FORENSIC SCIENCES RESEARCH
2022, VOL. 7, NO. 2, 163–171
https://doi.org/10.1080/20961790.2020.1751379

ORIGINAL ARTICLE

OPEN ACCESS Check for updates

Comprehensive analyses for genetic diversities of 19 autosomal STRs in Chinese Kazak group and its phylogenetic relationships with other continental populations

Yijie Wang[a,b], Xiaoye Jin[a,b,c], Wenqing Zhang[a,b], Wei Cui[a,b,c], Tingting Kong[a,b], Chong Chen[a,b,c], Yuxin Guo[a,b,c], Haotian Meng[a,b,c] and Bofeng Zhu[a,b]

[a]Key Laboratory of Shaanxi Province for Craniofacial Precision Medicine Research, College of Stomatology, Xi'an Jiaotong University, Xi'an, China; [b]Clinical Research Center of Shaanxi Province for Dental and Maxillofacial Diseases, College of Stomatology, Xi'an Jiaotong University, Xi'an, China; [c]College of Forensic Medicine, Xi'an Jiaotong University Health Science Center, Xi'an, China

ABSTRACT

Short tandem repeats (STRs) play an essential role in forensic genetics due to their high degree of polymorphisms, wide distributions and easy detection method. In this study, allelic frequencies and forensic statistical parameters of the 19 autosomal STR loci in a Kazak ethnic group were calculated, and its genetic relationships with reference populations were assessed in order to understand population structure better and enrich population genetic data for forensic practice in Chinese Kazak ethnic group. There were 226 identified alleles with the corresponding allelic frequencies ranging from 0.0008 to 0.5295 in the 628 unrelated healthy Kazak individuals in Xinjiang Uygur Autonomous Region. All autosomal STRs were conformed to the Hardy-Weinberg equilibrium after Bonferroni's correction. The cumulative power of discrimination and the combined probability of exclusion of all the 19 autosomal STRs were 0.999 999 999 999 999 999 999 997 162 and 0.999 999 994 484, respectively. Furthermore, the D_A distances and Fixation index values of pairwise populations, principal component analysis, multidimensional scaling analysis, phylogenetic tree analysis and structure analysis were conducted to probe the genetic relationships between the Kazak group and other reference populations. The population genetic results showed that these 19 autosomal STR loci were characterised by high genetic diversities in the Kazak group. Furthermore, the studied Kazak group had close genetic relationships with the Uyghur group and the Uzbek group. The present results may facilitate understanding the genetic background of the Chinese Xinjiang Kazak group.

ARTICLE HISTORY
Received 29 August 2019
Accepted 6 March 2020

KEYWORDS
Forensic sciences; forensic genetics; Kazak ethnic group; autosomal STRs; population genetics; interpopulation differentiation; phylogenetic reconstruction

Introduction

The short tandem repeats (STRs), also called microsatellite DNA, are a kind of genetic markers with a repeat unit of 2–6 bp [1]. As the most common genetic marker, STRs have wide applications in forensic DNA analyses because of their wide distributions in human genome, high polymorphisms and following the Mendelian codominant inheritance [2].

The Kazak is an ethnic group in China. Kazak individuals were about 1 462 500 in China by the data from the sixth national population census in 2010 [3]. Most Kazaks reside in Xinjiang Uygur Autonomous region, Gansu and Qinghai provinces. Kazaks generally believe in Islam, and have a written language themselves. Their language belongs to the Turkic branch of the Altaic language family.

To extend the population genetic database of the Kazak ethnic minority and reveal its exclusively genetic background, the individuals of the Kazak group in the Nilka county in Northwest China were analysed by 19 autosomal STR loci of the AGCU® EX20 kit (AGCU ScienTech, Wuxi, China). In order to illustrate population genetic differentiations between the Kazak group and other reference populations published previously, the data acquired above were compared with the same available 15 autosomal STR loci of these reference populations.

Material and methods

DNA extraction and ethics statement

Kazak individuals have been recruited according to the personal information provided by the participants. Bloodstain samples (peripheral blood) were obtained from 628 independent healthy Kazak individuals living in the Nilka county in Northwest China. Written informed consents were signed by all the participants in this study. Genomic DNA extraction was conducted

CONTACT Bofeng Zhu zhubofeng7372@126.com

Supplemental data for this article are available online at https://doi.org/10.1080/20961790.2020.1751379.

with the Chelex-100 method [4]. The research was performed with the permission of the human and ethical research principles approved by the Ethics Committee of Xi'an Jiaotong University Health Science Center, China. A sample collection procedure was proceeded strictly to insure that no blood kinships existed within three generations of analysed individuals.

Polymerase chain reaction and STR typing

Nineteen autosomal STR loci and an amelogenin gene were co-amplified using the AGCU® EX20 kit (AGCU ScienTech). Temperature cycling condition for polymerase chain reaction (PCR) comprised the initial denaturation for 95 °C at 2 min; 10 cycles for 94 °C at 30 s, 60 °C at 1 min and 72 °C at 1 min; and then 20 cycles for 90 °C at 30 s, 58 °C at 1 min and 72 °C at 1 min; and a final elongation cycle for 72 °C at 10 min. Then the PCR products of autosomal STRs were analysed using the ABI 3500 Genetic Analyzer (Thermo Fisher Scientific, Waltham, MA, USA). The STR genotyping analyses were performed by GeneMapper ID-X software v3.2 (Applied Biosystems, Foster City, CA, USA). Deionised water was selected as the negative control and DNA 9947 A was the positive control.

