is used for forensic DNA analysis such as personal identification and parentage testing, providing important evidence for solving missing persons cases, identifying victims, and resolving paternity cases with mutations [3]. Before using any human polymorphic genetic marker in the application of forensics, studies on the frequencies of alleles and genotypes are required. As a result, forensic geneticists must work to look into and present genetic marker data in various populations. Furthermore, the frequency of allele distribution, the number of alleles, and other parameters in forensics vary across populations for A-STR loci, so population diversities and genetic markers have become hot topics in forensic study.

Tibetans are an ancient and unique ethnic group, most of whom live on the Qinghai-Tibet Plateau all year round. The historical, cultural, and biological studies related to the Tibetan have attracted the attention of many scholars. Lu et al. [4] studied the ancestral composition of Tibetans and revealed their origin, showing that the origins of the Tibetans are more complex and ancient than expected. Population migration and integration have always been a hot topic of research. Qi et al. [5] showed that there are two major prehistoric migrations of modern humans and that Tibetans had already adapted to the high-altitude environment. He et al. [6] used genome-wide data to reveal the expansion and integration of Tibetans with other populations in the Tibetan Plateau region. Yang et al. [7] explored the genetic impact of Tibetan expansion on the Balti people. In addition, linguistics is closely related to genetics, which has a large amount of cultural information and can be a good reflection of the historical activities of the nation. The Sino-Tibetan language family is the second largest in the world, and the study of its origins and differences can provide important clues to the culture and history of the Tibetans. The study of the Sino-Tibetan language family also provides a basis for human activities such as migration and cultural exchange of the Tibetans [8,9].

For forensic genetics, the study of STR on the Tibetan has been constantly enriched. Studies on A-STR allele frequencies in the Tibetan from Qinghai Province, Tibet Lhasa, and Nepal were published as early as 2007 [10–12]. In addition, Li et al. [13] also conducted a study on the Y-STR of the Tibetans in Tibet Lhasa. More STR studies are being carried out to provide richer population genetics data for Tibetans across different geographical areas. In 2016, Gou et al. [14] reported the allele frequency distribution and forensic parameters in the Tibetan population using 21 non-combined DNA index system STR markers. Recently, research on the sex chromosome STR of the Tibetan has been continuously deepening. He et al. [15] explored the population genetic data of 19 X-STR markers in the Tibetan group in Sichuan Province and provided key forensic parameters. Cao et al. [16] conducted an analysis in 27 Y-STR markers in Qinghai Tibetan males. Song et al. [17] conducted a genetic diversity and phylogenetic analysis on 29 Y-STR markers in the Sichuan Tibetan group. Li et al. [18] conducted Y-STR studies on the Tibetan population from 11 different regions of the Tibetan Plateau and provided haplotype data of 41 Y-STR markers. As an important corridor for ethnic migration, the Tibetan-Yi corridor has a complex population genetic structure [19]. He et al. [20] explored the forensic characteristics and genetic structure of the Ü-Tsang and Kham Tibeto-Burman-speaking Tibetans based on the data of 27 Y-STR markers. In addition, Fan et al. [21] examined haplotype data of 41 Y-STR markers from three major populations in the Tibetan-Yi corridor, including the Han, Tibetan, and Yi populations.

As one part of the Tibetan-Yi corridor, the Ngawa Tibetan and Qiang Autonomous Prefecture (NTQAP) is located on the southeastern margin of the Tibetan Plateau and in the northwestern part of Sichuan Province. It is adjacent to Qinghai Province and Gansu Province, with geographical coordinates between north latitude 30°5′-34°9′ and east longitude 100°0′-104°7’. It is situated on the transitional zone of the eastern platform area and the western geosynclinal area, and the overall contour of the land surface is a typical plateau, with high terrain and an average elevation between 3500 and 4000 m. The elevation of the mountains decreases gradually from south to north, and the elevation of the valley decreases gradually from northwest to southeast. Temperature decreases accordingly from southeast to northwest and with altitude from low to high. Tibetans are the main population of the NTQAP, with a current population of approximately 537,100, in addition to the Qiang, Han, and Hui ethnic groups. The Tibetan population of NTQAP is second only to that of Ganzi Tibetan Autonomous Prefecture in China, outside the Tibet Autonomous Region. The region is rich in ethnic resources, which together constitute the distinctive population structure and genetic characteristics of the area. So, it is essential to investigate the genetic polymorphism and genetic structure for A-STR loci of the Tibetan population in the region. However, it is not clear the genetic information including allele frequencies, polymorphism information, and linkage disequilibrium (LD) by 36 A-STR loci in the Tibetan population from NTQAP. Therefore, this study can provide basic data for forensic genetics and enable objective judgements on genetic markers between closely related individuals and groups, ensuring the scientific and precise outcomes of forensic identification.

2. Materials and methods

2.1. Regents

We used the Microreader 36 A Direct ID System (V1.1) kit (Microreader 36 A kit) to test the Tibetan group in the NTQAP, which included thirty-six A-STR loci, one Y-indel, one Y-STR, and one sex-determining locus, namely TH01, TPOX, vWA, CSF1PO, Penta D, Penta E, FGA, SE33, D1S1656, D3S3045, D5S818, D18S51, D6S1043, D15S659, D6S477, D3S1358, D13S317, D7S820, D16S539, D8S1132, D7S3048, D2S441, D4S2366, D8S1179, D14S608, D19S433, D22S1045, D2S1338, D5S2500, D10S1248, D10S1435, D12S391, D18S535, D19S253, D21S11, D11S2368, DYS391, and amelogenin (Microread Genetics Co., Ltd., Beijing, Lot: 10412432). We used Chelex-100 for DNA extraction (Bio-Rad Laboratories Co., Ltd., America, Lot: 64175704).

2.2. Sampling and genomic DNA extraction

We sampled blood from 131 healthy and unrelated citizens (63 males and 68 females) of the Ngawa Tibetan in the Ruoergai county of NTQAP while observing the principle of informed consent. The FTA card and Chelex-100 were used to extract genomic DNA [22]. A blood spot sample measuring 1 mm in diameter was taken from each using a puncher. It was subsequently put in a 600 μl centrifuge tube. 400 μl of double distilled water (ddH_2O) was then poured into the centrifuge tube shaken for 5 s, and waited for 30 min at room temperature. A high-speed centrifuge (Thermo Fisher, USA) was applied to centrifuge the tubes for 3 min at 12,000 rpm. After removing the supernatant, 35 μl of 5 % Chelex-100 was supplemented, and the tube was kept in a water bath at 56 °C for 30 min. Then it was shook for 5 s and placed in a bath of boiling water for 8 min, and finally, centrifuged at top speed for 3 min. The 20 μl of supernatant was transferred into a fresh 600 μl tube and then amplified using PCR. The sample was stored in the refrigerator at −20 °C for use.

2.3. PCR amplification and autosomal STR genotype

Following the manufacturer's guidelines, a Microreader 36 A kit system was used to amplify the 36 autosomal STR loci. An Applied Biosystems Veriti® 96-Well Thermal Cycler (Applied Biosystems, Life Technology, USA) was used to create the PCR amplification. For details, a 10 μL PCR reaction volume including Microreader 2.5 × Master Mix III 4 μl, Microreader 36 A D 5 × Primer Mix 2 μl, ddH_2O 1 μl, and the DNA sample 3 μl. The program settings for PCR are as follows: pre-denaturation at 95 °C for 5 min; denaturation at 94 °C for 20 s, annealing at 59 °C for 90s, extension at 59 °C for 90 s, cycle number 28, then extension at 60 °C for 60 min, maintenance at 4 °C. The electrophoresis reagent was prepared by taking 1 μl of PCR product or Microreader 36 A Direct ID System Allelic ladder from Ngawa Tibetan samples mixed with 0.5 μl of molecular weight internal standard QD650 and 8.5 μl of formamide deionized, denatured at 95 °C for 3 min and then cooled in a −20 °C refrigerator for 3 min. Capillary electrophoresis was performed in a 3500Dx Genetic Analyzer (Applied Biosystems, Life Technology, USA) and STR type was analyzed using GeneMapper ID-X software (Thermo Fisher, USA) [23]. A control sample 9948 was included for the STR typing quality control.

