Genetic polymorphisms and forensic efficiency of 19 X chromosomal STR loci for Xinjiang Mongolian population

Ling Chen [1], **Yuxin Guo** [2,3], **Cheng Xiao** [1], **Weibin Wu** [1], **Qiong Lan** [1], **Yating Fang** [1], **Jiangang Chen** [4], **Bofeng Zhu** Corresp. [1,2,3]

[1] Southern Medical University, Department of Forensic Biology, School of Forensic Medicine, Guangzhou, Guangdong, China

[2] Xi'an Jiaotong University, Key Laboratory of Shaanxi Province for Craniofacial Precision Medicine Research, College of Stomatology, Xi'an, Shaanxi, China

[3] Xi'an Jiaotong University, Clinical Research Center of Shaanxi Province for Dental and Maxillofacial Diseases, College of Stomatology, Xi'an, Shaanxi, China

[4] Ministry of Public Security, Institute of Forensic Science, Beijing, China

Corresponding Author: Bofeng Zhu
Email address: zhubofeng@mail.xjtu.edu.cn

Aims: X-chromosomal short tandem repeat (X-STR) loci are playing an increasingly important role in some complex kinship cases in recent years. To investigate the forensic efficiency of X chromosomal STRs of Mongolian minority group from Xinjiang Uygur Autonomous Region, China, and further depict the genetic relationship among Xinjiang Mongolians and other populations, 267 blood samples from unrelated healthy Xinjiang Mongolians were amplified by AGCU X-19 STR kit.

Results: No deviations for all 19 X-STR loci were observed from the HWE after Bonferroni correction (P>0.0026) in female samples. The most frequent allele was allele 10 at locus DXS10164 with the frequency 0.5663. The PIC values of the 19 X-STR loci were more than 0.5 with the highest polymorphism at the locus DXS10135. The cumulative power of discrimination were 0.99999999999999999999988761005481 in females and 0.999999999999903 in males, respectively; and the cumulative mean exclusion chances were 0.9999999969738068321121 in duos and 0.999999999998952 in trios, respectively. The 7 linkage groups were extremely informative, with all the haplotype diversities greater than 0.9487. No linkage disequilibrium was observed for a significance level of 0.00029 (p=0.05/171) after Bonferroni correction. The MDS plot, phylogenetic tree based on the 11 overlapping X-STR loci all presented that the Xinjiang Mongolian population was genetically different from other Asian populations, including the Mongolians from Inner Mongolia Autonomous Region, China

Conclusion: This study indicated that the 19 X-STR multiplex PCR system was of high utility value for both forensic practices and population genetic researches in Xinjiang Mongolian group.

1 **Title:**

2 Genetic polymorphisms and forensic efficiency of 19 X chromosomal STR loci for Xinjiang Mongolian

3 Population

4 **Running Title:**

5 Analysis of 19 X-STR in Mongolian

7 Ling Chen[1,#], Yuxin Guo[2,3,#], Cheng Xiao[1], Weibin Wu[1], Qiong Lan[1], Yating Fang[1], Jiangang Chen[4], Bofeng Zhu[1,2,3,*]

9 [1]Department of Forensic Biology, School of Forensic Medicine, Southern Medical University, Guangzhou, Guangdong 510515, P. R. China

10 [2]Key Laboratory of Shaanxi Province for Craniofacial Precision Medicine Research, College of Stomatology, Xi'an Jiaotong University, Xi'an, Shaanxi

11 710004, P. R. China

12 [3]Clinical Research Center of Shaanxi Province for Dental and Maxillofacial Diseases, College of Stomatology, Xi'an Jiaotong University, Xi'an, Shaanxi

13 710004, P. R. China

14 [4]Institute of Forensic Science, Ministry of Public Security, Beijing, 100038, P. R. China

15 [#] These authors contributed equally to this work.

Abstract

18   **Aims:** X-chromosomal short tandem repeat (X-STR) loci are playing an increasingly important role in some complex kinship cases in

19   recent years. To investigate the forensic efficiency of X chromosomal STRs of Mongolian minority group from Xinjiang Uygur

20   Autonomous Region, China, and further depict the genetic relationship among Xinjiang Mongolians and other populations, 267 blood

21   samples from unrelated healthy Xinjiang Mongolians were amplified by AGCU X-19 STR kit.

22   **Results:** No deviations for all 19 X-STR loci were observed from the HWE after Bonferroni correction (P>0.0026) in female samples.

23   The most frequent allele was allele 10 at locus DXS10164 with the frequency 0.5663. The PIC values of the 19 X-STR loci were more

24   than 0.5 with the highest polymorphism at the locus DXS10135. The cumulative power of discrimination were

25   0.99999999999999999999988761005481 in females and 0.999999999999903 in males, respectively; and the cumulative mean

26   exclusion chances were 0.99999999969738068321121 in duos and 0.999999999998952 in trios, respectively. The 7 linkage groups were

27   extremely informative, with all the haplotype diversities greater than 0.9487. No linkage disequilibrium was observed for a significance

28   level of 0.00029 (p=0.05/171) after Bonferroni correction. The MDS plot, phylogenetic tree based on the 11 overlapping X-STR loci all

29   presented that the Xinjiang Mongolian population was genetically different from other Asian populations, including the Mongolians

30   from Inner Mongolia Autonomous Region, China.

31 **Conclusion:** This study indicated that the 19 X-STR multiplex PCR system was of high utility value for both forensic practices and

32 population genetic researches in Xinjiang Mongolian group.