Statistical analyses

Allelic frequencies, forensic statistical parameters and linkage disequilibrium (LD) tests of 19 autosomal STR loci of the Kazak group were calculated by the online programme (STRAF 1.0.5: STR Analysis for Forensics) [5]. Then pairwise D_A distances and fixation index (F_{st}) values based on the same 15 autosomal STR loci (19 autosomal STR loci with the exception of the D12S391, D6S1043, Penta D and Penta E loci) were performed with Dispan programme [6] and Genepop software v4.0.10 [7], respectively. The D_A distances and F_{st} values could be used for subsequent analyses. SPSS v19.0 (IBM Corp, Armonk, NY, USA) was applied to obtain the result of multidimensional scaling (MDS) on the basis of the pairwise F_{st} values, and the analysis of population structure was performed with STRUCTURE programme v2.3.4 [8] based on the raw data of the same 15 autosomal STR loci of the Kazak group and 14 reference populations. In addition, the heatmaps of pairwise D_A and F_{st} values of the Kazak group and 14 reference populations were generated using ggplot2 package in *R* statistical software v3.4.4 (https://www.r-project.org/). Moreover, according to the data of allelic frequencies of the same 15 autosomal STR loci, the principal component analysis (PCA) was yielded with the XLSTAT programme (https://www.xlstat.com/en/), and MEGA software v6.0 [9] was performed to construct phylogenetic tree of the Kazak group and other reference populations based on the data of D_A distances by the neighbour-joining method.

Results and discussion

Hardy-Weinberg equilibrium tests and linkage disequilibrium analyses of 19 autosomal STR loci

As illustrated in Table 1, there were no significant deviations from the Hardy-Weinberg equilibrium (HWE) at all 19 autosomal STR loci after Bonferroni's correction ($P > 0.002632 = 0.05/19$) [10], with the lowest probability values of tests (P_{HWE}) at locus Penta E (0.0180). In Supplementary Table S1, LD tests of 171 pairs at 19 autosomal STR loci were carried out in the Kazak group using the online software (STRAF 1.0.5: STR Analysis for Forensics). After Bonferroni's correction ($P = 0.05/171$), there were no significant LDs of most pairwise autosomal STR loci except four pairs, including *D12S391/D7S820* ($P = 0.0001$), *D18S51/D19S433* ($P < 0.0001$), *D19S433/D8S1179*

Table 1. Forensic statistical parameters of the 19 STR loci in the Kazak ethnic group ($n = 628$).

Loci	MP	PD	PIC	PE	TPI	H_O	H_E	P_{HWE}
CSF1PO	0.1135	0.8865	0.6903	0.4569	1.7740	0.7182	0.7377	0.1750
D12S391	0.0408	0.9592	0.8357	0.7431	3.9747	0.8742	0.8532	0.9540
D13S317	0.0586	0.9414	0.7946	0.6429	2.8288	0.8232	0.8203	0.6930
D16S539	0.0643	0.9357	0.7812	0.5832	2.3969	0.7914	0.8098	0.6870
D18S51	0.0342	0.9658	0.8461	0.7303	3.7831	0.8678	0.8613	0.7530
D19S433	0.0582	0.9418	0.7934	0.6398	2.8036	0.8217	0.8167	0.2600
D21S11	0.0597	0.9403	0.7875	0.6038	2.5323	0.8025	0.8098	0.7300
D2S1338	0.0297	0.9703	0.8607	0.7018	3.4130	0.8535	0.8743	0.2830
D3S1358	0.1133	0.8867	0.6944	0.5016	1.9625	0.7452	0.7396	0.8910
D5S818	0.1093	0.8907	0.7070	0.5572	2.2429	0.7771	0.7449	0.4440
D6S1043	0.0362	0.9638	0.8472	0.7527	4.1316	0.8790	0.8632	0.9910
D7S820	0.0691	0.9309	0.7708	0.5861	2.4154	0.7930	0.8013	0.7800
D8S1179	0.0605	0.9395	0.7861	0.6038	2.5323	0.8025	0.8111	0.7720
FGA	0.0379	0.9621	0.8387	0.6861	3.2371	0.8455	0.8557	0.4000
Penta D	0.0520	0.9480	0.8049	0.6582	2.9623	0.8312	0.8278	0.8750
Penta E	0.0120	0.9880	0.9216	0.8307	6.0385	0.9172	0.9272	0.0180
TH01	0.0805	0.9195	0.7489	0.5098	2.0000	0.7500	0.7838	0.2470
TPOX	0.1905	0.8095	0.5729	0.3128	1.3083	0.6178	0.6277	0.2820
vWA	0.0678	0.9322	0.7750	0.5774	2.3609	0.7882	0.8035	0.0830

MP: matching probability; PD: power of discrimination; PIC: polymorphism information content; PE: probability of exclusion; TPI: typical paternity index; H_O: observed heterozygosity; H_E: expected heterozygosity; P_{HWE}: probability values of Hardy–Weinberg equilibrium tests.

($P < 0.0001$) and *D7S820/D8S1179* ($P < 0.0001$). LDs of pairwise loci might be related with some factors like population stratification, genetic drift, natural selection, genetic linkage and so on [11]. However, these four pairwise autosomal STRs were located on the different chromosomes. The specific reason of such phenomenon needs further verification.

