Heliyon 9 (2023) e21384

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Insertion/deletion polymorphism for genetic background and forensic performance exploration of the Sui group from Guizhou

Xiaoxue Wang, Hongling Zhang, Qiyan Wang, Meiqing Yang, Yubo Liu, Qianchong Ran, Chunli Huang, Jiang Huang *, Zheng Ren **, Xiaoye Jin ***

Department of Forensic Medicine, Guizhou Medical University, Guiyang, 550004, Guizhou, China

ARTICLE INFO

Keywords:
Forensic genetics
InDels
STRs
Population genetic analyses
Guizhou Sui

ABSTRACT

Insertion/deletion polymorphisms (InDels) as ideal genetic markers for forensic genetics are appreciated by scholars both nationally and internationally because they integrated the favorable features of single nucleotide polymorphisms (SNPs) and short tandem repeats (STRs). Nevertheless, with the limited identification efficiency of InDels, the multiplex amplification systems of InDels might just be applied as the supplementary methods in paternity testing with respect to commonly used STRs.

In the current research, we successfully genotyped 105 unrelated individuals from the Guizhou Sui population based on a six-color fluorescence multiplex panel that could simultaneously detect 64 genetic markers (59 autosomal InDels, two autosomal miniSTRs and three Y chromosomal genetic markers). In addition, frequency distributions and forensic statistical parameters of these loci in the Sui group were assessed using the STRAF software. Phylogenetic relationships among the Sui group and other reference populations were dissected by two methods (principal component analysis and phylogenetic trees) based on 59 InDels. The combined discrimination power and probability of exclusion values of 61 autosomal genetic markers in the Sui group were nearly equal to 1–1.90063 $\times$ 10^{-27} and 0.999998272, respectively. Furthermore, we observed that the Sui group from Guizhou had closer genetic affinities with East Asian populations with respect to other continental populations.

In summary, we stated that the multiplex amplification system might be utilized as a prospective independent tool for human individual identification and parentage testing in the Sui group residing in Guizhou.

1. Introduction

Insertion/deletion polymorphisms (InDels) are diallelic markers generated through one mutation event [1]. InDels stand out among genetic markers as desirable genetic markers in forensic genetics and are preferred by academics both domestically and globally because they partially integrate the beneficial traits of STRs and SNPs. On the one hand, the amplicon of InDels is short and they display low mutation rate and high genetic stability. On the other hand, InDels are length polymorphisms, so they can be typed quickly and

* Corresponding author.
** Corresponding author.
*** Corresponding author.
E-mail addresses: mmm_hj@126.com (J. Huang), mmm_rz@163.com (Z. Ren), 1115259825@qq.com (X. Jin).

https://doi.org/10.1016/j.heliyon.2023.e21384
Received 11 May 2023; Received in revised form 8 September 2023; Accepted 20 October 2023
Available online 29 October 2023

accurately with capillary electrophoresis (CE) technology at low prices [2–4]. Researchers in forensic science have developed a series of multiplex amplification InDel panels. A novel InDel typing system better suited for East Asian populations, was developed by Fan et al. to increase forensic application efficiency of InDels, which was a reliable tool for intercontinental population differentiation and human identification [5]. Huang et al. exploited a novel 32-plex InDel panel for personal identification [6]. Tao et al. developed the SifaInDel 45-plex panel for complex kinship analysis [7]. Working on the CE platform, Li et al. designed and created a multiplex amplification panel of 18 autosomal multi-InDels for the forensic investigation of highly degraded specimen [8]. Besides, a new panel of 52 ancestral informative InDels was developed by Zhou et al. for biogeographic ancestry inferences of individuals [9]. Nevertheless, due to the limited identification efficiency of InDels, the multiplex amplification systems of InDels might just be considered as the supplementary methods in paternity testing in relatively to commonly applied STRs. Comfortingly, in the prior investigation by Liu et al., they built a 64-plex panel including two autosomal miniSTRs, 59 autosomal InDels and three Y chromosomal genetic markers by the PCR-CE and six dyes labeled technology. The system is not only suitable for the identification of highly degraded samples and mixed samples, it is also more efficient in the forensic investigation of complex cases [10]. Nevertheless, the forensic performance of the 64-plex panel in various populations should be further evaluated.