33

Introduction

35 As one of the most famous nomadic nations in the world, the Mongols mainly live in East-Central Asian. With a population of 5.8

36 million, Mongolian was the eighth largest ethnic minority on Chinese population data according to the 6th China population census in

37 2010 (http://www.stats.gov.cn/tjsj/pcsj/rkpc/6rp/indexch.htm). Their indigenous dialects are known as the Mongolian language, which

38 belongs to Altaic language family. The traditional Mongolian script was created in the early 13th century based on the script of Huihu

39 (Janhunen 2011). Based on Chinese historical texts, the ancestry of Mongolians can be traced back to Donghu, a nomadic confederation

40 of tribes with same ethnic origin, different dialects and names. After the Donghu were defeated by Xiongnu, the Xianbei and Wuhuan

41 survived as the main remnants of the confederation. As recorded by the Chinese histories, Xianbei split into three prominent groups: the

42 Rouran, the Khitan and the Shiwei. A subtribe of Shiwei, called "Shiwei Menggu", was held to be the origin of Mogolians. In the

43 thirteenth century, Genghis Khan united a large group of Mongolic-speaking tribes and founded the Mongol Empire. With the expansion

44 of the Mongol Empire, the Mongolians settled over almost all Eurasia and carried on military campaigns from the Adriatic Sea to

45 Indonesian Java island and from Japan to Palestine. In the late 14th century, Mongolia was divided into two parts: Western Mongolia

46   (Oirats) and Eastern Mongolia (Khalkha, Inner Mongols, Barga, Buryats) after the fall of the Mongol Empire. In the modern period, the

47   most prominent Mongol groups are the Inner Mongols concentrated in Inner Mongolia Autonomous Region and Northeast China, the

48   Oirats concentrated in Xinjiang Uyghur Autonomous Region of China, and the Khalkha, known as Outer Mongols, concentrated in State

49   of Mongolia, and approximately sixty percent Mongolians in the world settled in China.

50   In recent years, autosomal short tandem repeats (STRs) have been extensively applied in paternity testing and individual identification.

51   However, autosomal STRs may not be effective in some deficiency paternity cases, such as tests of half-sister without the father's DNA,

52   grandmother-granddaughter without the parent's DNA, paternal aunt-niece or maternal aunt/maternal uncle-nephew without the parent's

53   and grandparent's DNA. X-STRs have special characteristics that make them particularly useful in the forensic cases described above

54   (Chen et al. 2014; Szibor 2007; Szibor et al. 2003b). Therefore, X-STR loci have drawn more and more attention in forensic sciences

55   and been reported its special role for forensic purpose (Liu et al. 2013; Sun et al. 2013). At present, X chromosomal genetic information

56   of Mongolians had been reported by Hou et al. using 9 X-STR Loci, Zhang et al. using 34 X-markers (18-STRs and 16-Indels), and Tao

57   et al. using 12 X-STR loci (Hou et al. 2007; Tao et al. 2018; Zhang et al. 2015). However, these researches only involved in Inner

58   Mongolians, and few studies had aimed at Western Mongols. AGCU-X19 STR kit was a recently developed multiplex amplification

59   system, which can amplify 19 X-chromosomal STR loci simultaneously, including DXS8378, DXS7423, DXS10148, DXS10159,

60   DXS6809, DXS7424, DXS10164, DXS10162, DXS7132, DXS10079, DXS6789, DXS101, DXS10103, DXS10101, HPRTB,

61 DXS10075, DXS10074, DXS10135 and DXS10134. And Yang et al had indicated it could be used as a supplementary tool in kinship

62 tests in China (Yang et al. 2016). Further, genetic polymorphisms of this X-STR system have been investigated in a number of

63 populations (Liu et al. 2017; Meng et al. 2017; Yang et al. 2017; Zhang et al. 2016). In the present study, we used this 19 X-STR

64 multiplex PCR system to obtain genetic information from Western Mongols, including the allele frequencies and forensic parameters of

65 these 19 X-STR loci, the haplotypic diversities of 7 X-STR linkage groups. And we also estimated the population differentiations

66 between the Western Mongols and other previously reported groups.

67

68 **Materials and Methods**

69 **Population samples**

70 The blood samples were collected from 267 unrelated healthy Xinjiang Mongolians (156 males and 111 females) living in Bortala

71 Mongol Autonomous Prefecture, Xinjiang Uyghur Autonomous Region, China. Genomic DNA was extracted using the Chelex-100

72 method (Walsh et al. 1991).

73

74 **Compliance with ethics guidelines**

75 This study was carried according to the approval for research involving human and animals by the ethical committee of Xi'an Jiaotong

76   University Health Science Center, China. Informed consents were signed by all participants prior to sample collection.

77

78   **PCR amplification and STR typing**

79   DNA samples were amplified in a 25μl PCR reaction volume using the AGCU-X19 STR kit (AGCU ScienTech Incorporation, Wuxi,

80   Jiangsu, China), following the manufacturer's instructions. The PCR products were separated through capillary electrophoresis with an

81   ABI 3500xl Genetic Analyzer (Thermo Fisher Scientific, MA, USA) and analyzed with GeneMapper ID-X (Thermo Fisher Scientific,

82   MA, USA).

83

84   **Quality control**

85   This study was carried out strictly following the ISFG recommendations on the analysis of DNA polymorphisms (Schneider 2007) and

86   guidelines for publication of population data (Carracedo et al. 2013; Carracedo et al. 2014). The experiments were conducted under the

87   laboratory internal control standards. Negative control (deionized water for amplified reaction) and positive control (the 9947A DNA

88   sample, Promega, Madison, WI, USA) were genotyped along with each batch of samples.

89

90   **Statistical analyses**

91   Allelic frequencies of all the 19 X-STR loci were calculated using the PowerStats v1.2 program (Promega, Madison, WI, USA). Hardy-

92   Weinberg equilibrium (HWE) and pair-wise linkage disequilibrium in female samples were conducted by the software of Genepop

93   (http://genepop.curtin.edu.au/). Differences of allele frequencies in males and females were assessed by the standard analysis of variance

94   (ANOVA) method using statistical software SPSS Version 19.0 (IBM, USA). And the haplotype frequencies in male samples were also

95   calculated by SPSS Version 19.0. Polymorphism information content (PIC), Heterozygosity (Het), power of exclusion (PE) , power of

96   discrimination in females (PDF), power of discrimination in males (PDM), mean exclusion chance (MEC) for deficiency cases, normal

97   trios and duo cases were calculated using the free online-calculation tool of ChrX-STR.org 2.0 database (http://www.chrx-str.org). The

98   geographical distributions of the 15 Asian populations were drew by the R software (https://www.r-project.org/). The DA distance and

99   the phylogenetic tree were constructed using Poptree2 Software (Takezaki et al. 2010). Based on the DA distances, the multidimensional

100 scaling (MDS) plot was constructed by SPSS Version 19.0.