Allelic frequencies and forensic statistical parameter analyses

Raw genotypes of 19 autosomal STR loci applied in the studied Kazak group were shown in Supplementary Table S2. The allelic frequencies of all 226 alleles ranged from 0.0008 to 0.5295 (Table 2). As shown in Table 1, the matching probability (MP) ranged from 0.0120 at locus *Penta E* to 0.1905 at locus *TPOX*. Among all of the loci, *Penta E* showed the maximum values of power of discrimination (PD) (0.9880), polymorphism information content (PIC) (0.9216), probability of exclusion (PE) (0.8307), typical paternity index (TPI) (6.0385), observed heterozygosity (H_O) (0.9172) and expected heterozygosity (H_E) (0.9272). Conversely, the minimum values of the above six parameters observed at locus *TPOX* were 0.8095, 0.5729, 0.3128, 1.3083, 0.6178 and 0.6277, respectively. In addition, the cumulative PD (CPD) value of all 19 autosomal STR loci was 0.999 999 999 999 999 999 999 997 162, and the cumulative PE (CPE) value of these loci was 0.999 999 994 484. Based on the results, these 19 autosomal STR loci could be used for individual identifications in the Kazak group.

Population genetic analyses of the studied Kazak group and 14 reference populations based on the same 15 autosomal STR loci

There were 14 reference populations of which raw genotyping could be available. The autosomal STRs data of 14 reference populations published previously consisted of three continents, namely, one population from Europe (Portuguese [12]), five from America (Mexican [13], Asian American, African American, Caucasian American and Hispanics American [14]), and eight from China (Beijing Han [15], Sichuan Han [16], Inner Mongolian Russian [17], Qinghai Tibetan [18], Ningxia Hui [19], Chinese Xibe [20], Uyghur [21] and Uzbek groups [22]). Therefore, we assessed genetic relationships between the Kazak group and these 14 reference populations.

First, PCA among these populations was performed by XLSTAT programme. As illustrated in Figure 1A, the first two principal components occupied 70.58% of the total variance, in detail, the first component occupied 56.73% and second component occupied 13.85%. On the other hand, the population distributions were distinctly depicted in Figure 1A. The 15 populations were predominantly distributed in four different regions. African American distributed in the upper left quadrant; Portuguese, Caucasian American, Mexican and Hispanics American clustered together and distributed in the left; Asian American, Beijing Han, Sichuan Han, Inner Mongolian Russian, Qinghai Tibetan, Ningxia Hui and Xibe clustered together and spread in the right; interestingly, the Kazak group clustered nearly to the Uzbek and Uyghur groups and located in the middle of the graph. The distribution indicated that the Kazak group has close genetic relationships with these two groups. To further investigate genetic associations among these populations, a dimensionality reduction analysis named MDS was carried out according to the computational pairwise F_{st} values (Figure 1B). This figure reflected four main population clusters, which was consistent with the PCA plot shown in Figure 1A. Moreover, the studied Kazak group was closer with the Uzbek and Uyghur groups in comparison with the other populations. The results proposed that the genetic relationships between the Kazak ethnic group and these two groups were considerably closer than other populations.

Next, phylogenetic analysis among the above populations was conducted, as displayed in Figure 2. Two main clusters could be easily found. The first one mainly consisted of the American populations (marked by blue colour) and the European population (marked by yellow), and the second one was mainly composed of the Chinese populations (marked by red). The results showed that the cluster of 15 populations was approximately consistent with their distributions of geographical locations. Furthermore, the studied Kazak group labelled by a red pentagram was closely clustered with Uyghur, Uzbek groups and the other Chinese populations.

D_A distance is an indicator of genetic differences between different species or between different populations of the same species [23]. Similarly, F_{st} value is commonly used for analyzing the population differentiations. In order to visually reflect genetic differentiations of these 15 populations, heatmaps of D_A distances and F_{st} values of pairwise populations were provided. As illustrated in Supplementary Figure S1A, the colour scales ranging from grey to orange stood for the D_A distance variations of pairwise populations on each block. The colour of the block closer to light orange represented bigger D_A distance variations and more genetic differentiations of the pairwise populations, and *vice versa*. Results revealed that grey colours were observed between the studied Kazak and Uyghur and Uzbek groups, implying less genetic differentiations among these groups. Similarly, the colour scales ranging from light blue to pink denoted the F_{st} value

Table 2. Allelic frequencies of the 19 STR loci in the Kazak ethnic group ($n = 628$).