The accreditation by the China National Accreditation Service for Conformity Assessment (CNAS) indicates that our laboratory has the competence to supply testing and calibration services in compliance with relevant recognized standards. In addition, the laboratory has passed the accreditation of the Accreditation Criteria for the Competence of Testing and Calibration Laboratories (ISO/IEC 17025:2017). The Specification of Parentage Testing by China (GB/T 37223-2018) provides a standard for Autosome STR genotype analysis. Every method in this research was carried out in compliance with applicable guidelines and regulations, such as the STR population study guidelines for International Society for Forensic Genetics (ISFG) [24].

2.4. Data analysis

After obtaining the data of STR typing for the Ngawa Tibetan, the Modifed-PowerStats software was used to obtain allele frequencies for 36 A-STR loci, Hardy–Weinberg equilibrium (HWE) and its *p*-value, and forensic parameters such as probability of exclusion (PE), typical paternity index (TPI), polymorphism information content (PIC), matching probability (MP), discrimination power (DP). The Arlequin v3.5 software was applied to carry out the test of LD for 36 A-STR loci, as well as to calculate genetic differentiation between the previously published Chinese population and the Ngawa Tibetan population [25].

We also collected the allele frequencies of autosomal STR from 26 relevant Chinese populations in previously published studies on PubMed, including 14 Han populations [26–39], five Tibetan [40–43], two Yi [44,45], one Li [46], one She [47], one Uyghur [48], one Xibe [49], and one Hui [43]. The approximate geographic locations of the Ngawa Tibetan population and the other 26 populations were depicted on a map created using the R project software on the website https://www.r-project.org/(version 4.0.5). The modified PHYLIP program (version 3.698) constructed a Nei's standard genetic distance between the Ngawa Tibetan group and other 26 related Chinese groups [50]. Furthermore, OriginPro software (version 9.9.5.171) generated a heatmap showing the standard genetic distance between these groups [51]. The Multivariate Statistical Package (MVSP) software (version 3.22) evaluated principal component analysis (PCA) with normalized data on allele frequencies at 15 A-STR markers from 27 populations. According to the standard genetic distance, IBM SPSS 25 and MEGA-X software (version 10.0.5) were utilized to generate the multidimensional scaling plot (MDS) and neighbor-joining (NJ) phylogenetic tree [52], respectively. We conducted the population structure analysis of the Ngawa Tibetan group and other published Chinese groups by the STRUCTURE v2.3.4 software based on 15 A-STR loci [53]. For the parameters, we set the length of the burning period as 10,000 and the steps of the Markov Chain Monte Carlo (MCMC) as 50,000 and kept all others at

Table 1
The corresponding forensic statistical parameters for 36 A-STR loci of the Tibetan population from Ngawa Tibetan and Qiang Autonomous Prefecture, Southwest China.

Loci	Allele number	Genotype number	H_{exp}	H_{obs}	MP	DP	PIC	PE	TPI	H	h	p
TH01	5	11	0.63711	0.64885	0.18863	0.81137	0.58479	0.35365	1.42391	0.35115	0.64885	0.73599
D5S818	6	14	0.71013	0.72519	0.13653	0.86347	0.65701	0.46830	1.81944	0.27481	0.72519	0.65461
D21S11	14	41	0.84677	0.87023	0.05402	0.94598	0.82239	0.73510	3.85294	0.12977	0.87023	0.40021
D18S51	14	35	0.82945	0.84733	0.05961	0.94039	0.80232	0.68961	3.27500	0.15267	0.84733	0.52468
D6S1043	13	45	0.85608	0.81679	0.04015	0.95985	0.83456	0.63057	2.72917	0.18321	0.81679	0.24481
D15S659	9	30	0.84090	0.81679	0.05134	0.94866	0.81293	0.63057	2.72917	0.18321	0.81679	0.51687
D6S477	10	27	0.82537	0.83206	0.06171	0.93829	0.79453	0.65983	2.97727	0.16794	0.83206	0.76795
D3S1358	7	14	0.73262	0.75573	0.12744	0.87256	0.67820	0.51961	2.04688	0.24427	0.75573	0.50419
D13S317	7	24	0.82528	0.83206	0.06311	0.93689	0.79386	0.65983	2.97727	0.16794	0.83206	0.76599
D7S820	7	19	0.78661	0.82443	0.08700	0.91300	0.74682	0.64513	2.84783	0.17557	0.82443	0.25645
D16S539	6	17	0.77910	0.79389	0.09143	0.90857	0.73722	0.58776	2.42593	0.20611	0.79389	0.62537
CSF1PO	8	18	0.75214	0.75573	0.11509	0.88491	0.70395	0.51961	2.04688	0.24427	0.75573	0.86447
Penta D	10	25	0.81183	0.85496	0.06859	0.93141	0.77929	0.70467	3.44737	0.14504	0.85496	0.17842
D8S1132	9	32	0.85890	0.83206	0.04423	0.95577	0.83425	0.65983	2.97727	0.16794	0.83206	0.44328
D7S3048	10	38	0.87314	0.87786	0.03560	0.96440	0.85197	0.75046	4.09375	0.12214	0.87786	0.78367
D2S441	8	22	0.79974	0.84733	0.07838	0.92162	0.76568	0.68961	3.27500	0.15267	0.84733	0.14971
vWA	7	20	0.79665	0.83206	0.08292	0.91708	0.75962	0.65983	2.97727	0.16794	0.83206	0.27674
D8S1179	9	28	0.82910	0.80153	0.05821	0.94179	0.79988	0.60188	2.51923	0.19847	0.80153	0.46156
TPOX	5	12	0.61371	0.54962	0.21322	0.78678	0.55398	0.23472	1.11017	0.45038	0.54962	0.14713
Penta E	19	70	0.92331	0.93130	0.02080	0.97920	0.91016	0.85969	7.27778	0.06870	0.93130	0.62730
D14S608	8	27	0.83511	0.83206	0.05868	0.94132	0.80644	0.65983	2.97727	0.16794	0.83206	0.99643
D4S2366	7	20	0.73071	0.70992	0.10949	0.89051	0.68533	0.44378	1.72368	0.29008	0.70992	0.64365
D3S3045	7	23	0.75760	0.70992	0.10378	0.89622	0.71371	0.44378	1.72368	0.29008	0.70992	0.23360
D19S433	11	34	0.83820	0.87023	0.05681	0.94319	0.81213	0.73510	3.85294	0.12977	0.87023	0.27717
D22S1045	7	17	0.76456	0.80153	0.10856	0.89144	0.71600	0.60188	2.51923	0.19847	0.80153	0.28389
D2S1338	10	36	0.83967	0.83969	0.05029	0.94971	0.81472	0.67466	3.11905	0.16031	0.83969	0.91998
FGA	15	37	0.84977	0.83206	0.04621	0.95379	0.82452	0.65983	2.97727	0.16794	0.83206	0.64640
D5S2500	9	29	0.82176	0.76336	0.05612	0.94388	0.79264	0.53290	2.11290	0.23664	0.76336	0.10080
D10S1435	6	14	0.76177	0.80153	0.11637	0.88363	0.71164	0.60188	2.51923	0.19847	0.80153	0.25357
D18S535	7	22	0.78188	0.77863	0.08840	0.91160	0.74337	0.55999	2.25862	0.22137	0.77863	0.99421
D1S1656	12	35	0.82431	0.80916	0.05705	0.94295	0.79526	0.61615	2.62000	0.19084	0.80916	0.72024
D12S391	11	34	0.84243	0.88550	0.05891	0.94109	0.81560	0.76590	4.36667	0.11450	0.88550	0.14915
D10S1248	7	19	0.77314	0.74046	0.09376	0.90624	0.73044	0.49358	1.92647	0.25954	0.74046	0.41886
SE33	28	88	0.94612	0.93893	0.01521	0.98479	0.93555	0.87542	8.18750	0.06107	0.93893	0.86074
D19S253	8	23	0.77076	0.74046	0.09877	0.90123	0.73093	0.49358	1.92647	0.25954	0.74046	0.45852
D11S2368	9	27	0.83740	0.87023	0.06136	0.93864	0.80855	0.73510	3.85294	0.12977	0.87023	0.26746

Hexp: expected heterozygosity; Hobs: observed heterozygosity; MP: matching probability; DP: discrimination power; PIC: polymorphism information content; PE: probability of exclusion; TPI: typical paternity index; H: Homozygotes; h: Heterozygotes; p: p value of the exact test in Hardy-Weninbegy equilibrium.

their default settings. We performed ten independent runs for each K and K ranged from 2 to 8. We computed the optimal K value by the formula $m(|L(K+1) - 2L(K) + L(K-1)|)/s[L(K)]$ (in the formula m represents the mean value, s represent standard deviation) for ΔK [54]. CLUMPP 1.1.2 software was used to merge the results of multiple replications (Q-matrix) of the optimal K value obtained by running STRUCTURE [55].