101

102 **Results**

103 **Allelic frequencies of 19 X-STR loci**

104 A total of 156 male and 111 female individuals from Xinjiang Mongolian group were analyzed. And the raw genotypes at 19 X-STR

105 loci were displayed in Supplementary Table 1. Exact tests of HWE after Bonferroni correction (p>0.05/19=0.0026) demonstrated no

106 significant deviations for all 19 X-STR loci in female samples. There were no significant differences in the allele frequency distributions

107 between males and females by the results of One-Way ANOWA (p>0.05/19=0.0026), which indicated that the allele frequency

108 distributions of all the 19 X-STR loci had no gender bias. Therefore, the allele frequencies of males and females were pooled together

109 according to the formula (2*femalefreq + malefreq)/3, which were showed in Supplementary Table 2. A total of 223 alleles were

110 observed at the 19 X-STR loci, and the allele numbers ranged from 5 at DXS7423 to 26 at DXS10148. The most frequent allele observed

111 was allele 10 at locus DXS10164 with the frequency 0.5663. Since all positive controls were genotyped correctly, we confirmed a series

112 of off-ladder alleles for the 267 Mongolian individuals, among which allele 18.1 and 19.1 at DXS10079 and DXS10162 were more than

113 0.1

114

115 **Forensic parameters of 19 X-STR loci in Xinjiang Mongolian group**

116 Forensic statistical parameters of the 19 X-STR loci were calculated based on the pooled allele frequencies. As displayed on Table 1,

117 all the 19 X-STR loci were found to be very informative (PIC>0.5). DXS10135, with the highest PIC (0.9184), Het (0.9235), PE

118 (0.8437), PDF (0.9890) and PDM (0.9235), had the most polymorphism. The lowest PIC (0.5334), Het (0.5915), PE (0.2808), PDF

119 (0.7750) and PDM (0.5915) were observed at DXS10164, indicating that DXS10164 was of the lowest polymorphism. The cumulative

120 PDF and PDM were 0.9999999999999999999988761005481 and 0.999999999999903, respectively. The cumulative PDF and PDM

121   both showed extremely high values, which confirmed that the 19 X-STR multiplex system would be powerful for individual

122   identification in the Xinjiang Mongolian group. The MEC Krüger, MEC Kishida, and MEC Desmarais Duo ranged from 0.3406 to

123   0.8448 for the deficiency cases, 0.5334 to 0.9183 for the normal trios and 0.3867 to 0.8539 for the duo cases, respectively. The

124   cumulative MECs were 0.999999983245493 in deficiency cases, 0.999999999998952 in normal trios, and 0.9999999969738068321121

125   in duo cases, respectively. The high values of these cumulative MECs suggested that the 19 X-STR multiplex system could be useful in

126   paternity tests, especially for some deficiency cases.

127

128   **Haplotypic structure of 7 linkage groups**

129   According to the study of Zhang et al, the 19 X-STR loci could be divided into 7 linkage groups (Zhang et al. 2016). And DXS8378-

130   DXS10148-DXS10135, DXS7132-DXS10079-DXS10075-DXS10074, DXS10103-DXS10101-HPRTB and DXS10134-DXS7423

131   (http://xdb.qualitype.de/xdb/linkageTable.jsf) were classified as linkage groups 1, 2, 3 and 7, respectively. DXS6809-DXS6789 (Pasino

132   et al. 2011; Szibor et al. 2003a) was specified as linkage group 4; DXS10159-DXS10162-DXS10164 (Meng et al. 2014; Ye et al. 2014a)

133   as linkage group 5; DXS7424-DXS101 (Edelmann et al. 2002; Szibor et al. 2003a) as linkage group 6. The haplotypic frequencies and

134   haplotypic diversities of the 7 linkage groups were showed in Supplementary Table 3. In this study, 127 haplotypes were observed for

135   linkage group 1, 117 haplotypes for linkage group 2, 84 haplotypes for linkage group 3, 43 haplotypes for linkage group 4, 53 haplotypes

136   for linkage group 5, 41 haplotypes for linkage group 6, 34 haplotypes for linkage group 7. The 7 linkage groups were proved to be fairly

137   informative in the Xinjiang Mongolians because of the high values of haplotype diversities, which ranged from 0.9487 at linkage group

138   7 to 0.9969 at linkage group 1. Among these linkage groups, the most common haplotypes were H 33-20 for linkage group DXS6809-

139   DXS6789 and H 16-24 for linkage group DXS7424-DXS101 with a frequency of 0.1090, followed by H 37-15 for linkage group

140   DXS10134-DXS7423 with a frequency of 0.1026. For linkage group DXS8378-DXS10148-DXS10135, which had a large variety of

141   haplotypes with relatively low frequencies, the most common haplotype was H 10-25.1-21, with a frequency of 0.0256. However, H 10-

142   25.1-21 was not found in the Xibe population (Meng et al. 2017) and was observed in the Guanzhong Han population (Zhang et al.

143   2016) with a relatively lower frequency of 0.0040. For linkage group DXS10103-DXS10101-HPRTB, the most common haplotype in

144   the Xinjiang Mongolians was H 19-30.2-12 with a frequency of 0.0449, followed by H 16-32-13 with 0.0385; while in the Japanese

145   population (Uchigasaki et al. 2013), H 19-30.2-12 and H 16-32-13 were at a frequency of 0.0091, H 16-30-13 with a frequency of 0.0411

146   was the most common haplotype instead. For linkage group DXS10159-DXS10162-DXS10164, H 24-18-10 had the highest frequency

147   of 0.0833; however, H 26-18-10 was much more common than H 24-18-10 in the Xibe (Meng et al. 2017) and Guangxi Han population

148   (Ye et al. 2014b). For the linkage group DXS7424-DXS101, H 16-24 was of a medium frequency of 0.0403 in the Kazak population

149   (Liu et al. 2017), where H 16-25 with a frequency of 0.0872 was the most common haplotype. These haplotype comparisons indicated

150   that the studied Xinjiang Mongolian population had a different pattern of haplotype distribution, compared to other Asian populations.