Alleles	*CSF1PO*	*D12S391*	*D13S317*	*D16S539*	*D18S51*	*D19S433*	*D21S11*	*D2S1338*	*D3S1358*	*D5S818*	*D6S1043*	*D7S820*	*D8S1179*	*FGA*	*Penta D*	*Penta E*	*TH01*	*TPOX*	*vWA*
5																0.0350			
6	0.0008														0.0239		0.1887		
7	0.0008									0.0295		0.0056			0.0048	0.0533	0.2420		
8			0.1943	0.0295						0.0016		0.2309	0.0072		0.0143	0.0111	0.1083	0.5295	
9	0.0398		0.1162	0.2086	0.0016					0.0613	0.0016	0.0804	0.0032		0.2667	0.0111	0.2906	0.1107	
9.2						0.0008													
9.3																	0.1624		
10	0.2723		0.1385	0.1282	0.0040					0.1011	0.0183	0.2126	0.0844		0.1465	0.0597	0.0080	0.0382	
10.2						0.0008													
11	0.2739		0.2293	0.2389	0.0080	0.0032				0.3877	0.1959	0.2564	0.0597		0.1807	0.0876		0.2779	
11.3												0.0016							
12	0.3248		0.2293	0.2269	0.0454	0.0549				0.2643	0.1791	0.1632	0.1274		0.1704	0.0963		0.0390	
12.2						0.0008													
13	0.0812		0.0621	0.1449	0.1911	0.2381		0.0008	0.0008	0.1433	0.1099	0.0446	0.3049		0.1234	0.0701		0.0048	0.0064
13.2						0.0350													
14	0.0064	0.0016	0.0303	0.0223	0.2317	0.2938			0.0725	0.0080	0.0860	0.0048	0.2142		0.0470	0.0637			0.1139
14.2						0.0764													
14.3						0.0008													
15		0.0151		0.0008	0.1377	0.1003			0.3368	0.0032	0.0064		0.1481		0.0175	0.0860			0.0621
15.2						0.1401													
16		0.0104			0.1306	0.0207		0.0088	0.3049				0.0414	0.0008	0.0048	0.1123			0.2086
16.2						0.0263													
17		0.1107			0.0709	0.0024		0.0804	0.2094		0.0510		0.0064	0.0008		0.0995			0.2866
17.2						0.0040													
17.3		0.0104																	
18		0.1943			0.0430			0.1083	0.0740		0.1401		0.0032	0.0151		0.0828			0.2134
18.2						0.0016													
18.3		0.0064																	
19		0.2102			0.0478			0.1990	0.0016		0.1473			0.0382		0.0334			0.0939
19.1																0.0008			
19.3		0.0080																	
20		0.1903			0.0271			0.1545			0.0565			0.0725		0.0223			0.0151
20.2														0.0016					
21		0.1019			0.0239			0.0287			0.0040			0.1115		0.0231			
21.2														0.0008					
21.3											0.0032								
22		0.0645			0.0191			0.0350						0.1306		0.0247			
22.2														0.0008					
22.3											0.0008								
23		0.0342			0.0143			0.1537						0.2006		0.0135			
23.2														0.0080					
24		0.0247			0.0032			0.0924						0.2189		0.0056			
24.2														0.0032					
25		0.0143			0.0008		0.0008	0.0979						0.1330		0.0072			
25.2														0.0008					
26		0.0032						0.0271						0.0502		0.0008			
27							0.0072	0.0032						0.0096					
28							0.0796	0.0104						0.0032					
28.2							0.0064												

(*continued*)

Table 2. Continued.

Alleles	*CSF1PO*	*D12S391*	*D13S317*	*D16S539*	*D18S51*	*D19S433*	*D21S11*	*D2S1338*	*D3S1358*	*D5S818*	*D6S1043*	*D7S820*	*D8S1179*	*FGA*	*Penta D*	*Penta E*	*TH01*	*TPOX*	*vWA*
29							0.1990												
29.2							0.0008												
30							0.3368												
30.2							0.0135												
30.3							0.0008												
31							0.0955												
31.2							0.0756												
32							0.0183												
32.2							0.1218												
33							0.0040												
33.1							0.0016												
33.2							0.0342												
34.2							0.0040												

variations of pairwise populations, as presented in Supplementary Figure S1B. The colour of the block closer to light blue represented smaller F_{st} values and less genetic differentiations of the pairwise populations, and *vice versa*. We observed that the block between Kazak and Uyghur or Uzbek groups mainly displayed lighter blue colours, which meant small genetic discrepancies between Kazak and these two groups. From the perspective of ethnic origin, the Kazaks, Uyghurs and Uzbeks are descendants of the Turkic tribes [24]. Especially, the Kazaks lived in mixed habitations with the Hans, Uyghurs and Mongolians [25]. Accordingly, there were more opportunities for genetic exchange due to the adjacent areas between Kazak group and Uzbek and Uyghur groups. Similar study results could be discerned from the previous research [26].

STRUCTURE analysis is a tool for inferring the genetic structure of a population using genetic markers. Population structure analysis was constructed among Kazak group and these 14 reference populations with K values defined at 2–6. As shown in Figure 3, the area between every two vertical black thin lines represented the genetic structure of one population. When $K = 2$, these 15 populations were roughly divided into three clusters on the basis of different colour compositions: four American populations and one European population (blue component was larger than orange component), Asian American and six Chinese populations (orange was larger than blue), and the other three Chinese populations including Kazak, Uyghur and Uzbek groups (blue and orange for about half-half), respectively. When $K = 3$, the result was basically the same as $K = 2$. Moreover, Kazak, Uyghur and Uzbek groups showed similar ancestral proportions from $K = 2$–6. Their compositions were much similar to each other in comparison with other continental populations. The Kazaks, Uyghurs and Uzbeks are important members of the ancient Silk Road of China. Yuan et al. [27] investigated the Uyghur group from Southern Xinjiang of China based on 38 STR loci. The above results revealed that the genetic distance between Uyghur group and Kazak group was close. Similarly, Xie et al. [28] found that Xinjiang Kazak ethnic group had the closer genetic relationship with the Uyghur group based on the genetic polymorphisms of 60 mitochondrial DNA (mtDNA) loci of Xinjiang Kazak group in China.

