Allele frequencies of 17 autosomal STR loci in the Va ethnic minority from Yunnan Province, Southwest China

Fei Guo 1

Received: 20 April 2017 / Accepted: 30 May 2017 / Published online: 7 June 2017
© Springer-Verlag Berlin Heidelberg 2017

Abstract Allele frequencies and forensic parameters for 17 autosomal short tandem repeat (STR) loci were investigated from 1542 individuals of Va ethnic minority in Yunnan Province in the southwest of China. Also, genetic distances between Va and 13 published populations were compared, and a neighbor-joining (NJ) phylogenetic tree was developed and visualized using averaged \( F_{ST} \) matrix. Results demonstrate that these loci are highly polymorphic in the Va population and can be applied in forensic, population genetic, and linguistic fields.

Keywords Population data · Autosomal short tandem repeat (STR) · PowerPlex 18D System · Va ethnic minority · Yunnan Province

The Va (also called “Wa”) ethnic minority, with a population of 429,709 at the 2010 census, lives in Ximeng, Menglian, and Lancang counties of Pu’er City and Cangyuan, Gengma, Shuangjiang, Zhenkang, and Yongde counties of Lincang City in southwestern Yunnan Province (Fig. S1). Traditionally, this area was called the Ava Hilly Region that is situated between the Lancang and Nu rivers and blocked by Nu mountain ridges. Ximeng and Cangyuan counties are the main places where the Va people live in compact communities. The Va language belongs to the Palaung-Va language group of the Austroasiatic language family, including three dialects—Parauk, Vo, and Awau, and an alphabetic script based on Parauk dialect was created in 1957 [1].

Blood samples of 1542 unrelated healthy individuals (1277 males and 265 females) were collected from Ximeng Va Autonomous County and Cangyuan Va Autonomous County in Yunnan Province, Southwest China, after informed consent. Genomic DNA was extracted from blood dried onto 903 paper using magnetic beads on the TECAN® Freedom EVO Workstation (TECAN, Männedorf, Switzerland). Extracted DNA was amplified with the PowerPlex® 18D System (Promega, WI, USA) in the GeneAmp® PCR System 9700 (Thermo Fisher Scientific, MA, USA) according to the manufacturer’s recommendations. Amplified products were separated and detected on the Applied Biosystems® 3130xl Genetic Analyzer (Thermo Fisher). Raw data was analyzed using the GeneMapper® ID Software v3.2 (Thermo Fisher). Allele frequencies, chi-square tests of the Hardy-Weinberg equilibrium (HWE), and forensic parameters including observed heterozygosity (\( H_o \)), polymorphism information content (PIC), power of discrimination (PD), and power of exclusion (PE) were calculated using the Modified Powerstats [2]. Genetic distances (\( F_{ST} \)) and \( p \) values for population comparisons were calculated by locus using the Arlequin ver 3.5 [3]. A neighbor-joining (NJ) phylogenetic tree was developed and visualized using the MEGA6 [4] based on averaged \( F_{ST} \) matrix.

Allele frequencies and forensic parameters for each locus in a Yunnan Va ethnic minority are listed in Table S1. A total of 199 unique alleles were counted, and the most informative locus was FGA (\( H_o = 0.8515; \) PIC = 0.8500) with 21 unique alleles and the lowest was TPOX (\( H_o = 0.5467; \) PIC = 0.5051) with 6 alleles. No statistically significant departure from the HWE was observed at any locus after the Bonferroni’s correction (\( \alpha = 0.0029 \)). The PD ranged from 0.9672 (FGA) to

Electronic supplementary material The online version of this article (doi:10.1007/s00414-017-1620-8) contains supplementary material, which is available to authorized users.

Fei Guo
nevergfaye@gmail.com

1 Department of Forensic Medicine, Criminal Investigation Police University of China, No. 83, Tawan Street, Huanggu District, Shenyang, Liaoning 110854, China
0.7534 (TPOX) with >0.999999999 for combined PD. The PE ranged from 0.7262 (D2S1338) to 0.2317 (TPOX) with 0.999972288 for combined PE. Table S2 shows pairwise $F_{ST}$ and $p$ values between Va and 10 populations in Yunnan Province [5–12] as well as 3 populations in neighboring countries [13–15] across 15 loci, where statistically significances after adjustment for multiple testing ($p < 0.0033$) differed from Miao [9] at all loci, from Nakhi [11], Han [12] and Japanese [15] at 14 loci, from Bai [5], Derung [7], Yi [11] and South Korean [14] at 13 loci, from Hani [8] and Nu [7] at 12 loci, from Tibetan [7] at 9 loci, from Dai [6] at 8 loci, and from Vietnamese [13] at 1 locus. Figure S2 indicates the clusters in a NJ tree based on the averaged $F_{ST}$ matrix by 15 loci between 14 populations (Table S3). In the figure, Yunnan Va clustered together with Vietnamese, which may result from a great opportunity for gene exchange within two populations or a disproportional size of two datasets. In addition, linguistics-related clusters were observed: (1) Palaung-Va (Va) and Kinh (Vietnamese) language groups of Austroasiatic language family; (2) Bai (Bai), Bodish (Tibetan), Lolo–Burmese–Nakhi (Yi, Hani and Nakhi) and Jingpho–Nungish–Luish (Derung and Nu) language subgroups from Tibeto-Burman language group and Sinitic language group (Han) of Sino-Tibetan language family; (3) Tai–Kadai language family (Dai); (4) Hmong–Mien language family (Miao). Results are generally in accordance with the biogeographical and linguistic distribution of the studied populations. In conclusion, the 17 autosomal STR loci provide highly polymorphic information in the Va ethnic minority for forensic individual identification and paternity testing as well as for population genetic and linguistic studies.

References