151

Linkage disequilibrium analyses

153 Pair-wise linkage disequilibrium in the female samples were analyzed by Genepop software before the 19 X-STR loci were applied for

154 forensic purposes and population genetics. In this study, 171 pair-wise comparisons were performed (Supplementary Table 4) and the

155 results revealed that 10 pairs of pair-wise loci had p-value < 0.05. Nevertheless, no linkage disequilibrium were observed for a

156 significance level of 0.00029 ($p = 0.05/171$) after Bonferroni correction. This indicated that there weren't non-random association of

157 alleles at different loci in the Xinjiang Mongolian group.

158

Inter-population comparisons based on the 11 overlapping X-STRs

160 To investigate the population differentiations between the studied Xinjiang Mongolians and other Asian groups, we collected allele

161 frequencies of 11 overlapping X-STR loci (DXS8378, DXS7423, DXS10148, DXS10134, DXS7132, DXS10079, DXS10103,

162 DXS10101, DXS10074, DXS10135 and HPRTB) from Xibe (Meng et al. 2017), Kazak (Liu et al. 2017), Inner Mongolian (Tao et al.

163 2018), Japanese (Uchigasaki et al. 2013), Shanghai Han (Zhang et al. 2012), Bhil (Shrivastava et al. 2015), Malay (Samejima et al.

164 2012), Taiwanese (Chen et al. 2014), Ili Uygur (Guo et al. 2016), Guanzhong Han (Zhang et al. 2016), Tibetan (Yang et al. 2017),

165 Southern Han (Yang et al. 2017), Hui (Yang et al. 2017) and Korla Uygur (Yang et al. 2017). For a better understanding, the geographical

166   distributions of these populations mentioned above were showed in Figure 1. The D_A distance (Nei et al. 1983) was often used in forensic

167   science for measuring genetic distance between populations, and the D_A distance was more efficient in obtaining the correct tree topology

168   for microsatellite data than other distance measures including D_{ST} and F_{ST} (Takezaki & Nei 1996; Takezaki & Nei 2008). Therefore,

169   through the comparisons of allelic frequencies of the 11 overlapping X-STR loci, we calculated the D_A distances between these Asian

170   populations. As shown in Supplementary Table 5, the largest D_A distance was observed between Xinjiang Mongolian and Bhil (0.078),

171   and the smallest D_A distance between Guanzhong Han and Taiwanese (0.009). Based on these D_A distances, a multidimensional scaling

172   (MDS) plot was constructed. As shown in Figure 2, the Southern Han, Guanzhong Han, Shanghai Han, Taiwanese, Japanese, Hui,

173   Malay, as well as Inner Mongolian were clustered together at the center part of this plot; however, only the Xinjiang Mongolian was

174   located in the lower right corner, keeping a long distance from other populations. Thus, we can see that the Xinjiang Mongolian

175   population was obviously different from other Asian populations, including Inner Mongolian.

176

177   **Phylogenetic analyses based on the 11 overlapping X-STRs**

178   To evaluate the evolutionary relationship between the Xinjiang Mongolian and other 14 Asian populations, we constructed a

179   phylogenetic tree using the D_A distance and the unweighted pair-group method with arithmetic means method (UPGMA). As shown in

180   Figure 3, populations with the same ethno-linguistic origin such as Taiwanese, Guanzhong Han, Shanghai Han, Southern Han, Japanese

181 were clustered together; populations with close geographic distance like Korla Uygur, Kazak were clustered as a terminal clade; only

182 the Xinjiang Mongolian was segregated as a distant outlier, indicating that the Xinjiang Mongolian population was genetically distinct

183 from those compared Asian populations. On the whole, the phylogenetic tree showed a genetic relationship similar to the result of MDS

184 plot.

185

Discussion

187 Results of this study showed that the 19 X-STR multiplex PCR system was very useful for forensic applications and population genetic

188 researches in Xinjiang Mongolian population. Compared with Investigator Argus X-12 amplification kit (Qiagen, Hilden, Germany),

189 the AGCU 19 X-STR multiplex PCR system enriched eight new X-STR loci (DXS10159, DXS10162, DXS7424, DXS10164,

190 DXS6789, DXS101, DXS6809 and DXS10075). The 19 X-STR multiplex PCR system has been studied in several Chinese populations

191 such as Zhejiang Han (Yang et al. 2016), Guanzhong Han (Zhang et al. 2016), Kazak (Liu et al. 2017), Xibe (Meng et al. 2017), Uygur

192 (Guo et al. 2016), and these studies also indicated its potential in paternity testing and individual identification. However, there has been

193 no data of the set of 19 X-STR for Western Mongolians so far. There were about 170,000 Mongolians residing in Xinjiang Uyghur

194 Autonomous Region, China according to the 6th China population census in 2010

195 (http://www.stats.gov.cn/tjsj/pcsj/rkpc/6rp/indexch.htm). Thus, we performed a study on the polymorphism of 19 X-STRs in Xinjiang

196   Mongolians. In summary, the forensic statistical parameters of the 19 X-STR loci for Xinjiang Mongolian population showed that most

197   loci owned high PIC, Het and PE, and the new added eight X-STRs were highly polymorphic except DXS10164 (PIC=0.5334). The

198   combined power of discrimination was more than 0.999999999999 for both the female samples and the male samples. Moreover, the

199   combined mean exclusion chance was more than 0.99999999999 in normal trio cases and over 0.99999999 in duo cases. These results

200   suggested that the 19 X-STR multiplex system could provide highly polymorphic information suitable for individual identification and

201   paternity testing in Xinjiang Mongolian populations. Especially, the combined mean exclusion chance was more than 0.9999999 in

202   deficiency cases. This meant that the 19 X-STRs could be helpful for deficiency paternity cases.