3. Results

3.1. Linkage disequilibrium (LD) and forensic data for the Ngawa Tibetan group

A-STR genotypes obtained from the samples were used to test for LD in the Ngawa Tibetan population. Following Bonferroni correction ($p = 0.05/630 \approx 0.00008$), we did not observe significant LD changes between these 36 A-STR markers except for the pair of D13S317 and D18S535 loci (Table S1), suggesting that almost all markers were statistically independent. The allele frequencies of 36 A-STR markers were calculated by STR genotyping of the Ngawa Tibetan population and the results are displayed in Table S2. The *p*-values for HWE and other forensic parameters of 36 A-STR in the Ngawa Tibetan population were reported in Table 1, which showed that all loci are HWE following Bonferroni correction ($p = 0.05/36 \approx 0.00139$). Based on the genetic data of the 36 A-STR loci in the Ngawa Tibetan population, the total discrimination power (TDP) was calculated as $1–2.2552 \times 10^{-42}$ and the cumulative probability of exclusion (CPE) was calculated as $1–1.3031 \times 10^{-16}$, respectively.

A total of 345 alleles were identified in 36 A-STR loci, with the highest allele frequency of 0.00382 and the lowest allele frequency of 0.55343 (Table 1). We could also obtain allele numbers ranging from 5 at the TH01 locus and TPOX locus to 28 at the SE33 locus. Many genotypes were observed in different loci, among which the largest and smallest number of genotypes was 88 in locus SE33 and 11 in locus TH01, respectively. In addition, the highest PIC was 0.93555 in locus SE33 and the lowest PIC was 0.55398 in locus TPOX between these A-STR markers. We also computed the averages of several relevant parameters including DP of 0.92062, PE of 0.61538, expected heterozygosity (Hexp) of 0.80286, and observed heterozygosity (Hobs) of 0.80471 in 36 A-STRs. Thirty-four of 36 A-STR markers had values of Hobs and DP greater than 0.8, apart from locus TPOX and TH01. Additionally, locus SE33 exhibited the highest levels of Hobs and DP, which were found to be 0.93893 and 0.98479, respectively.

3.2. Population structures and comparisons among 27 Chinese-related populations

To compare the Ngawa Tibetan population data with other Chinese populations, allele frequency data for A-STR were collected for 26 Chinese-related populations, including those in close geographical proximity. More details of the 27 groups are shown in Table 2. According to the allelic frequencies of A-STR makers, including CSF1PO, vWA, D19S433, D21S11, D18S51, D5S818, D6S1043, D3S818, D12S391, D13S317, D7S820, D8S1179, D16S539, D2S1338 and FGA, we computed the genetic distance matrix to accomplish population comparisons (Table S3). The approximate geographic locations of 27 Chinese populations are depicted in Fig. 1, as well as

Table 2
Information of Ngawa Tibetan and other 26 relative Chinese reference populations.

Group	Region	Sample size	Number of A-STR	Reference
Han	Tianjin	565	20	Chen et al. [26]
	Hebei	1008	19	Sun et al. [27]
	Shandong	511	23	wang et al. [28]
	Shanxi-Yuncheng	525	23	Gao et al. [29]
	Shanghai	676	20	Xie et al. [30]
	Zhejiang	3510	19	wang et al. [31]
	Fujian-Xiamen	5141	20	Lu et al. [32]
	Hubei	3078	21	Xiao et al. [33]
	Hunan	560	21	Zou et al. [34]
	Chongqing	671	19	Zou et al. [35]
	Sichuan	2793	21	He et al. [36]
	Guangdong	1533	23	Yang et al. [37]
	Yunnan	2068	20	Zhang et al. [38]
	Qinghai	2000	20	Wang et al. [39]
Tibetan	Ngawa	131	36	–
	Sichuan-Chengdu	200	23	Liu et al. [43]
	Tibet-Chamdo	2249	18	Li et al. [40]
	Gansu-Tianzhu	168	15	Yao et al. [41]
	Gansu-Gannan	635	15	Yao et al. [41]
	Tibet-Lhasa	1220	15	He et al. [42]
Yi	Sichuan	1016	19	Cheng et al. [44]
	Yunnan	559	20	Zhang et al. [45]
Li	Hainan	653	19	Fan et al. [46]
She	Fujian	154	38	Yuan et al. [47]
Uyghur	Xinjiang	290	38	Yuan et al. [48]
Xibe	Xinjiang	222	19	Meng et al. [49]
Hui	Ningxia	183	23	Liu et al. [43]

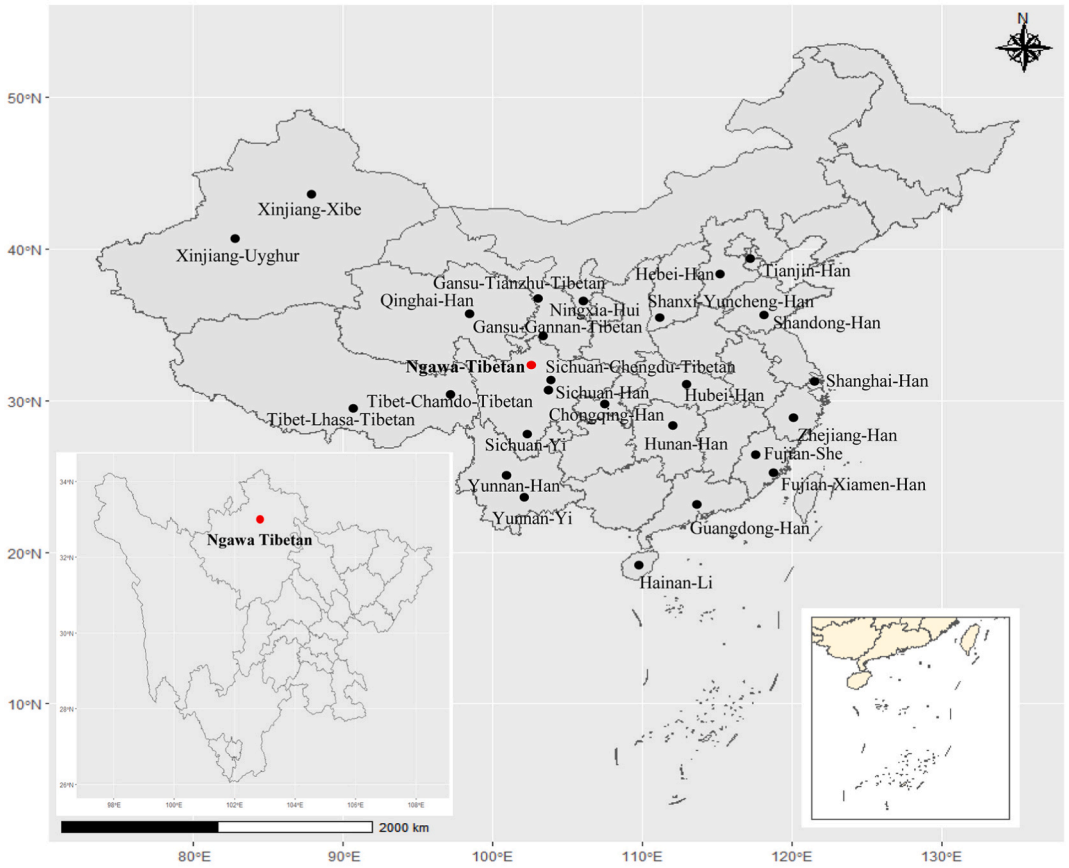

Fig. 1. The distribution of 27 populations in China. Map showing the approximate geographic locations of the 27 populations in China and the Ngawa Tibetan population in Sichuan Province.