Genetic relationship explorations of the studied Kazak group and 20 reference populations based on allelic frequencies of 15 overlapping STRs

The above analyses were based on the raw data or allelic frequencies of 14 reference populations that

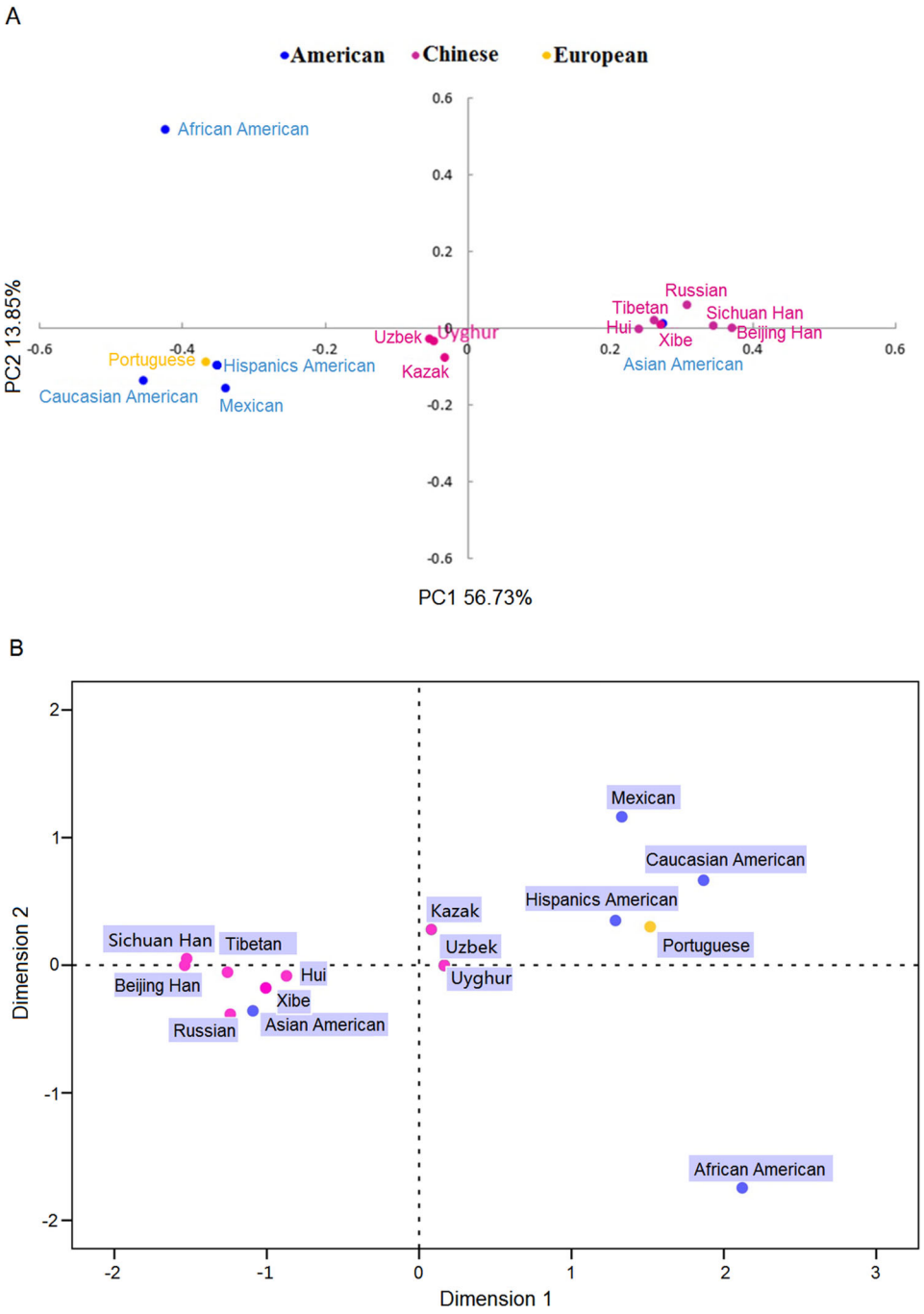

Figure 1. A principal component analysis (PCA) plot (A) was carried out based on allelic frequencies of 15 STR loci with XLSTAT programme, and a two-dimensional multidimensional scaling (MDS) plot (B) was generated based on pairwise F_{st} values of the Kazak group and 14 reference populations by SPSS v19.0.

could be obtained. To explore the genetic background of the studied Kazak group in Nilka county, the genetic relationships between the Kazak group and 20 reference populations (only allelic frequencies were available without raw data) were performed based on allelic frequencies of the same 15 autosomal STR loci mentioned above. These reference populations included Northern Han [29], Shanghai Han [30], Guangdong Han [31], Henan Han [32], Hui [33], Tibetan [34], Ewenki [35], Kazak in Xinjiang [36], Kazak in Barkol [37], Zhuang [38], Dai [38], Li [39], Yi [40], Dong [41], She [42], Bouyei [43], Tujia [44], Gelao [44], Italian [45] and Estonian [46] populations.