The Sui population is a member of the Tai-Kadai language family and officially recognized as an ethnic minority in China. Almost 90 % of the population of Sui resides in Guizhou province, with the remainder spread out through other regions like Guangxi, Yunnan, and Sichuan [11,12]. The characters, medical culture, and costumes of the Sui people have been studied by several researchers [13–15], which has promoted our understanding of them. Nonetheless, the source of the Sui population remains controversial and exploring of the genetic structure for the Sui population is still limited [16,17]. As a result, we used the 64-plex panel previously developed by Liu et al. [10] to perform the population genetic analysis of the Guizhou Sui population. Firstly, forensic statistical parameters of these 64 loci in Sui group living in Guizhou were calculated to provide fundamental information for kinship testing as well as personal identification. Secondly, frequency distribution comparisons of 59 overlapping InDels among the Sui and previously reported populations were conducted to explore genetic divergences of these loci. Finally, phylogenetic relationships of these populations were analyzed deeply to dissect the genetic structure of the studied Sui group.

2. Materials and methods

2.1. Sample collection

In this investigation, there were 105 bloodstain specimens (52 males and 53 females) collected from unrelated healthy Sui participants. The present work was performed in accordance with the Guizhou Medical University Ethics Committee guidelines. We also collected genetic data of 59 InDels in different continental populations from the 1000 Genomes Project [18]. Besides, the previous reported Guizhou Han (CHG), Guizhou Dong (CDG), Tibet Tibetan (CTT), Qinghai Tibetan (CTQ) and Hunan Han (CHH) populations were also considered as reference populations [10,19–21]. The details of the populations were shown in Supplementary Table 1.

2.2. PCR amplification and allele genotyping

Firstly, we obtained 1 cm^2 blood cards from the bloodstain samples. Afterward, the samples were extracted for DNA with the Chelex-100 method and DNA was quantified with Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific, Wilmington, USA). We next conducted the multiplex amplification of 64 loci for each sample using the GeneAmp PCR System 9700 apparatus (Thermo Fisher Scientific, Foster City, CA, United States). The above involved PCR reagents and thermal cycling conditions were reported in Liu et al. [19]. The amplified product was then mixed thoroughly. Later we isolated and detected the mixture by ABI 3500xL Genetic Analyzer (Thermo Fisher Scientific, Foster City, CA, USA). Finally, GeneMapper ID-X software v1.5 (Thermo Fisher Scientific, Foster City, CA, USA) was applied for genotyping.

2.3. Statistical analysis

Firstly, we performed a series of calculations based on STRAF software [22] for the forensic parameters of 59 InDels and two miniSTRs in the Guizhou Sui population, including Hardy–Weinberg equilibrium (HWE) testing, linkage disequilibrium (LD) analysis, allelic frequencies, expected heterozygosity (He), observed heterozygosity (Ho), polymorphic information content (PIC), matching probability (PM), discrimination power (PD), probability of exclusion (PE), and typical paternity index (TPI). Next, we used the see package v0.7.4 of R software v4.2.2 to draw the raincloud plot of the forensic parameters. Besides, we estimated the allelic frequencies and the pairwise fixation index (F_{ST}) of 59 InDels among the Sui group and 31 previously published populations by applying STRAF. Next, on the basis of allelic frequencies of 59 InDels, genetic distances (D_A) of these 32 populations were calculated with the DISPAN program. Furthermore, the phylogenetic tree was carried out using the MEGA program [23] on the basis of D_A values. Finally, we performed the principal component analysis (PCA) of these populations by the factoextra package v1.0.7 and the ggplot2 package v3.4.1 of R software v4.2.2 based on allelic frequencies of 59 InDels.

3. Results

3.1. LD and HWE results of 59 InDels and two miniSTRs in the Sui group residing in Guizhou

The linkage disequilibrium tests (*p*-values) for all pairs of 59 InDels and two miniSTRs in the Sui population from Guizhou province were given in Supplementary Table S2. The smallest *p*-value (0.0001) was observed between rs3030496 and rs5877451 loci. After Bonferroni correction ($p > 0.05/\ 1830 = 0.00002732$) being employed, we observed that *p*-values of these pairwise loci weren't statistically significant. Therefore, we concluded that 61 loci complied with linkage equilibrium in Guizhou Sui population. The *p*-values for Hardy-Weinberg equilibrium were shown in Supplementary Table S3. The results demonstrated that the HWE *p*-values of these 61 loci in the Guizhou Sui population fluctuated from 0.0090 to 1.0000, with just one locus deviating slightly from HWE. Following the Bonferroni correction ($p > 0.05/61 = 0.0008197$), HWE *p*-values of all loci were all within admissible range. As a result, we thought that the 61 loci were in line with HWE in the Sui group.