the geographic locations of the Ngawa Tibetan population in Sichuan Province. To provide a more visual representation of the genetic distances between the 27 populations, a heatmap was constructed based on the genetic distance matrix (Fig. 2). From Fig. 2 and Table S3, we found that the Tibet Chamdo Tibetan is the closest group to the Ngawa Tibetan with a genetic distance of 0.016168, followed by the Gansu Gannan Tibetan (0.016305), the Tibet Lhasa Tibetan (0.018191), the Sichuan Chengdu Tibetan (0.026211), and the Gansu Tianzhu Tibetan (0.027047). The nearest other minority to the Ngawa Tibetan is the Sichuan Yi (0.032293) and the nearest Han Chinese is the Shandong Han (0.033249). The Fujian She (0.090624) had the greatest genetic distance, followed by the Yunnan Yi (0.059303) and Hainan Li (0.065222).

According to the MDS graph calculated with the model for Euclidean distance (Fig. 3.), we found that all Tibetan nationality populations are clustered in the upper right of the graph, while numerous minorities are dispersed around the graph. Moreover, the various Han populations were closer together, mostly grouped in the middle of this graph. As ethnic minorities, the Xinjiang Xibe and the Sichuan Yi were closely located with Han-Chinese groups. In the top left corner, the Fujian She was far apart from the other populations. According to Nei's standard distance matrix, we constructed the NJ phylogenetic tree to reveal the evolutionary history and relation of these Chinese populations (Fig. 4.). Results showed evident differences between the Tibetan and Han Chinese populations. Some of the ethnic minorities among these Chinese groups created distinct clusters that set them apart from other groups, such as the Fujian She, Hainan Li, and Yunnan Yi. The Ngawa Tibetan was relatively close to Gansu Gannan Tibetan and formed a cluster. Many Tibetan groups formed the cluster with each other from adjacent regions, especially the Sichuan Chengdu Tibetan, the Tibet Chamdo Tibetan, and the Tibet Lhasa Tibetan. Furthermore, the Fujian She developed a distinct cluster that clearly indicates a distant relationship from the other groups, whereas both the Yunnan Yi and Hainan Li had a close relationship with groups in southern China.

Additionally, we carried out PCA using the normalized allelic frequencies for 15 A-STR loci in 27 groups (Fig. 5). In this graph, the first and second principal components illustrated 30.365 % and 18.172 % of the total variance, respectively. Between these 27 Chinese populations, we observed a distinct population stratification that was broadly split into three segments. The majority of China's ethnic minorities were scattered on the left side of the graph, while the Yunnan Yi and the Hainan Li were distributed in the upper right of the graph. Besides, the Hainan Li and the Yunnan Yi showed a relatively close genetic distance from the Han Chinese group in southwest China. Although all Tibetan groups were gathered on the left of the graph, there are still considerable distances between groups. It is

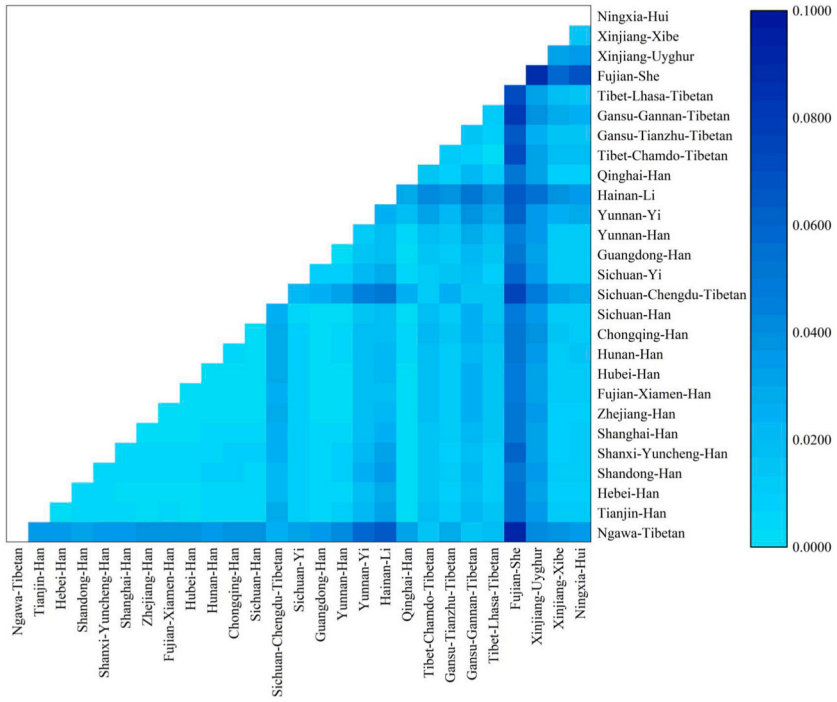

Fig. 2. Heatmap of the standard genetic distance for 27 Chinese populations.

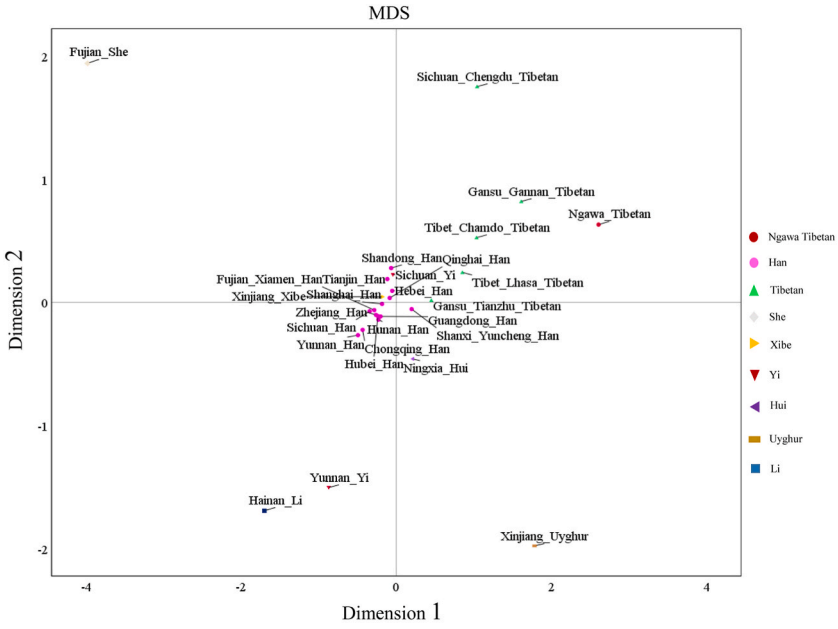

Fig. 3. Multidimensional scaling (MDS) plot of the genetic distance of 27 populations in China. The MDS plot displays the genetic relationships between the Ngawa Tibetan group and the other 26 relevant groups based on the genetic distance matrix.

interesting to note that Fujian She showed an obvious distant relationship from other populations.