203   As mentioned above, the 19 X-STR loci were grouped into 7 linkage groups. The results showed the 7 linkage groups could provide

204   highly haplotype diversity information for the Xinjiang Mongolian population. The obtained X-chromosomal haplotype frequencies of

205   the 7 linkage groups could be essential for calculate likelihood ratio of kinship. The previous researches showed the haplotyping of the

206   X-STR loci can be used to analyze some complex kinship testing (Edelmann et al. 2002; Szibor 2007). The MDS plot and the

207   phylogenetic tree based on the 11 overlapping X-STRs all presented that the Xinjiang Mongolian population was genetically different

208   from other Asian populations, including the Mongolians from Inner Mongolia Autonomous Region, China.

209   The Xinjiang Mongolians were the subgroup of the Oirats, whose ancestral home was in the Altai region of western Mongolia. The

210   Oirats share some history, geography, culture and language with the Eastern Mongols. However, they were often at war with the Eastern

211   Mongols, only few times united as a larger Mongol entity. Gradually, the Oirats and Eastern Mongols had developed separate identities

212   after the collapse of the Mongol Empire (https://en.wikipedia.org/wiki/Main_Page).

213   Previous studies based on X chromosomal STR loci indicated that Mongolians were closely related with Chinese Han groups (Hou et

214   al. 2007; Tao et al. 2018; Zhang et al. 2015). In fact, the Mongolians involved in these studies were only Inner Mongolians, descendants

215   of the Eastern Mongols. And through the analysis of autosomal SNPs, Xing et al indicated that Deedu Mongolians, living in the Qinghai-

216   Tibetan Plateau, shared genetic ancestry with other Mongolians as well as Tibetan populations (Xing et al. 2013); Nakayama et al

217   demonstrated that the largest fraction of the ancestry of Outer Mongols was mainly shared by ethnic groups in Central and Northeast

218   China (Daur, Hezhen, Oroqen, and Xibe), and by Siberians (Yakut), and this fraction was present to a lesser extent in Southern

219   populations including Japanese, Chinese Han, and other ethnic groups in Southeast Asia (Nakayama et al. 2017). Thus, we can see that

220   many factors like the sample sizes, the origins of the selected samples, the choice of the genetic markers and the coverage of reference

221   populations may lead to the discrepancy in the analysis of genetic relationship. In summary, to achieve a better understanding of genetic

222   relationships between those populations, larger sample sizes, more genetic markers, more population data and further investigations are

223   needed.

224

225   **Conclusion**

226 In this study, 267 unrelated individuals from Chinese Mongolian minority living in Xinjiang Uyghur Autonomous Region of China were

227 genotyped and analyzed for 19 X-STR loci for the first time. These loci were proved to be highly polymorphic and demonstrated high

228 power of discrimination in Xinjiang Mongolian minority. The present study indicated that the 19 X-STR multiplex PCR system could

229 be of high utility value for forensic practices, especially for the deficient paternity cases. We also found that the Xinjiang Mongolian

230 population was genetically different from other Asian populations, including the Mongolian population from Inner Mongolia

231 Autonomous Region, China.

232

Acknowledgements

234 The authors would like to thank Jiangwei Yan for helping us to analysis data.

235

References

237 Carracedo A, Butler JM, Gusmao L, Linacre A, Parson W, Roewer L, and Schneider PM. 2013. New guidelines for the publication of genetic population data.
238   *Forensic Sci Int Genet* 7:217-220. 10.1016/j.fsigen.2013.01.001
239 Carracedo A, Butler JM, Gusmao L, Linacre A, Parson W, Roewer L, and Schneider PM. 2014. Update of the guidelines for the publication of genetic population
240   data. *Forensic Sci Int Genet* 10:A1-2. 10.1016/j.fsigen.2014.01.004
241 Chen MY, Ho CW, Pu CE, and Wu FC. 2014. Genetic polymorphisms of 12 X-chromosomal STR loci in Taiwanese individuals and likelihood ratio calculations
242   applied to case studies of blood relationships. *Electrophoresis* 35:1912-1920. 10.1002/elps.201300645
243 Edelmann J, Hering S, Kuhlisch E, and Szibor R. 2002. Validation of the STR DXS7424 and the linkage situation on the X-chromosome. *Forensic Sci Int* 125:217-
244   222.