3.2. Frequency distributions and forensic statistical indexes of 59 InDels and two miniSTRs in the Sui group from Guizhou

Allele frequencies of 59 InDels and two miniSTRs in Guizhou Sui population were listed in Supplementary Table S3. Amount of 136 alleles were noticed at the 61 loci. Not surprisingly, D1S1656 and D3S1358 loci displayed maximum amount of alleles in the Guizhou Sui population, with allele counts reaching 11 and 7, respectively. With respect to insertion allele of 59 InDels, their frequencies varied from 0.3000 to 0.6905. For two miniSTRs, their allelic frequencies varied from 0.0048 to 0.3333, with the median value of 0.0786.

Forensic parameters for 59 InDels and two miniSTRs of the Guizhou Sui population were demonstrated in Fig. 1 and Supplementary Table S3. As to two miniSTRs, values of the He, PIC, Ho, PM, PD, PE, and TPI were 0.8506, 0.8289, 0.8571, 0.0483, 0.9517, 0.7090, and 3.5000 for the D1S1656 locus, respectively; for the other locus D3S1358, they were 0.7293, 0.6759, 0.7714, 0.1322, 0.8678, 0.5471, and 2.1875, respectively. As to 59 InDels, values of the He, PIC, Ho, PM, PD, PE and TPI ranged from 0.4220 (rs60922184) to 0.5024 (rs67365630), 0.3318 (rs60922184) to 0.3750 (rs67365630, rs66739142, rs2308292, and rs10546179), 0.3238 (rs10535391) to 0.5714 (rs3061475), 0.3448 (rs111626822) to 0.4500 (rs10581929), 0.5500 (rs10581929) to 0.6552 (rs111626822), 0.0738 (rs10535391) to 0.2580 (rs3061475), and 0.7394 (rs10535391) to 1.1667 (rs3061475), with the median values of 0.4935, 0.3705, 0.4857, 0.3798, 0.6202, 0.1753 and 0.9722, respectively. At the same time, values of the combined PM, PD and PE of 59 InDels were 2.9749×10^{-25}, $1–2.9749 \times 10^{-25}$ and 0.999986889, respectively. After two miniSTRs were adding into these 59 InDels, the values of CPM, CPD and CPE of 61 loci were 1.90063×10^{-27}, $1–1.90063 \times 10^{-27}$ and 0.999998272, respectively. Furthermore, we evaluated haplotype frequencies of two Y-InDel (rs759551978, and rs199815934), as presented in Supplementary Table S4. Two haplotypes were identified in the Guizhou Sui population with frequencies varying from 0.3077 to 0.6923.

3.3. Allelic frequency distribution of 59 InDels among different populations

Comparison of allelic frequency distributions of 59 InDels among the studied Sui and 31 previously reported populations was

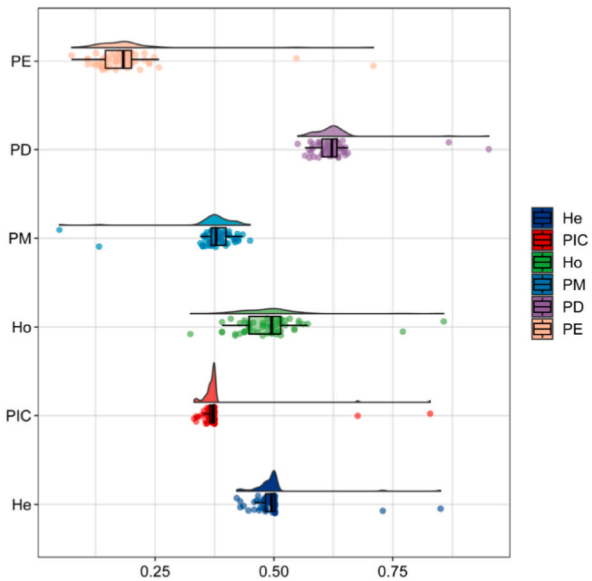

Fig. 1. The raincloud plot of forensic parameters for the 59 InDels and two STRs in the Guizhou Sui population. PIC, polymorphic information content; He, expected heterozygosity; Ho, observed heterozygosity; PM, matching probability; PD, discrimination power; PE, probability of exclusion.

conducted, as shown in Fig. 2. It is possible to interpret the color transition from green to white and then to orange as representing the range from the lowest to the highest allelic frequencies. Loci which exhibited pronounced allelic distribution differences among these populations could be easily discerned from Fig. 2. For instance, rs66739142, rs71698233, rs3083268, rs10579944, rs1611025 and rs144378883 loci showed comparatively high insertion allele frequencies in African populations. Oppositely, rs6481 and rs111626822 loci displayed relatively low insertion allelic frequencies in African populations, so did rs35173752 and rs5833522 loci in South Asian and European populations. Besides, we observed that rs35453727 and rs71698233 loci indicated higher insertion allelic frequencies in the other continental populations than those in East Asian populations. The above-mentioned loci could be ideal markers of ancestral information to differentiate these continental populations.