Supplementary Figure S2 showed the first two principal components occupied to 63.82% of the total variance, equivalently, the first, second components occupied 51.07%, 12.75%, respectively. It was obvious from the figure that these populations were distributed into three clusters with most Chinese populations distributed in the left, European populations distributed in the right, and Kazak groups in different regions distributed between the two population clusters mentioned above. As expected, Kazak

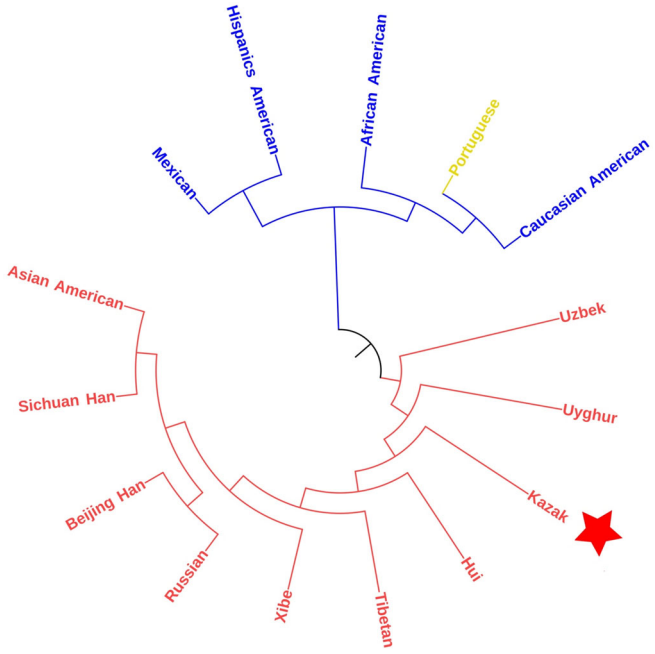

Figure 2. A neighbour-joining tree showing the genetic relationships of the Kazak group and 14 reference populations was constructed based on pairwise D_A distances by MEGA software v6.0.

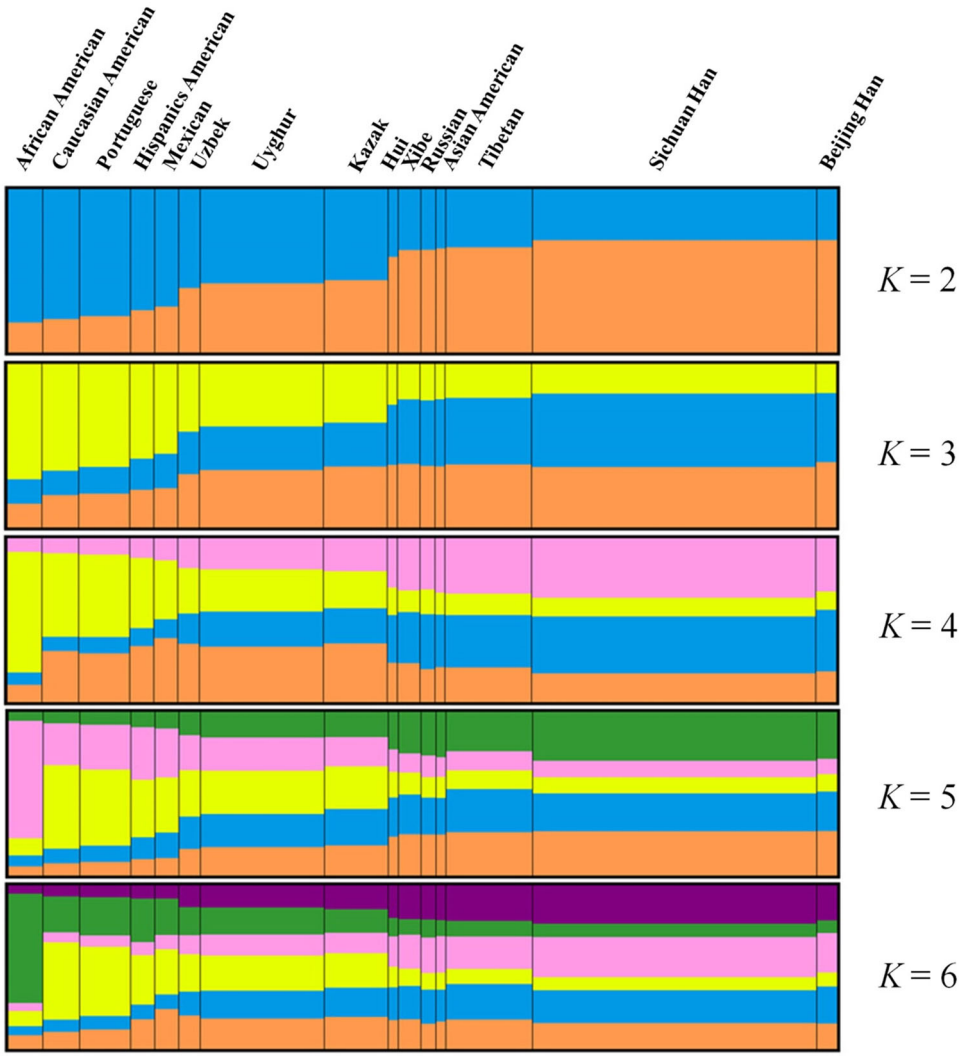

Figure 3. Population structure analysis was performed based on the raw data of the same 15 autosomal STR loci of the Kazak group and 14 reference populations *via* STRUCTURE programme v2.3.4 ($K = 2$–6). Population names were labeled above the plot.

group in Nilka county was located closely to Kazak in Xinjiang and Kazak in Barkol in the PCA plot, indicating that the Kazak groups in these three different regions had closer genetic relationships. As shown in Supplementary Figure S3, the similar results were also seen in phylogenetic tree constructed by MEGA software v6.0, which showed the genetic relationships of the Kazak group in Nilka county and 20 reference populations. For the above results, we found that the studied Kazak group had closer genetic affinities with the Kazak groups living in different regions.