We did population comparisons between the Ngawa Tibetan group and the other 26 related Chinese groups at 15 A-STR makers by computing the Fst values and their corresponding p values, as shown in Table S4. The calculations showed that the Hainan Li is the population that differs most from the Ngawa Tibetan, with significant differences at three loci. The Yunnan Yi and the Fujian She had notable genetic differences with the Ngawa Tibetan at two A-STR markers. The Sichuan Chengdu Tibetan and Yunnan Han had notable differences with the Ngawa Tibetan at one marker after Bonferroni correction ($p = 0.05/390 \approx 0.00013$). Nevertheless, none of these

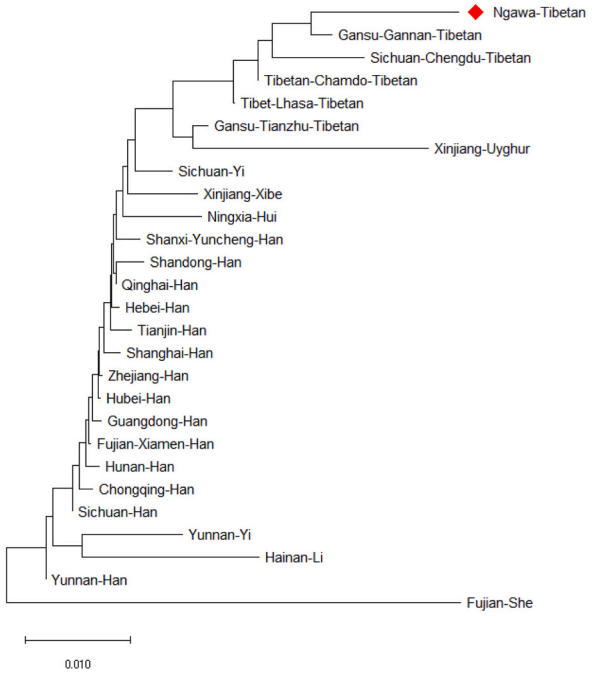

Fig. 4. The evolutionary tree of 27 populations in China. The phylogenetic tree displays the genetic relationships between the Ngawa Tibetan group and the other 26 relevant groups.

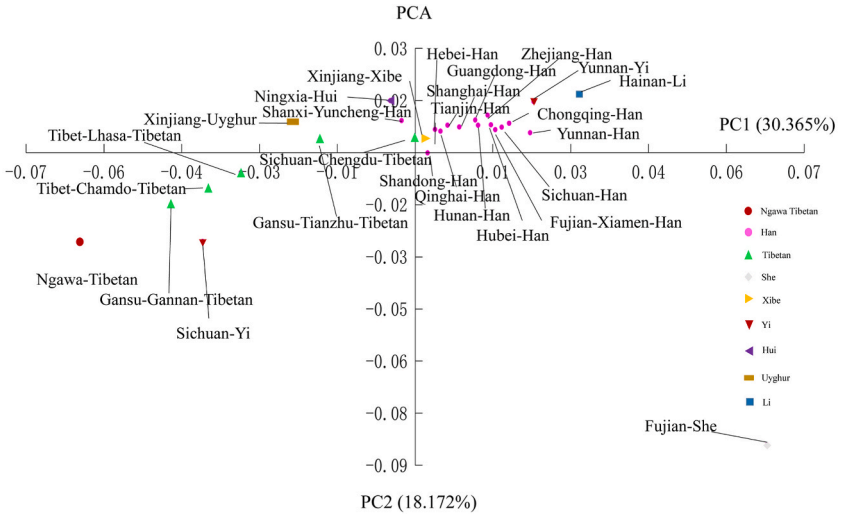

Fig. 5. Principal component analysis (PCA) of 27 populations in China. PCA estimates the genetic distance based on the allele frequencies of 15 A-STR markers in the 27 Chinese populations.

A-STR markers showed any statistically significant differentiation between the Ngawa Tibetan and other groups.

We performed STRUCTURE analysis based on raw data typing of 15 A-STR loci, including CSF1PO, vWA, D19S433, D21S11, D18S51, D5S818, D6S1043, D3S818, D12S391, D13S317, D7S820, D8S1179, D16S539, D2S1338 and FGA. In this analysis, we selected eight of 26 Chinese-related populations that are close to the Ngawa Tibetan population, all of which come from the Tibetan-Yi corridor or southwest China. We obtained that ΔK has a clear peak at K = 2 by using the computing formula, showing the optimum cluster number K = 2 (Fig. 6A). In the STRUCTURE plot (Fig. 6B), each vertical line represents individual data from different populations, and its different colored lengths indicate the distribution of population components among the deductive K groups (K = 2). In the plot, we found similar population component distributions at K = 2, indicating that these populations have no clear population structure and may have similar ancestry. These results may be influenced by geographic proximity and perhaps improved by increasing the number of loci detected.

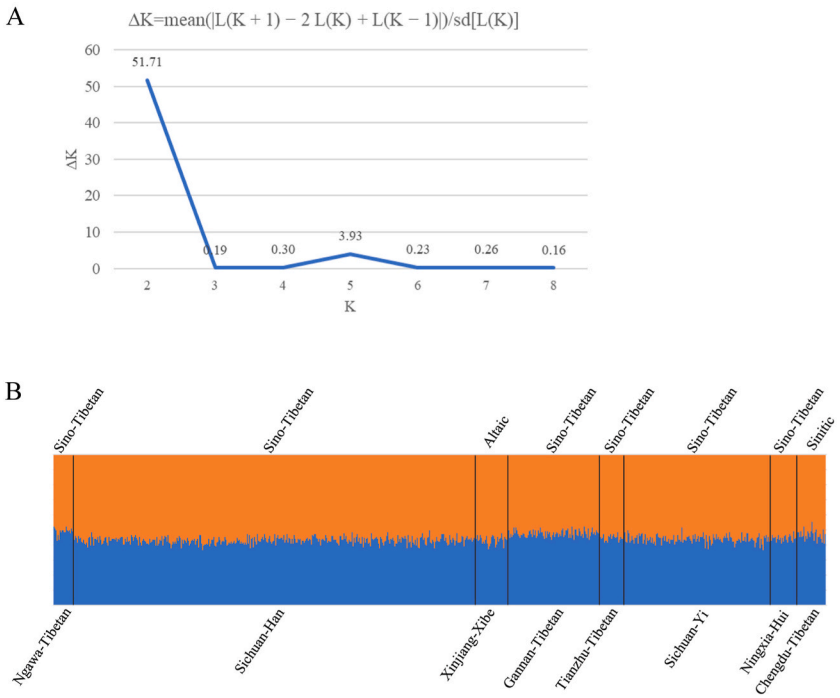

Fig. 6. The STRUCTURE plot is based on 15 A-STR markers with the optimum cluster number K = 2. A. The ΔK values correspond to different K values. B. The cluster plot of STRUCTURE analysis in the optimal K value (K = 2). Population names are listed beneath the plot and population languages are above the plot.

4. Discussion

The Tibetan people are one of the 56 ethnic populations in China and are indigenous to the Qinghai-Tibet Plateau. According to the results of the seventh population census in 2020, the Tibetan population is 7,060,731, making it the eighth-largest ethnic minority in China (http://www.stats.gov.cn/sj/pcsj/rkpc/7rp/indexch.htm). Tibetans are mostly distributed in the Tibet Autonomous Region, Diqing Tibetan Autonomous Prefecture of Yunnan Province, western Sichuan Province, Gannan Tibetan Autonomous Prefecture of Gansu Province, Qinghai Province, and other regions in China. In the Tibetan-Yi corridor, the Tibetan and Han, Yi and Qiang groups have formed a complex population structure. As one section of the Tibetan-Yi corridor, the NTQAP is a typical plateau terrain, with high altitude, low temperature, hypobaric environment, lack of oxygen, and high ultraviolet radiation making the Tibetan population in this region have a unique natural selection and cultural background. So it is important to study the genetic polymorphism and genetic structure for A-STR markers of the Tibetan population in the region. To provide richer genetic data, we selected a total of 36 A-STR loci containing all CODIS loci and 23 non-CODIS loci for the Tibetan population in the NTQAP, of which SE33 has not been reported in the Tibetan group.