Meanwhile, we also noticed that these 59 InDels could be divided into four clusters on the left of the heatmap. In addition, these 32 populations were also classified into five branches on the top of the heatmap, which were basically consistent with their continental origins.

3.4. Genetic divergence evaluation between the Sui from Guizhou and other populations using 59 InDels

Genetic differentiations among Guizhou Sui and other reference populations were explored by F_{ST} and D_A genetic distances, as shown in Supplementary Tables S5–6 and Fig. 3. A heatmap based on D_A values could be spotted in the top right corner of Fig. 3. The transition in color from pink to pale yellow indicated the progression of D_A values from small to large. Similarly, the heatmap of F_{ST} values were shown in the lower-left corner of Fig. 3. The color shifting from dark blue to light blue symbolized the progression of F_{ST} values from small to large. Meanwhile, the size of the circles represented the genetic distances among these 32 populations: the larger the circle was, the larger genetic distances of paired populations. In general, it was obvious that the genetic distances between Guizhou Sui population and East Asian populations were relatively close with small circles. Whereas, compared with populations in Africa, the

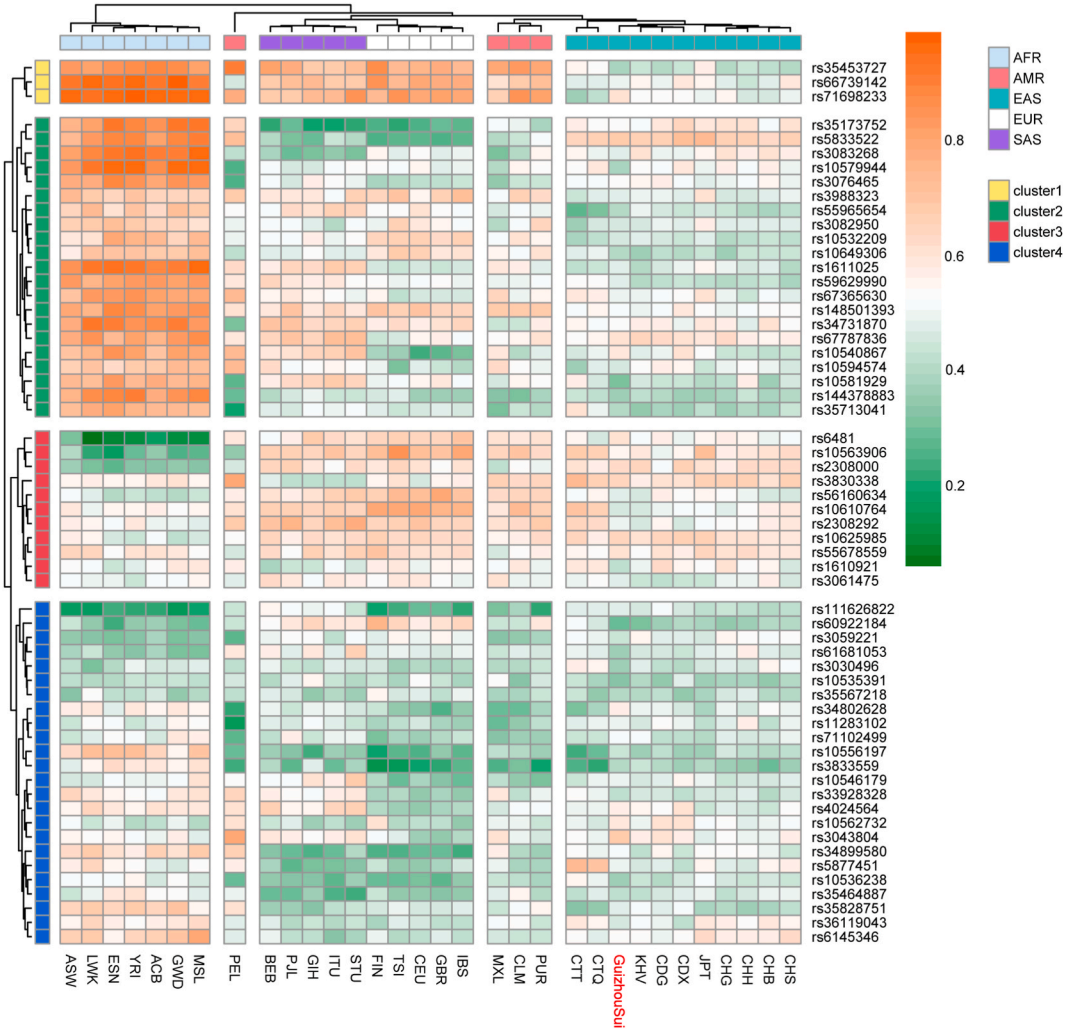

Fig. 2. The heatmap of the insertion allelic frequencies among the Guizhou Sui and other reference populations.

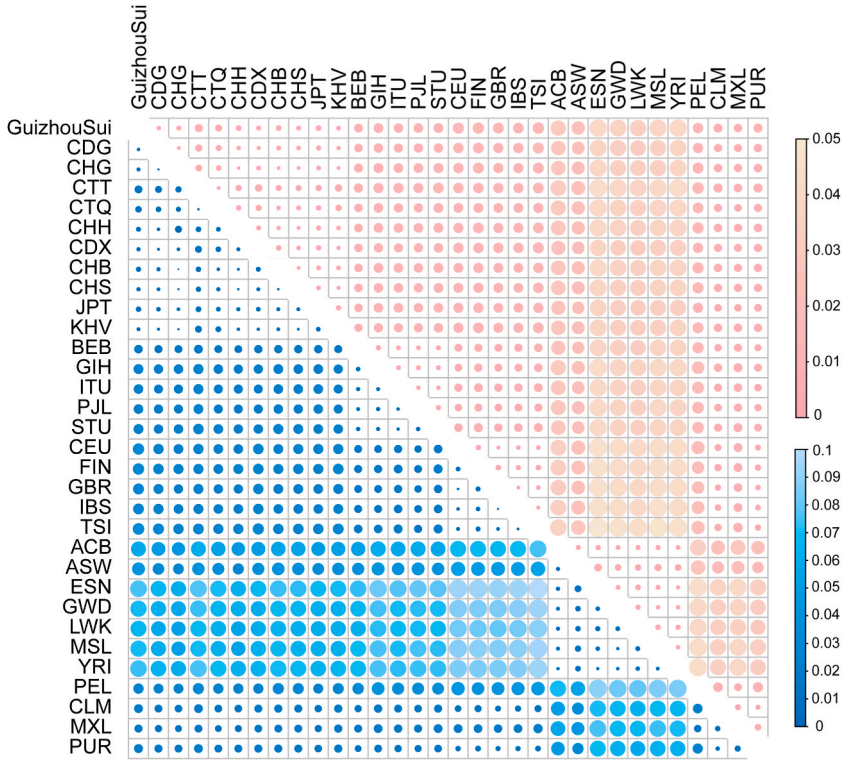

Fig. 3. Heatmaps of D_A (the upper right part) and F_{ST} (the lower left part) among the Guizhou Sui population and other 31 comparison populations.

Guizhou Sui population was discovered to be large circles representing the further genetic distances. On the one hand, the Sui population in Guizhou shared close genetic affinities with East Asian populations, particularly with the CDG, which has the smallest D_A (0.0018) and F_{ST} (0.0035) values. On the other hand, the furthest genetic affinities were found between the Sui population in Guizhou and the African populations, especially with the Esan population in Nigeria, the D_A and F_{ST} values were 0.0440 and 0.0753, respectively.

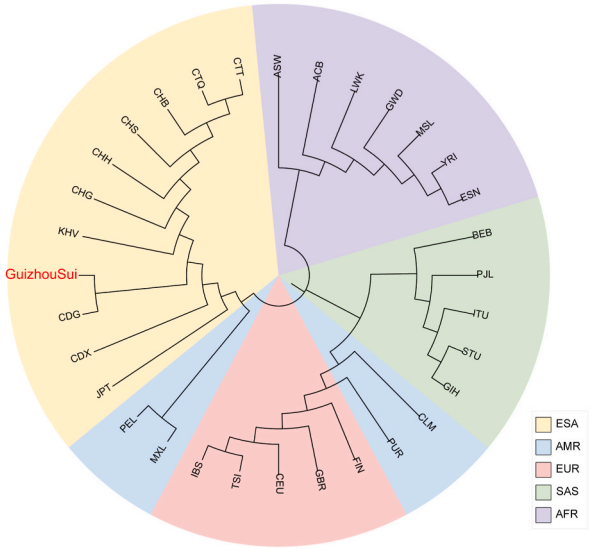

Fig. 4. Phylogenetic tree reconstructed of Guizhou Sui and other reference populations based on the neighbor-joining algorithm.