Conclusion

This project was undertaken to assess the genetic diversities of 19 autosomal STR loci of Kazak group in Nilka county. Furthermore, the phylogenetic relationships between Kazak group and other reference populations based on the same 15 autosomal STR loci were investigated. The following points emerged from the present investigation: (1) These 19 loci were highly polymorphic in the studied Kazak group and could be forward to apply for paternity testing and individual identification. (2) The Kazak groups in different regions had close genetic relationships. (3) The studied Kazak group was also closely related with Uyghur and Uzbek groups. Overall, the current study made several noteworthy contributions to studying the genetic background of Kazak group.

Acknowledgements

We are grateful to the donors for participating in this study.

Authors' contributions

Bofeng Zhu participated in the design of the study and revised the final draft of the manuscript. Yijie Wang performed the majority of the experiments and wrote the manuscript. Xiaoye Jin and Wenqing Zhang performed a part of experiments and statistical analyses. Wei Cui, Tingting Kong, Chong Chen, Yuxin Guo and Haotian Meng collected the samples and helped proofread. All authors contributed to the final text and approved it.

Compliance with ethical standards

This study was performed in accordance with the human and ethical research principles approved by the Ethics Committee of Xi'an Jiaotong University Health Science Center, China, and all procedures were approved by the Ethics Committee of Xi'an Jiaotong, University, China. Written informed consents were signed by all the participants for sample collection and use in this study.

Disclosure statement

The authors declare that they have no conflict of interest.

References

[1] Ruitberg CM, Reeder DJ, Butler JM. STRBase: a short tandem repeat DNA database for the human identity testing community. Nucleic Acids Res. 2001;29:320–322.
[2] Edwards A, Civitello A, Hammond HA, et al. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am J Hum Genet. 1991;49:746–756.
[3] National Bureau of Statistics of China. Sixth National Population Census of China (NPCC). Beijing (China): National Bureau of Statistics of China. 2012. http://www.stats.gov.cn/tjsj/pcsj/rkpc/6rp/indexch.htm
[4] Walsh PS, Metzger DA, Higuchi R. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. BioTechniques. 2013;54:506–513.
[5] Gouy A, Zieger M. STRAF—A convenient online tool for STR data evaluation in forensic genetics. Forensic Sci Int Genet. 2017;30:148–151.
[6] Ota A, Ota A. DISPAN: genetic distance and phylogenetic analysis. State College (PA): Pennsylvania State University; 1993.
[7] Rousset F. GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. J Mol Ecol Resour. 2008;8:103–106.
[8] Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000;155:945–959.
[9] Tamura K, Peterson D, Peterson N, et al. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 2011;28:2731–2739.
[10] Goeman JJ, Solari A. Multiple hypothesis testing in genomics. Statist Med. 2014;33:1946–1978.
[11] Slatkin M. Linkage disequilibrium — understanding the evolutionary past and mapping the medical future. Nat Rev Genet. 2008;9:477–485.
[12] Almeida C, Ribeiro T, Oliveira AR, et al. Population data of the GlobalFiler® Express loci in South Portuguese population. Forensic Sci Int Genet. 2015;19:39–41.
[13] Aguilar-Velázquez JA, Martínez-Cortés G, Inclán-Sánchez A, et al. Population data of 23 STR loci (PowerPlex® Fusion System) in Mexican Mestizos from the West Region. Int J Legal Med. 2016;130:1489–1491.
[14] Hill CR, Duewer DL, Kline MC, et al. U.S. population data for 29 autosomal STR loci. Forensic Sci Int Genet. 2013;7:e82–e83.
[15] Ruan X, Wang WN, Yang YR, et al. [Genetic variability and phylogenetic analysis of 39 short tandem repeat loci in Beijing Han population]. Yi Chuan. 2015;37:683–691. Chinese.
[16] He G, Li Y, Wang Z, et al. Genetic diversity of 21 autosomal STR loci in the Han population from Sichuan province, Southwest China. Forensic Sci Int Genet. 2017;31:e33–e35.
[17] Zhu BF, Yang G, Shen CM, et al. Population genetic analysis of 15 autosomal STR loci in the Russian population of northeastern Inner-Mongolia, China. Mol Biol Rep. 2010;37:3889–3895.