In the current research, the STR typing test of healthy unrelated individuals from the Ngawa Tibetan population calculated that the TDP and CPE of the 36 A-STRs of this group are 1–2.2552 $\times$ 10^{-42} and 1–1.3031 $\times$ 10^{-16}, respectively, with highly informative polymorphisms. The LD values reflected the non-random association of alleles between different markers at different population levels and showed these A-STR markers can be utilized to compute the corresponding forensic parameters. In the Ngawa Tibetan population, the LD test results revealed no significant LD between A-STR loci except for the D13S317 and D18S535 loci after Bonferroni correction ($p = 0.05/630 = 0.00008$). Since D13S317 and D18S535 loci are not on the same chromosome, they should be independent of each other, which may be caused by random sampling error. Among the 36 A-STR markers in the Ngawa Tibetan group, 345 alleles were identified, with allele frequencies ranging from 0.00382 to 0.55343 and the allelic numbers ranging from 5 in both the TH01 locus and TPOX locus to 28 in the SE33 locus. Hardy–Weinberg equilibrium is the most important fundamental law in population genetics and can be used to describe the effects of reproduction on the gene frequency and genotype frequency of a population. We noted that without Bonferroni correction, all 36 A-STR loci of this population obeyed HWE ($p = 0.05$). The 36 A-STR loci could offer highly informative polymorphisms for forensic genetics and play fundamental roles in the research of population comparison and population differentiation.

For polymorphic genetic markers, it is essential to study their allele frequencies and genotype frequencies in specific populations. Compared with other populations, the Ngawa Tibetan population had a larger number of alleles and a wide range of allele frequencies. Among the six Tibetan populations, the Hobs and DP values were quite high in the Ngawa Tibetan population. The 36 A-STR loci had the highest values of TDP and CPE in the Ngawa Tibetan population compared to the other 26 Chinese-related populations. Among these Chinese populations, the TPOX locus usually had the lowest values of Hobs, DP, PE, and PIC and the highest value of MP, while

the opposite was true for the Penta E locus. As in most populations, the TPOX locus was similar in the Ngawa Tibetan population. It is noteworthy that SE33 had the highest values for forensic parameters such as DP, PIC, PE, TPI, Hobs, and Heterozygotes in the Tibetan population. Moreover, the SE33 locus also had the smallest value of MP, all of which shows its great potential for forensic applications for the Ngawa Tibetan population. The SE33 locus is highly polymorphic, and its high genetic variability and heterozygote nature also make its forensic application challenging. Bhinder et al. [56] showed that the addition of the SE33 locus to forensic DNA profiling aided forensic identification. Borsuk et al. [57] showed the importance of careful evaluation and confirmation of SE33 alleles and improved sequence recognition of SE33 in capillary electrophoresis. Therefore, the availability of relevant population genetic data for the SE33 locus is extremely valuable for its application in the Tibetan population. Because there are differences in the number of alleles, allelic frequencies, and forensic parameters in A-STR loci in different populations, this study may provide genetic diversity and available markers for forensic identification of the Ngawa Tibetan population.

The population genetic distance matrix is a visual matrix that measures genetic differences and is generally used to determine the similarity of different populations. In this study, a Nei's standard genetic distance matrix was constructed based on 15 A-STR loci common to 27 populations in China, and a heatmap was constructed accordingly. The results indicated a relatively close genetic distance between Tibetan populations. We constructed the MDS plot based on genetic distance matrices and explored the genetic distance and evolutionary differentiation of the populations after dimensionality reduction. It was found that Tibetan groups are scattered at a distance from other groups, and Han Chinese are mostly clustered in the middle of this plot. Among them, the Ngawa Tibetan showed a closer genetic relationship with the Gansu Gannan Tibetan and the Tibet Chamdo Tibetan. The NJ evolutionary tree constructed from the genetic distance matrix showed small evolutionary differences between the Ngawa and other Tibetan groups, but the ethnic groups showed large genetic evolutionary differences from the Han Chinese, demonstrating that evolutionary relationships are closely related to the geographical location. We also normalized the allele frequencies of the 27 populations to construct a PCA plot. The results were similar to the MDS plot, with Han groups clustered into clusters and with close genetic distances for the more geographically proximate groups such as the Sichuan Han and Chongqing Han. Although Tibetan populations were far away from other populations, they also had relatively close genetic distances from each other. Based on the Fst values and corresponding *p*-values of these 15 A-STR loci between populations, we found significant population differentiation and genetic distance between the Ngawa Tibetan population and the Hainan Li, the Fujian She, the Yunnan Yi, the Yunnan Han, and the Sichuan Chengdu Tibetan populations.

The study showed that Han Chinese populations form a genetic group, differing from several other Chinese groups such as the She, Yi, Tibetan, Uyghur, Hui, and Li, and the genetic relationship between Han Chinese groups in various provinces is consistent with their geographic distribution. In the analysis of genetic distance and genetic evolution, it can be clearly seen that the Han population is divided into two clusters: the southern Han population and the northern Han population. However, there are also several non-Han Chinese groups such as the Xinjiang Xibe and the Ningxia Hui which are close to the Han genetic group. We can find that these Tibetan populations have retained their own genetic characteristics and are distant from the Han Chinese and ethnic minorities, forming their own unique substructure and diversity. Among the neighboring populations, the Tibetan-Yi corridor and western Chinese populations showed relatively close genetic relationships with the Ngawa Tibetan population, such as the Gansu Gannan Tibetan, the Tibet Chamdu Tibetan, the Sichuan Yi, the Ningxia Hui, and the Xinjiang Xibe. In addition, the genetic relationship between the northern Han populations or neighboring Han populations and the Ngawa Tibetan population is closer than other Han populations, such as the Shandong Han, the Qinghai Han, the Shanxi Yuncheng Han, and the Hebei Han.

In the STRUCTURE analysis, we observed no significant differences among the eight neighboring populations at the optimal K value ($K = 2$). Since there is little difference in the color composition of each vertical line, this implies that these populations share a similar ancestry. Therefore, the Ngawa Tibetan population and other neighboring populations did not have a clear population structure in this analysis. The results may be influenced by geographic proximity and most populations come from the same language family (Sino-Tibetan language). For population structure and population comparisons of the Ngawa Tibetan group and other Chinese groups, our genetic results showed its complex demographic history and potential substructure within the population, which plays vital roles in the evolution, integration, and differentiation of many populations. Based on the results of these analyses, we can speculate that Tibetan populations migrated and integrated less after prehistoric migration and retained the genetic characteristics of the group, showing clear genetic differentiation from other groups and genetic differences among Tibetan populations within different areas of the Tibetan-Yi corridor. Overall, studies of population structure, genetic differentiation, and phylogenetic relationship have shown that geographical distribution, historical factors, and cultural integration have influenced both genetic structure and ethnic group diversity.

In conclusion, our study represents the first investigation into the allelic frequencies and correlated forensic parameters in the Ngawa Tibetan population in the Tibetan-Yi corridor by 36 A-STR markers, which enriches the A-STR data in forensic database and provides highly polymorphic genetic information for the application forensic science. The 36 A-STR loci, particularly for the SE33 locus, have rich genetic polymorphisms and can be used as powerful genetic markers for individual identification, parentage testing, and population genetic analysis of the Ngawa Tibetan population. This study provides basic data of STR genetic markers for the forensic genetics identification of the Ngawa Tibetan population in southwest China, thereby ensuring the scientific accuracy of the identification results and providing support for the promotion of the application of their genetic research results in judicial expertise in the NTQAP. Moreover, we not only reveal the substructures and diversities of the Tibetan population and other Chinese populations but also show the integration and differentiation of the Chinese populations.

Ethics statement

This study was reviewed and approved by the ethics committee of Southwest Medical University, with the approval number: KY2021168.

Data availability statement

All data used for the analyses in this report are available from the corresponding author upon reasonable request.

Funding

This work was funded by the Research Foundation of Luzhou City (No. 2021-SYF-37), the Teaching Reform Project of Postgraduate Education in Southwest Medical University (No. YJG202222), and partially funded by the National Natural Science Foundation of China (Nos. 82073263, 81672887, and 30371493).

