

**Electronic Supplementary Materials**

<https://doi.org/10.1631/jzus.B1800279>

# **Genetic distribution and forensic evaluation of multiplex autosomal short tandem repeats in the Chinese Xinjiang Mongolian group<sup>\*#</sup>**

Yuan-yuan WEI<sup>1,2</sup>, Xiao-ye JIN<sup>1,2,3</sup>, Qiong LAN<sup>4</sup>, Wei CUI<sup>1,2,3</sup>, Chong CHEN<sup>1,2,3</sup>,  
Ting-ting KONG<sup>1,2</sup>, Yu-xin GUO<sup>1,2,3</sup>, Jian-gang CHEN<sup>†‡5</sup>, Bo-feng ZHU<sup>†‡1,2,4</sup>

**Data S1 Materials and methods****Sample collections**

After written informed consent of participants being obtained, bloodstain samples of 1133 unrelated healthy Mongolians (679 men and 454 women) were collected. The studied individuals have lived in Xinjiang Uygur autonomous region for at least three generations and no consanguinity relationships among these individuals existed. The current research was carried out following the human and ethical research principles of Xi'an Jiaotong University Health Science Center, China and authorized by the ethical committee of Xi'an Jiaotong University Health Science Center, China.

**Multiplex PCR and STR typing**

Under the specifications provided by AGCU EX20 kit, multiplex PCR of 19 STR loci and one Amelogenin gene was performed on the GeneAmp PCR system 9700 (Applied Biosystems, Foster City, CA, USA). By the capillary electrophoresis on the ABI 3500xl Genetic Analyzer (Applied Biosystems, Foster City, CA, USA), PCR products were separated and detected. Allele genotypes of 20 loci were determined on the GeneMapper ID Software v3.2.1 (Applied Biosystems, Foster City, CA, USA) by comparing with the allelic ladder of the kit. Control DNA 9947A and sterile water were used as positive and negative control, respectively. Peak detection threshold was set to 50 relative fluorescence units when analyzing the allele peaks.

**Statistical analyses**

Powerstats (version1.2) spreadsheet (Promega, Madison, WI, USA) was utilized to estimate allele frequencies and forensic parameters including polymorphism information content, match probability, discrimination power, probability of exclusion and observed heterozygosity. The *P*-values for Hardy-Weinberg equilibrium tests and expected heterozygosity were conducted by Arlequin software v3.5 (Excoffier and

Lischer, 2010). SHEsis online tool (Shi and He, 2006) was used to assess linkage disequilibrium (LD) of pairwise loci. Pairwise genetic distances ( $D_A$ ) among 21 populations which were estimated by DISPAN software, were visualized by *R* program (<https://www.r-project.org/>). Furthermore, Arlequin software v3.5 was utilized to calculate *Fst* values and their corresponding *P*-values between Mongolian group and other compared populations. On the basis of *Fst* values mentioned above, MDS of 21 populations was plotted by SPSS software v18.0.

## References:

- Excoffier, L. and Lischer, H.E.L., 2010. *Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows*, **10**(3):564-567.  
[doi: 10.1111/j.1755-0998.2010.02847.x]
- Shi, Y.Y. and He, L., 2006. *SHEsis, a powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci (vol 15, pg 97, 2005)*, **16**(10):851-851.  
[doi:10.1038/sj.cr.7310101]