245 Guo YX, Chen JG, Wang Y, Yan JW, Chen J, Yao TH, Zhang LP, Yang G, Meng HT, Zhang YD, Mei T, Liu YS, Dong Q, and Zhu BF. 2016. Genetic
246   polymorphism analyses of a novel panel of 19 X-STR loci in the Chinese Uygur ethnic minority. *J Zhejiang Univ Sci B* 17:367-374.
247   10.1631/jzus.B1500228
248 Hou QF, Yu B, and Li SB. 2007. Genetic polymorphisms of nine X-STR loci in four population groups from Inner Mongolia, China. *Genomics Proteomics*
249   *Bioinformatics* 5:59-65. 10.1016/S1672-0229(07)60015-1
250 Janhunen J. 2011. *The Mongolic Languages*. London: Routledge.
251 Liu QL, Li ZD, Li CT, Zhao H, Wu YD, Li Q, and Lu DJ. 2013. X chromosomal recombination--a family study analyzing 26 X-STR Loci in Chinese Han three-
252   generation pedigrees. *Electrophoresis* 34:3016-3022. 10.1002/elps.201300204
253 Liu YS, Meng HT, Mei T, Zhang LP, Chen JG, Zhang YD, Chen J, Guo YX, Dong Q, Yan JW, and Zhu BF. 2017. Genetic diversity and haplotypic structure of
254   Chinese Kazak ethnic group revealed by 19 STRs on the X chromosome. *Gene* 600:64-69. 10.1016/j.gene.2016.11.018
255 Meng HT, Han JT, Zhang YD, Liu WJ, Wang TJ, Yan JW, Huang JF, Du WA, Guo JX, Wang HD, Zhang YH, Zhou RH, Zhu BF, and Wei X. 2014. Diversity
256   study of 12 X-chromosomal STR loci in Hui ethnic from China. *Electrophoresis* 35:2001-2007. 10.1002/elps.201400045
257 Meng HT, Shen CM, Zhang YD, Dong Q, Guo YX, Yang G, Yan JW, Liu YS, Mei T, Shi JF, and Zhu BF. 2017. Chinese Xibe population genetic composition
258   according to linkage groups of X-chromosomal STRs: population genetic variability and interpopulation comparisons. *Ann Hum Biol*:1-8.
259   10.1080/03014460.2017.1318951
260 Nakayama K, Ohashi J, Watanabe K, Munkhtulga L, and Iwamoto S. 2017. Evidence for Very Recent Positive Selection in Mongolians. *Mol Biol Evol* 34:1936-
261   1946. 10.1093/molbev/msx138
262 Nei M, Tajima F, and Tateno Y. 1983. Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. *J Mol Evol* 19:153-170.
263 Pasino S, Caratti S, Del Pero M, Santovito A, Torre C, and Robino C. 2011. Allele and haplotype diversity of X-chromosomal STRs in Ivory Coast. *Int J Legal*
264   *Med* 125:749-752. 10.1007/s00414-011-0591-4
265 Samejima M, Nakamura Y, Nambiar P, and Minaguchi K. 2012. Genetic study of 12 X-STRs in Malay population living in and around Kuala Lumpur using
266   Investigator Argus X-12 kit. *Int J Legal Med* 126:677-683. 10.1007/s00414-012-0705-7
267 Schneider PM. 2007. Scientific standards for studies in forensic genetics. *Forensic Sci Int* 165:238-243. 10.1016/j.forsciint.2006.06.067
268 Shrivastava P, Jain T, Gupta U, and Trivedi VB. 2015. Genetic polymorphism study on 12 X STR loci of investigator Argus X STR kit in Bhil tribal population of
269   Madhya Pradesh, India. *Leg Med (Tokyo)* 17:214-217. 10.1016/j.legalmed.2014.11.004
270 Sun K, Zhao S, Tian H, Zhang S, and Li C. 2013. Development of the 16 X-STR loci typing system and genetic analysis in a Shanghai Han population from China.
271   *Electrophoresis* 34:3008-3015. 10.1002/elps.201300234
272 Szibor R. 2007. X-chromosomal markers: past, present and future. *Forensic Sci Int Genet* 1:93-99. 10.1016/j.fsigen.2007.03.003
273 Szibor R, Edelmann J, Hering S, Plate I, Wittig H, Roewer L, Wiegand P, Cali F, Romano V, and Michael M. 2003a. Cell line DNA typing in forensic genetics--
274   the necessity of reliable standards. *Forensic Sci Int* 138:37-43.

275 Szibor R, Krawczak M, Hering S, Edelmann J, Kuhlisch E, and Krause D. 2003b. Use of X-linked markers for forensic purposes. *Int J Legal Med* 117:67-74.
276   10.1007/s00414-002-0352-5

277 Takezaki N, and Nei M. 1996. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* 144:389-399.

278 Takezaki N, and Nei M. 2008. Empirical tests of the reliability of phylogenetic trees constructed with microsatellite DNA. *Genetics* 178:385-392.
279   10.1534/genetics.107.081505

280 Takezaki N, Nei M, and Tamura K. 2010. POPTREE2: Software for constructing population trees from allele frequency data and computing other population
281   statistics with Windows interface. *Mol Biol Evol* 27:747-752. 10.1093/molbev/msp312

282 Tao R, Zhang J, Bian Y, Dong R, Liu X, Jin C, Zhu R, Zhang S, and Li C. 2018. Investigation of 12 X-STR loci in Mongolian and Eastern Han populations of
283   China with comparison to other populations. *Sci Rep* 8:4287. 10.1038/s41598-018-22665-3

284 Uchigasaki S, Tie J, and Takahashi D. 2013. Genetic analysis of twelve X-chromosomal STRs in Japanese and Chinese populations. *Mol Biol Rep* 40:3193-3196.
285   10.1007/s11033-012-2394-1

286 Walsh PS, Metzger DA, and Higuchi R. 1991. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques*
287   10:506-513.

288 Xing J, Wuren T, Simonson TS, Watkins WS, Witherspoon DJ, Wu W, Qin G, Huff CD, Jorde LB, and Ge RL. 2013. Genomic analysis of natural selection and
289   phenotypic variation in high-altitude mongolians. *PLoS Genet* 9:e1003634. 10.1371/journal.pgen.1003634

290 Yang X, Wu W, Chen L, Liu C, Zhang X, Chen L, Feng X, Wang H, and Liu C. 2016. Development of the 19 X-STR loci multiplex system and genetic analysis
291   of a Zhejiang Han population in China. *Electrophoresis* 37:2260-2272. 10.1002/elps.201500540

292 Yang X, Zhang X, Zhu J, Chen L, Liu C, Feng X, Chen L, Wang H, and Liu C. 2017. Genetic analysis of 19 X chromosome STR loci for forensic purposes in four
293   Chinese ethnic groups. *Sci Rep* 7:42782. 10.1038/srep42782

294 Ye Q, Tang J, Chen Z, Li F, Yu X, Wang P, Lin H, and Shi M. 2014a. Analysis of linkage disequilibrium and linkage for 12 short tandem repeat loci on chromosome
295   X. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* 31:782-785. 10.3760/cma.j.issn.1003-9406.2014.06.023

296 Ye Q, Tang J, Chen Z, Li F, Yu X, Wang P, Lin H, and Shi M. 2014b. [Analysis of linkage disequilibrium and linkage for 12 short tandem repeat loci on
297   chromosome X]. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* 31:782-785. 10.3760/cma.j.issn.1003-9406.2014.06.023

298 Zhang S, Bian Y, Li L, Sun K, Wang Z, Zhao Q, Zha L, Cai J, Gao Y, Ji C, and Li C. 2015. Population genetic study of 34 X-Chromosome markers in 5 main
299   ethnic groups of China. *Sci Rep* 5:17711. 10.1038/srep17711