3.5. Neighbor-joining phylogenetic tree

The phylogenetic tree of the studied Sui and other comparsion populations was reconstructed according to their paired D_A values, which were presented in Fig. 4. Five diverse colors stood for the five continents spreading throughout different geographic regions. According to the analyses of the phylogenetic tree, it was clear that the Guizhou Sui population clustered with the East Asian populations. Additionally, the Guizhou Sui population and CDG were classified on the same branch, signifying that the Guizhou Sui population and CDG got closer phylogenetic relationships than the other nine populations in East Asia.

3.6. PCA of the Sui group and other comparison populations

On the basis of allele frequencies of 59 InDels in the Sui group from Guizhou and 31 previously published populations, the PCA of these populations was constructed, as displayed in Fig. 5. The first three principal components we chose could explain 58.34 %, 18.09 % and 6.36 % variation among these populations. Moreover, the different colors and shapes in Fig. 5 represented continental origins of populations.

For PC1, we noticed that the 32 populations can be dispersed into two parts, including the populations in Africa on the right side and the non-African populations on the left side. For PC2, these populations could be further classified into three parts: the European populations, the East Asian populations and other reference populations. Due to loose aggregation degree than other continents, the PUR and CLM populations in the America were dispersed into the South Asian populations. For PC3, the South Asian and the European populations could be distinguished from other intercontinental populations. Taking these three principal components into account, the Sui population in Guizhou generally clustered with other East Asian populations, suggesting that Sui population shared close genetic relationships with these East Asian populations.

4. Discussion

Based on the 64-plex panel, several academics have evaluated genetic distributions and forensic application efficiency of the panel in some populations to date [19–21]. Nevertheless, the forensic performance and the genetic structure of the Guizhou Sui population have not been explored utilizing the 64-plex panel. Therefore, this study focused on the values of the forensic applications of the panel in the Guizhou Sui population. Furthermore, genetic structure of the Guizhou Sui population was also dissected by multiple methods in comparison with other reference populations.

Prior studies indicated that the SNPs which could be used for personal identification had an average heterozygosity ≥0.4 [24]. As similar to SNPs, InDels usually displayed di-allelic mutations. In our study, the average heterozygosity of the 59 InDels was more than 0.4 in the Guizhou Sui population, revealing that these loci might be used in personal identification. A previous research by Botstein et al. stated that genetic markers were reasonably informative when their PIC values were more than 0.25 [25]. For this current work, we discovered that the PIC values of 59 InDels were higher than 0.3 in the Guizhou Sui population, signifying that these InDels could offer reasonable genetic information for the Guizhou Sui population. Next, we compared the forensic efficiency of the various kits, as displayed in Supplementary Table S7. It was discovered that the 64-plex panel could offer lower CPM and higher CPD and CPE values than other previously developed InDels [26–28], emphasizing the value of the application in forensic parentage testing and personal