[18] Yan J, Shen C, Li Y, et al. Genetic analysis of 15 STR loci on Chinese Tibetan in Qinghai Province. Forensic Sci Int. 2007;169:e3–e6.
[19] Wang Z, Huang P, Chen L, et al. Genetic polymorphisms of 15 STR loci in Chinese Hui population. J Forensic Sci. 2005;50:1–9.
[20] Meng H, Guo Y, Dong Q, et al. Autosomal-STR based genetic structure of Chinese Xibe ethnic group and its relationships to various groups. Int J Legal Med. 2016;130:1501–1503.
[21] Jin X, Wei Y, Chen J, et al. Phylogenic analysis and forensic genetic characterization of Chinese Uyghur group *via* autosomal multi STR markers. Oncotarget. 2017;8:73837–73845.
[22] Jin X, Cui W, Wei Y, et al. Genetic characteristics of 19 STRs in Chinese Uzbek ethnic and its phylogenetic relationships with other 24 populations. Int J Legal Med. 2018;132:729–731.
[23] Takezaki N, Nei M. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. Genetics. 1996;144:389–399.
[24] Palstra FP, Heyer E, Austerlitz F. Statistical inference on genetic data reveals the complex demographic history of human populations in central Asia. Mol Biol Evol. 2015;32:1411–1424.
[25] Sheng X, Wang Y, Zhang J, et al. Forensic investigation of 23 autosomal STRs and application in Han and Mongolia ethnic groups. Forensic Sci Res. 2018;3:138–144.
[26] Yuan JY, Wang XY, Shen CM, et al. Genetic profile characterization and population study of 21 autosomal STR in Chinese Kazak ethnic minority group. Electrophoresis. 2014;35:503–510.
[27] Yuan L, Liu H, Liao Q, et al. Genetics analysis of 38 STR loci in Uygur population from Southern Xinjiang of China. Int J Legal Med. 2016;130:687–688.
[28] Xie T, Hu L, Guo YX, et al. Genetic polymorphism analysis of mitochondrial DNA from Chinese Xinjiang Kazak ethnic group by a novel mitochondrial DNA genotyping panel. Mol Biol Rep. 2019;46:17–25.
[29] Xie B, Chen L, Yang Y, et al. Genetic distribution of 39 STR loci in 1027 unrelated Han individuals from Northern China. Forensic Sci Int Genet. 2015;19:205–206.
[30] Xie J, Shao C, Zhou Y, et al. Genetic distribution on 20 STR loci from the Han population in Shanghai, China. Forensic Sci Int Genet. 2014;9:e30–e31.
[31] Yang L, Zhang X, Zhao L, et al. Population data of 23 autosomal STR loci in the Chinese Han population from Guangdong Province in southern China. Int J Legal Med. 2018;132:133–135.
[32] Wang HD, Wu D, Feng ZQ, et al. Genetic polymorphisms of 20 short tandem repeat loci from the Han population in Henan, China. Electrophoresis. 2014;35:1509–1514.
[33] Deng YJ, Zhu BF, Shen CM, et al. Genetic polymorphism analysis of 15 STR loci in Chinese Hui ethnic group residing in Qinghai province of China. Mol Biol Rep. 2011;38:2315–2322.
[34] Wang Z, Zhou D, Jia Z, et al. Developmental Validation of the Huaxia Platinum System and application in 3 main ethnic groups of China. Sci Rep. 2016;6:31075.
[35] Chen L, He Y, Li S. Genetic analysis of 15 STR loci of Chinese Ewenki ethnic population. J Forensic Sci. 2006;51:1408–1409.
[36] Zhang H, Yang S, Guo W, et al. Population genetic analysis of the GlobalFiler STR loci in 748 individuals from the Kazakh population of Xinjiang in northwest China. Int J Legal Med. 2016;130:1187–1189.
[37] Zhang L, Xu C, Chen H, et al. [Genetic polymorphism of 19 STR loci in Xinjiang Barkol Kazakh population]. J Cent South Univ. 2012;37:934–938. Chinese.
[38] Zhang X, Du L, Huang Z, et al. Genetic variation of 20 autosomal STR loci in three ethnic groups (Zhuang, Dai and Hani) in the Yunnan province of southwestern China. Forensic Sci Int Genet. 2017;31:e41–e42.
[39] Chen J, Xie B, Yang Y, et al. Genetic variability and forensic efficiency of 39 microsatellite loci in the Li ethnic group from Hainan Island in South China Sea. Ann Hum Biol. 2017;44:467–474.
[40] Li J, Guo F, He W, et al. Allele frequencies of 15 autosomal STR loci in the Yi ethnic minority from Yunnan Province in the far southwest of China. Forensic Sci Int Genet. 2017;26:e23–e24.
[41] Zhang L. Population data for 15 autosomal STR loci in the Dong ethnic minority from Guizhou Province, Southwest China. Forensic Sci Int Genet. 2015;16:237–238.
[42] Yuan L, Ou Y, Liao Q, et al. Population genetics analysis of 38 STR loci in the She population from Fujian Province of China. Legal Med. 2014;16:314–318.
[43] Zhang L. Population data for 15 autosomal STR loci in the Bouyei ethnic minority from Guizhou Province, Southwest China. Forensic Sci Int Genet. 2015;17:108–109.
[44] Sun H, Xu S, Long F, et al. Forensic and population genetic analysis of Han, Miao, Tujia and Gelao populations from Zunyi (Southwest China) on 15 autosomal short tandem repeat loci. Forensic Sci Int Genet. 2016;25:e20–e21.
[45] Turrina S, Ferrian M, Caratti S, et al. Evaluation of genetic parameters of 22 autosomal STR loci (PowerPlex® Fusion System) in a population sample from Northern Italy. Int J Legal Med. 2014;128:281–283.
[46] Sadam M, Tasa G, Tiidla A, et al. Population data for 22 autosomal STR loci from Estonia. Int J Legal Med. 2015;129:1219–1220.