300 Zhang S, Zhao S, Zhu R, and Li C. 2012. Genetic polymorphisms of 12 X-STR for forensic purposes in Shanghai Han population from China. *Mol Biol Rep*
301   39:5705-5707. 10.1007/s11033-011-1379-9

302 Zhang YD, Shen CM, Meng HT, Guo YX, Dong Q, Yang G, Yan JW, Liu YS, Mei T, Huang RZ, and Zhu BF. 2016. Allele and haplotype diversity of new
303   multiplex of 19 ChrX-STR loci in Han population from Guanzhong region (China). *Electrophoresis* 37:1669-1675. 10.1002/elps.201500425

304

Figure legend

306 Figure 1. The geographical distribution of the 15 Asian populations.

307 Figure 2. The multidimensional scaling (MDS) plot based on DA distance for the 15 Asian populations.

308 Figure 3. The phylogenetic tree generated by the DA distance and the unweighted pair-group method with arithmetic means method

309 (UPGMA).

310

Figure 1

The geographical distribution of the 15 Asian populations.

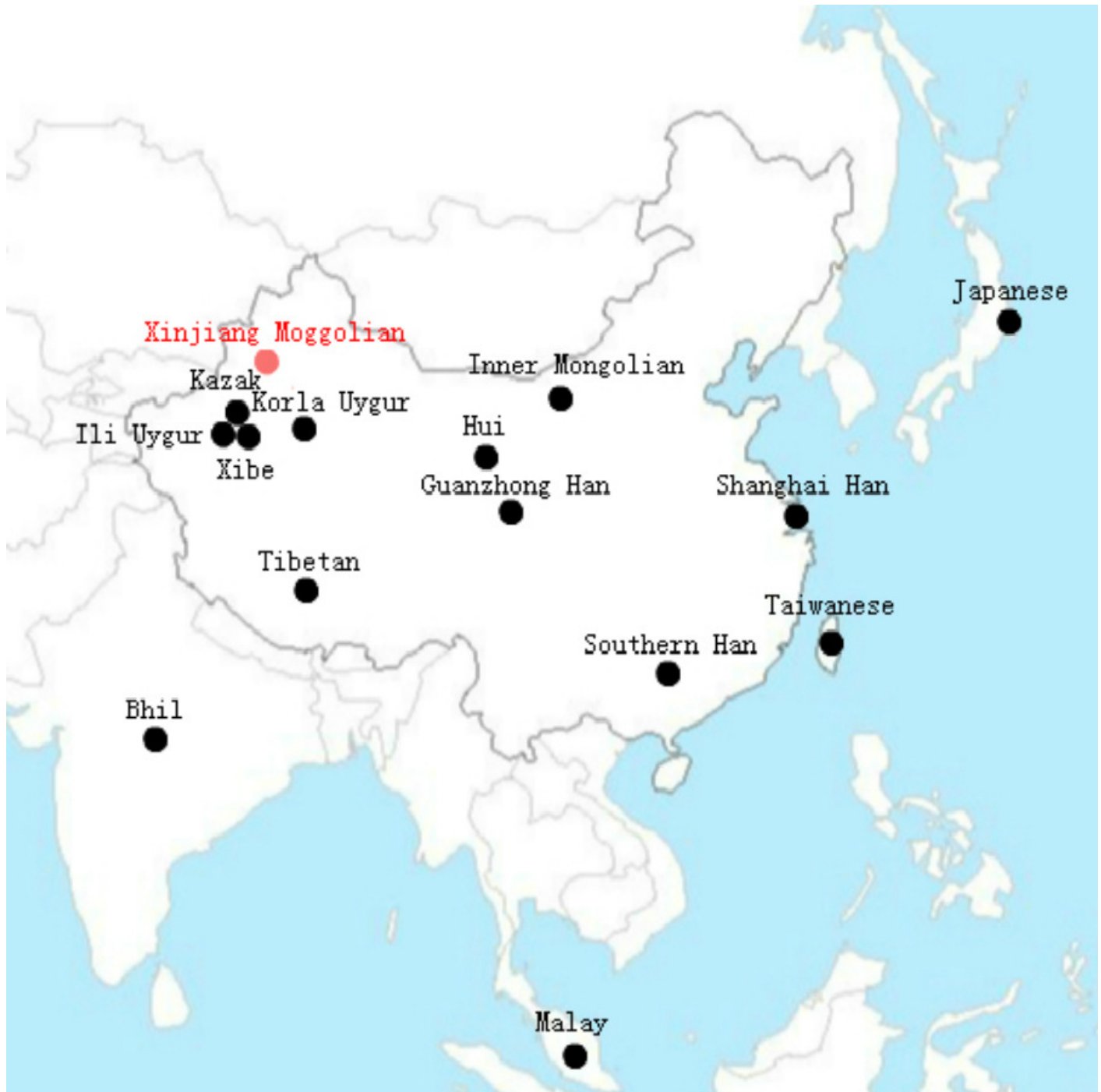
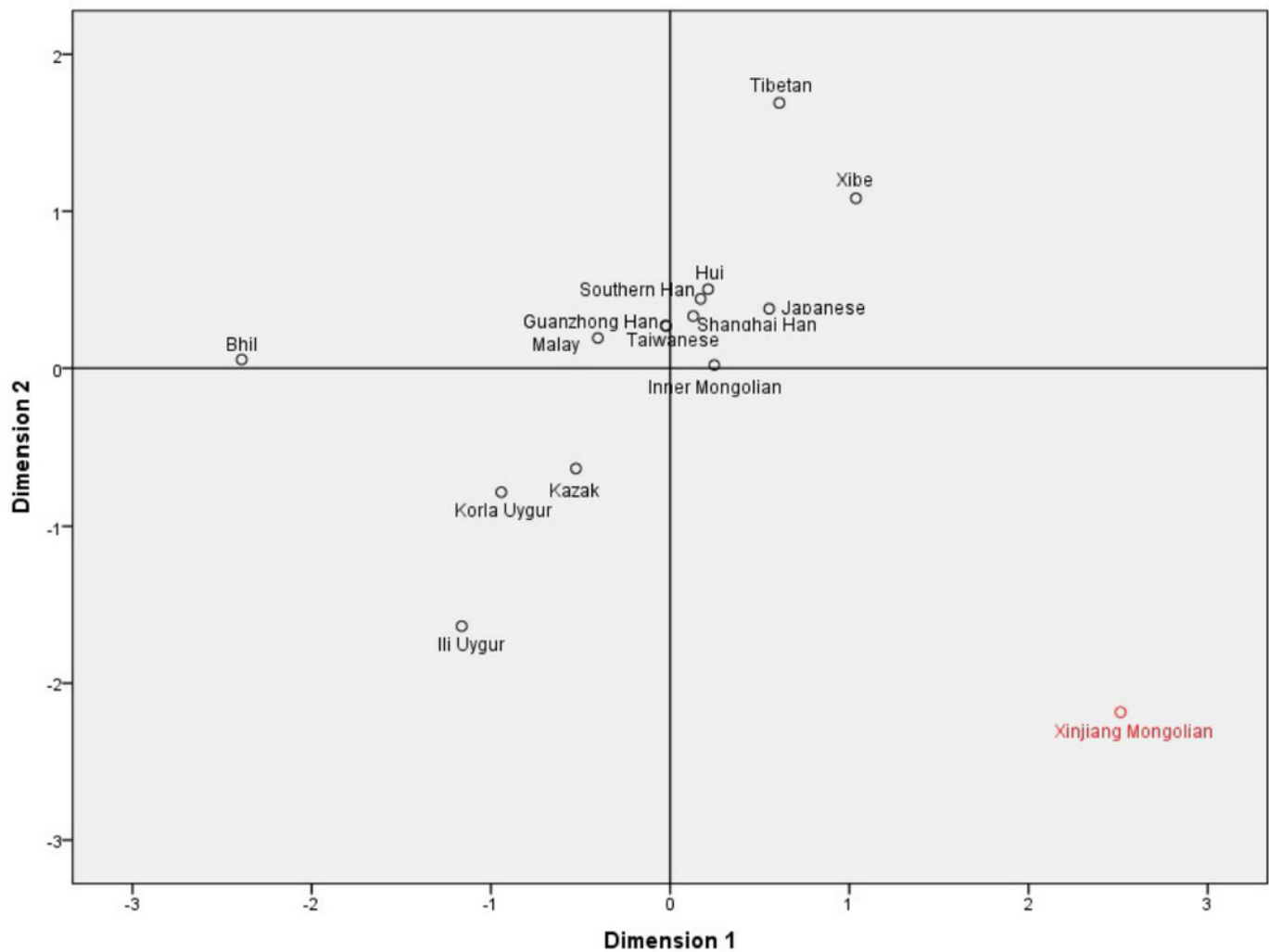
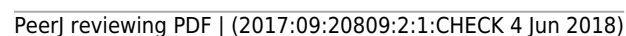