CRediT authorship contribution statement

Binghui Song: Formal analysis, Investigation, Writing - original draft, Writing - review & editing. **Jiewen Fu:** Investigation, Methodology, Writing - review & editing. **Kan Guo:** Investigation, Project administration. **Jie Qian:** Investigation, Writing - review & editing. **Ting He:** Investigation, Writing - review & editing. **Lisha Yang:** Resources, Writing - review & editing. **Jingliang Cheng:** Formal analysis, Investigation, Methodology, Supervision, Visualization, Writing - review & editing. **Junjiang Fu:** Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing - original draft, Writing - review & editing.

Declaration of competing interest

The authors declare no any competing interests.

Acknowledgements

The authors thank people in the Research Center for Preclinical Medicine, Southwest Medical University. We truly thank Dr. Mazaher Maghsoudloo for editing our manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e23005.

References

[1] B.R. McCord, Q. Gauthier, S. Cho, M.N. Roig, G.C. Gibson-Daw, B. Young, F. Taglia, S.C. Zapico, R.F. Mariot, S.B. Lee, G. Duncan, Forensic DNA analysis, Anal. Chem. 91 (1) (2019) 673–688.
[2] B. Song, J. Fu, J. Qian, L. Yang, J. Cheng, J. Fu, Genetic polymorphism and population genetic structure analysis of 21 autosomal STR loci for a han-Chinese population from Luzhou of Southwest China, Genes 14 (7) (2023) 1419.
[3] N. Wyner, M. Barash, D. McNevin, Forensic autosomal short tandem repeats and their potential association with phenotype, Front. Genet. 11 (2020) 884.
[4] D. Lu, H. Lou, K. Yuan, X. Wang, Y. Wang, C. Zhang, Y. Lu, X. Yang, L. Deng, Y. Zhou, Q. Feng, Y. Hu, Q. Ding, Y. Yang, S. Li, L. Jin, Y. Guan, B. Su, L. Kang, S. Xu, Ancestral origins and genetic history of Tibetan highlanders, Am. J. Hum. Genet. 99 (3) (2016) 580–594.
[5] X. Qi, C. Cui, Y. Peng, X. Zhang, Z. Yang, H. Zhong, H. Zhang, K. Xiang, X. Cao, Y. Wang, Ouzhuluobu, Ciwangsangbu Basang, Bianba, T. Wu Gonggalanzi, H. Chen, H. Shi, B. Su, Genetic evidence of paleolithic colonization and neolithic expansion of modern humans on the Tibetan plateau, Mol. Biol. Evol. 30 (8) (2013) 1761–1778.
[6] G. He, M. Wang, X. Zou, P. Chen, Z. Wang, Y. Liu, H. Yao, L.H. Wei, R. Tang, C.C. Wang, H.Y. Yeh, Peopling history of the Tibetan Plateau and multiple waves of admixture of Tibetans inferred from both ancient and modern genome-wide data, Front. Genet. 12 (2021), 725243.
[7] X.Y. Yang, A. Rakha, W. Chen, J. Hou, X.B. Qi, Q.K. Shen, S.S. Dai, X. Sulaiman, N.T. Abdulloevich, M.E. Afanasevna, K.B. Ibrohimovich, X. Chen, W.K. Yang, A. Adnan, R.H. Zhao, Y.G. Yao, B. Su, M.S. Peng, Y.P. Zhang, Tracing the genetic legacy of the Tibetan empire in the Balti, Mol. Biol. Evol. 38 (4) (2021) 1529–1536.
[8] M. Zhang, S. Yan, W. Pan, L. Jin, Phylogenetic evidence for Sino-Tibetan origin in northern China in the late neolithic, Nature 569 (7754) (2019) 112–115.
[9] L. Liu, J. Chen, J. Wang, Y. Zhao, X. Chen, Archaeological evidence for initial migration of neolithic proto Sino-Tibetan speakers from yellow river valley to Tibetan plateau, Proc. Natl. Acad. Sci. U.S.A. 119 (51) (2022), e2212006119.
[10] L.L. Kang, J. Zhao, K. Liu, X. Li, Allele frequencies of 15 STR loci of Tibetan lived in Tibet Lassa, Forensic Sci. Int. 168 (2–3) (2007) 236–240.
[11] J. Yan, C. Shen, Y. Li, X. Yu, X. Xiong, H. Mu, Y. Huang, Y. Deng, J. Yu, Genetic analysis of 15 STR loci on Chinese Tibetan in Qinghai province, Forensic Sci. Int. 169 (1) (2007) e3–e6.
[12] M. Ota, Y. Droma, B. Basnyat, Y. Katsuyama, H. Asamura, H. Sakai, H. Fukuhsima, Allele frequencies for 15 STR loci in Tibetan populations from Nepal, Forensic Sci. Int. 169 (2–3) (2007) 234–238.
[13] L.L. Kang, K. Liu, Y. Ma, Y chromosome STR haplotypes of Tibetan living Tibet lassa, Forensic Sci. Int. 172 (1) (2007) 79–83.
[14] H. Gou, J. Zhou, Y. Zhang, C. Wang, X. Ma, H. Ma, S. Li, X. Hu, M. Shang, J. Zhang, M. Wang, Y. Zhou, Y. Ye, X. Song, J. Wang, X. Lu, B. Ying, Allele frequency distribution of 21 forensic autosomal STR loci of Goldeneye DNA ID 22NC Kit in Chinese Tibetan group, Forensic Sci Int Genet 22 (2016) e21–e24.
[15] 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 (2) (2018) 409–413.
[16] S. Cao, P. Bai, W. Zhu, D. Chen, H. Wang, B. Jin, L. Zhang, W. Liang, Genetic portrait of 27 Y-STR loci in the Tibetan ethnic population of the Qinghai province of China, Forensic Sci Int Genet 34 (2018) e18–e19.
[17] F. Song, M. Xie, B. Xie, S. Wang, M. Liao, H. Luo, Genetic diversity and phylogenetic analysis of 29 Y-STR loci in the Tibetan population from Sichuan Province, Southwest China, Int. J. Leg. Med. 134 (2) (2020) 513–516.