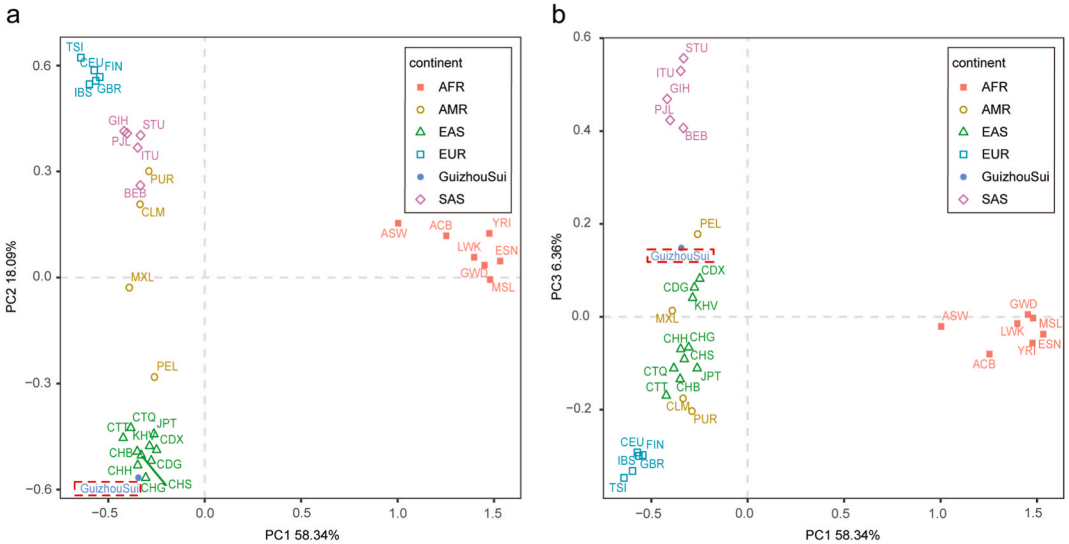

Fig. 5. Principal component analysis of Guizhou Sui and other reference populations. **a** PCA results based on PC1 and PC2; **b** PCA results based on PC1 and PC3.

identification. In addition, we observed that forensic efficiency of the 64-plex panel was also greater than some STR panels [29–31], indicating that the 64-plex panel might be viewed as an independent and valuable tool to conduct forensic research in the Guizhou Sui population. What is more, the physical distances between the majority of InDels from the 64-plex panel and the commonly used STRs on the same chromosome were greater than 10 MB (Supplementary Table S8), signifying that these InDels might be employed for forensic research in combination with commonly employed STRs. Even so, we should evaluate linkage disequilibrium of these InDels and STRs before them being employed for forensic research.

For genetic markers, previous studies have indicated that loci with significant frequency variations in different populations could be used as potential ancestral informative markers for population origin inferences [32–34]. In our study, loci rs35453727 and rs71698233 were found with relatively low allelic frequencies in the East Asian populations. Also, the allelic frequency of locus rs35173752 was relatively lower in the South Asian and European populations than those in other populations, which was consistent with a prior research by Zheng et al. [21]. In contrast, loci rs10579944, rs144378883, rs1611025 and rs3083268 shared high allelic frequencies in African populations, which were consistent with the findings previously reported by Wan et al. [20]. Overall, these loci might be utilized as potential ancestral informative markers in the identification of the intercontinental populations.

Through the genetic analysis of the Sui population in Guizhou, we observed that the Guizhou Sui population possessed closer genetic affinities with East Asian populations than other reference populations, which was akin to the preceding research by Yang et al. [26]. There was no doubt that the Guizhou Sui population and several nearby populations shared close genetic affinities, such as CDG, Kinh in Ho Chi Minh City, Vietnam (KHV), CHG, and Chinese Dai in Xishuangbanna (CDX). However, there existed relatively large genetic divergences between the Guizhou Sui population and Han Chinese in Beijing (CHB), CTQ, and CTT populations. The exploration of genetic relationships among the Guizhou Sui population with the CHG and Tibetan populations was already conducted in the previous research [12], and obtained results were in agreement with the results in the present research. The reasons for these results might be due to the long-time isolated geographical location and long-term endogamy of the Guizhou Sui population [35]. In addition, we noticed that the Guizhou Sui population exhibited closer genetic affinity with the CDG than other East Asian populations. Possible reason was that the Sui and Dong people in Guizhou spoke the same Tai-Kadai language, which led to more communication. Nonetheless, there were some limitations in this study. On the one hand, the sample size of the Guizhou Sui population in this study was relatively small. In the future, it is necessary to obtain more Sui individuals from various regions and further assess their genetic relationships via the 64-plex panel. On the other hand, population genetic analyses in the Guizhou Sui population and other reference populations were performed only by 59 InDels. These InDels were mainly used to forensic personal identification and paternity test. Accordingly, it's essential for us to comprehend genetic structure of the Guizhou Sui population with ancestral informative markers and genome-wide data.

5. Conclusion

In the present work, the forensic application efficieiny and genetic structure of the Sui group from Guizhou were investigated utilizing a 64-plex panel. The findings suggested that the panel might be applied as a prospective independent tool for personal identification and paternity test in the Guizhou Sui population. Following population genetic analysis, it was obvious that the Guizhou Sui population shared closer genetic affinities with the East Asian populations than the other intercontinental reference populations, especially with the CDG.

Ethics statement

It was undertaken within the relevant guidelines and further authorized by the Ethics Committee of Guizhou Medical University (Approval number: 2021-224). Our research bloodstain specimens were obtained from 105 unrelated healthy participants who signed informed consents.

Data availability statement

Data will be made available on request.

Funding

This study was supported by the Guizhou Provincial Science and Technology Projects (ZK[2022] General 355); National Natural Science Foundation of China (No. 82160324); Guizhou "Hundred" High-level Innovative Talent Project, Qian Science Platform Talents: [2020]6012-2; and Guizhou Innovation training program for college students (S202210660028, S202210660030, S202010660039).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e21384.