Figure 2

The multidimensional scaling (MDS) plot based on DA distance for the 15 Asian
populations.

Figure 3

The phylogenetic tree generated by the DA distance and the unweighted pair-group method with arithmetic means method (UPGMA).

The forensic efficiency parameters of 19 X-STR loci in Xinjiang Mongolians.

1   **Table 1. The forensic efficiency parameters of 19 X-STR loci in Xinjiang Mongolians.**

Index	PIC	Het	PE	PDF	PDM	MEC Krüger	MEC Kishida	MEC Desmarais Duo
DXS8378	0.6157	0.6687	0.3815	0.8372	0.6687	0.4160	0.6157	0.4687
DXS7423	0.5430	0.6077	0.3003	0.7814	0.6077	0.3442	0.5430	0.3966
DXS10148	0.8980	0.9056	0.8069	0.9835	0.9056	0.8095	0.8980	0.8221
DXS10159	0.7550	0.7882	0.5773	0.9220	0.7882	0.5817	0.7550	0.6254
DXS10134	0.8494	0.8635	0.7217	0.9672	0.8635	0.7296	0.8494	0.7508
DXS7424	0.7185	0.7499	0.5096	0.9061	0.7499	0.5457	0.7185	0.5830
DXS10164	0.5334	0.5915	0.2808	0.7750	0.5915	0.3406	0.5334	0.3867
DXS10162	0.6804	0.7199	0.4598	0.8820	0.7199	0.4958	0.6804	0.5393
DXS7132	0.7047	0.7453	0.5018	0.8945	0.7453	0.5200	0.7047	0.5669
DXS10079	0.7883	0.8130	0.6234	0.9403	0.8130	0.6345	0.7883	0.6681
DXS6789	0.7679	0.7966	0.5927	0.9299	0.7966	0.6050	0.7679	0.6425
DXS101	0.8119	0.8321	0.6600	0.9516	0.8321	0.6707	0.8119	0.6992
DXS10103	0.7559	0.7863	0.5739	0.9240	0.7863	0.5876	0.7558	0.6271
DXS10101	0.8931	0.9015	0.7984	0.9819	0.9015	0.8001	0.8930	0.8141
HPRTB	0.6781	0.7244	0.4669	0.8777	0.7244	0.4844	0.6781	0.5362
DXS6809	0.7799	0.8051	0.6085	0.9368	0.8051	0.6232	0.7799	0.6572
DXS10075	0.6289	0.6845	0.4047	0.8448	0.6845	0.4296	0.6289	0.4839
DXS10074	0.7535	0.7847	0.5710	0.9225	0.7847	0.5845	0.7535	0.6244
DXS10135	0.9184	0.9235	0.8437	0.9890	0.9235	0.8448	0.9183	0.8539

2   PIC: polymorphism information content; Het: Heterozygosity; PE: probability of exclusion; PDF: power of discrimination in females; PDM: power of

3   discrimination in males; MEC Kr€ uger: mean exclusion chance for deficiency cases; MEC Kishida: mean exclusion chance for normal trios; MEC

4   Desmarais Duo: mean exclusion chance in duo cases.