[18] G. Li, Y. Lin, S. Lan, J. Zou, S. Li, F. Song, Y. Ye, Tibetan Y-STR trait in the eleven regions of the Qinghai-Tibet Plateau, Int. J. Leg. Med. 135 (5) (2021) 1793–1795.
[19] F.H. Chen, G.H. Dong, D.J. Zhang, X.Y. Liu, X. Jia, C.B. An, M.M. Ma, Y.W. Xie, L. Barton, X.Y. Ren, Z.J. Zhao, X.H. Wu, M.K. Jones, Agriculture facilitated permanent human occupation of the Tibetan Plateau after 3600 B.P, Science 347 (6219) (2015) 248–250.
[20] G. He, Z. Wang, Y. Su, X. Zou, M. Wang, X. Chen, B. Gao, J. Liu, S. Wang, Y. Hou, Genetic structure and forensic characteristics of Tibeto-Burman-speaking U-Tsang and Kham Tibetan Highlanders revealed by 27 Y-chromosomal STRs, Sci. Rep. 9 (1) (2019) 7739.
[21] G.Y. Fan, Z.Q. Zhang, P.Z. Tang, D.L. Song, X.K. Zheng, Y.J. Zhou, M.N. Liu, Forensic and phylogenetic analyses of populations in the Tibetan-Yi corridor using 41 Y-STRs, Int. J. Leg. Med. 135 (3) (2021) 783–785.
[22] P.S. Walsh, D.A. Metzger, R. Higushi, Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material, BioTechniques 10(4): 506-513 (April 1991), Biotechniques 54 (3) (2013) 134–513.
[23] J. Fu, S. Shen, J. Cheng, H. Lv, J. Fu, A case of Usher syndrome type IIA caused by a rare USH2A homozygous frameshift variant with maternal uniparental disomy (UPD) in a Chinese family, J. Cell Mol. Med. 24 (14) (2020) 7743–7750.
[24] B. Olaisen, W. Bar, B. Brinkmann, B. Budowle, A. Carracedo, P. Gill, P. Lincoln, W.R. Mayr, S. Rand, DNA recommendations 1997 of the international society for forensic genetics, Vox Sang. 74 (1) (1998) 61–63.
[25] L. Excoffier, H.E. Lischer, Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows, Mol Ecol Resour 10 (3) (2010) 564–567.
[26] P. Chen, B. Wang, G. He, F. Song, H. Luo, L. Zhang, J. Yu, Y. Han, Population genetic analyses of 20 autosomal STR loci in Chinese Han population from Tianjin in Northern China, Forensic Sci Int Genet 27 (2017) 184–185.
[27] L. Sun, K. Shi, L. Tan, Q. Zhang, L. Fu, X. Zhang, G. Fu, S. Li, B. Cong, Analysis of genetic polymorphisms and mutations at 19 STR loci in Hebei Han population, Forensic Sci Int Genet 31 (2017) e50–e51.
[28] D. Wang, T. Zhang, R. Xu, Z. Dong, C. Li, W. Cui, Genetic polymorphisms of 23 STR loci of Goldeneye 25A kit in Shandong Han population, Forensic Sci Int Genet 40 (2019) e256–e258.
[29] H. Gao, J. Yu, X. Feng, X. Wu, L. Luo, X. Li, C. Liu, P. Chen, Genetic polymorphism of 23 autosomal STR loci in han population from Yuncheng, Shanxi province, Zhong Nan Da Xue Xue Bao Yi Xue Ban 46 (4) (2021) 351–360.
[30] J. Xie, C. Shao, Y. Zhou, W. Zhu, H. Xu, Z. Liu, Z. Zhao, Genetic distribution on 20 STR loci from the Han population in Shanghai, China, Forensic, Sci Int Genet 9 (2014) e30–e31.
[31] Y. Wang, F. Chen, Y. Fu, S. Zhang, H. Zhang, M. Lai, E. Xu, Genetic polymorphisms of 19 autosomal STR loci in 3510 individuals from Han population of Zhejiang province, Southeast China, Forensic Sci. Int. 306 (2020), 110045.
[32] Y. Lu, P. Song, J.C. Huang, X. Wu, Genetic polymorphisms of 20 autosomal STR loci in 5141 individuals from the Han population of Xiamen, Southeast China, Forensic Sci Int Genet 29 (2017) e31–e32.
[33] C. Xiao, W. Zhang, T. Wei, C. Pan, D. Huang, Population data of 21 autosomal STR loci in Chinese Han population from Hubei province in Central China, Forensic Sci Int Genet 20 (2016) e13–e14.
[34] Y. Zou, J.J. Guo, Q.P. Li, D.H. Zuo, J.S. Liu, Y.D. Guo, J. Yan, L. Zha, J.F. Cai, L.M. Lan, Genetic polymorphisms of 21 STR loci in hunan province-based han population, Fa Yi Xue Za Zhi 32 (5) (2016) 356–362.
[35] X. Zou, Y. Li, P. Li, Q. Nie, T. Wang, Y. Hu, Y. Zhu, J. Li, R. Tang, Genetic polymorphisms for 19 autosomal STR loci of Chongqing Han ethnicity and phylogenetic structure exploration among 28 Chinese populations, Int. J. Leg. Med. 131 (6) (2017) 1539–1542.
[36] G. He, Y. Li, Z. Wang, W. Liang, H. Luo, M. Liao, J. Zhang, J. Yan, Y. Li, Y. Hou, J. Wu, Genetic diversity of 21 autosomal STR loci in the Han population from Sichuan province, Southwest China, Forensic Sci Int Genet 31 (2017) e33–e35.
[37] L. Yang, X. Zhang, L. Zhao, Y. Sun, J. Li, R. Huang, L. Hu, S. Nie, Population data of 23 autosomal STR loci in the Chinese Han population from Guangdong Province in southern China, Int. J. Leg. Med. 132 (1) (2018) 133–135.
[38] X. Zhang, L. Liu, R. Xie, G. Wang, Y. Shi, T. Gu, L. Hu, S. Nie, Population data and mutation rates of 20 autosomal STR loci in a Chinese Han population from Yunnan Province, Southwest China, Int. J. Leg. Med. 132 (4) (2018) 1083–1085.
[39] Z. Wang, B. Lu, X. Jin, J. Yan, H. Meng, B. Zhu, Genetic and structural characterization of 20 autosomal short tandem repeats in the Chinese Qinghai Han population and its genetic relationships and interpopulation differentiations with other reference populations, Forensic Sci Res 3 (2) (2018) 145–152.
[40] Z. Li, J. Zhang, H. Zhang, Z. Lin, J. Ye, Genetic polymorphisms in 18 autosomal STR loci in the Tibetan population living in Tibet Chamdo, Southwest China, Int. J. Leg. Med. 132 (3) (2018) 733–734.
[41] H.B. Yao, C.C. Wang, J. Wang, X. Tao, L. Shang, S.Q. Wen, Q. Du, Q. Deng, B. Xu, Y. Huang, H.D. Wang, S. Li, C. Bin, L. Ma, L. Jin, J. Krause, H. Li, Genetic structure of Tibetan populations in Gansu revealed by forensic STR loci, Sci. Rep. 7 (2017), 41195.
[42] G. He, Z. Wang, Y. Su, X. Zou, M. Wang, J. Liu, Y. Hou, Genetic variation and forensic characterization of highland Tibetan ethnicity reveled by autosomal STR markers, Int. J. Leg. Med. 132 (4) (2018) 1097–1102.
[43] 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 (1) (2019) 3371.
[44] 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 (1) (2021), 16327.
[45] X. Zhang, L. Hu, L. Du, H. Zheng, A. Nie, M. Rao, J. Bo Pang, S. Nie, Population data for 20 autosomal STR loci in the Yi ethnic minority from Yunnan Province, Southwest China, Forensic Sci Int Genet 28 (2017) e43–e44.
[46] H. Fan, X. Wang, Z. Ren, G. He, R. Long, A. Liang, T. Song, J. Deng, Population data of 19 autosomal STR loci in the Li population from Hainan Province in southernmost China, Int. J. Leg. Med. 133 (2) (2019) 429–431.
[47] 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 (5) (2014) 314–318.
[48] 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 (3) (2016) 687–688.
[49] 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 (6) (2016) 1501–1503.
[50] M.K. Shimada, T. Nishida, A modification of the PHYLIP program: a solution for the redundant cluster problem, and an implementation of an automatic bootstrapping on trees inferred from original data, Mol. Phylogenet. Evol. 109 (2017) 409–414.
[51] J.G. Moberly, M.T. Bernards, K.V. Waynant, Key features and updates for origin 2018, J Cheminform 10 (1) (2018) 5.
[52] S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, X. Mega, Molecular evolutionary genetics analysis across computing platforms, Mol. Biol. Evol. 35 (6) (2018) 1547–1549.
[53] J. Novembre, Stephens Pritchard, Donnelly on population structure, Genetics 204 (2) (2016) 391–393.
[54] G. Evanno, S. Regnaut, J. Goudet, Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study, Mol. Ecol. 14 (8) (2005) 2611–2620.
[55] M. Jakobsson, N.A. Rosenberg, CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure, Bioinformatics 23 (14) (2007) 1801–1806.
[56] M.A. Bhinder, M.Y. Zahoor, H. Sadia, M. Qasim, R. Perveen, G.M. Anjum, M. Iqbal, N. Ullah, W. Shehzad, M. Tariq, A.M. Waryah, SE33 locus as a reliable genetic marker for forensic DNA analysis systems, Turk. J. Med. Sci. 48 (3) (2018) 611–614.
[57] L.A. Borsuk, K.B. Gettings, C.R. Steffen, K.M. Kiesler, P.M. Vallone, Sequence-based US population data for the SE33 locus, Electrophoresis 39 (21) (2018) 2694–2701.