References

[1] A. Zidkova, A. Horinek, V. Kebrdlova, M. Korabecna, Application of the new insertion-deletion polymorphism kit for forensic identification and parentage testing on the Czech population, Int. J. Leg. Med. 127 (2013) 7–10, https://doi.org/10.1007/s00414-011-0649-3.

[2] M. Fondevila, C. Phillips, C. Santos, R. Pereira, L. Gusmão, A. Carracedo, J.M. Butler, M.V. Lareu, P.M. Vallone, Forensic performance of two insertion-deletion marker assays, Int. J. Leg. Med. 126 (2012) 725–737, https://doi.org/10.1007/s00414-012-0721-7.

[3] K. Sun, S.-H. Zhang, R.-X. Zhu, S.-M. Zhao, C.-T. Li, [Progress in InDel as a new generation of genetic marker], Fa Yi Xue Za Zhi 29 (2013) 134–139, 143.

[4] R. Pereira, C. Phillips, C. Alves, A. Amorim, A. Carracedo, L. Gusmão, A new multiplex for human identification using insertion/deletion polymorphisms, Electrophoresis 30 (2009) 3682–3690, https://doi.org/10.1002/elps.200900274.

[5] H. Fan, Y. He, S. Li, Q. Xie, F. Wang, Z. Du, Y. Fang, P. Qiu, B. Zhu, Systematic evaluation of a novel 6-dye direct and multiplex PCR-CE-based InDel typing system for forensic purposes, Front. Genet. 12 (2021), 744645, https://doi.org/10.3389/fgene.2021.744645.

[6] Y. Huang, C. Liu, C. Xiao, X. Chen, S. Yi, D. Huang, Development of a new 32-plex InDels panel for forensic purpose, Forensic Sci Int Genet 44 (2020), 102171, https://doi.org/10.1016/j.fsigen.2019.102171.

[7] R. Tao, J. Zhang, X. Sheng, J. Zhang, Z. Yang, C. Chen, Y. Bian, X. Liu, S. Zhang, C. Li, Development and validation of a multiplex insertion/deletion marker panel, SifaInDel 45plex system, Forensic Sci Int Genet 41 (2019) 128–136, https://doi.org/10.1016/j.fsigen.2019.04.008.

[8] J. Li, L. Lin, B. Jiang, C. Wang, M.M.J. Zeye, D. Wen, W. He, W. Qu, Y. Liu, L. Zha, An 18 Multi-InDels panel for analysis of highly degraded forensic biological samples, Electrophoresis 42 (2021) 1143–1152, https://doi.org/10.1002/elps.202000245.

[9] Y. Zhou, X. Jin, B. Wu, B. Zhu, Development and performance evaluation of a novel ancestry informative DIP panel for continental origin inference, Front. Genet. 12 (2021), 801275, https://doi.org/10.3389/fgene.2021.801275.

[10] Y. Liu, S. Mei, X. Jin, M. Zhao, B. Zhu, Independent development and validation of a novel six-color fluorescence multiplex panel including 61 diallelic DIPs and 2 miniSTRs for forensic degradation sample, Electrophoresis 43 (2022) 1423–1437, https://doi.org/10.1002/elps.202100225.

[11] X. Bin, R. Wang, Y. Huang, R. Wei, K. Zhu, X. Yang, H. Ma, G. He, J. Guo, J. Zhao, M. Yang, J. Chen, X. Zhang, L. Tao, Y. Liu, X. Huang, C.-C. Wang, Genomic insight into the population structure and admixture history of tai-kadai-speaking Sui people in southwest China, Front. Genet. 12 (2021), 735084, https://doi.org/10.3389/fgene.2021.735084.

[12] J. Guo, J. Ji, G. He, Z. Ren, H. Zhang, Q. Wang, M. Yang, Y. Nabijiang, Z. Zhang, J. Zhang, J. Huang, C.-C. Wang, Genetic structure and forensic characterisation of 19 X-chromosomal STR loci in Guizhou Sui population, Ann. Hum. Biol. 46 (2019) 246–253, https://doi.org/10.1080/03014460.2019.1623911.

[13] J. Liu, New achievements in the study of shui characters: comments on shui character bank, Journal of the Qiannan Normal College for Nationaliti 42 (2022) 9–13.

[14] G. Zhu, On the "Origins and Development" of the Medical Culture of the Shui People in Guizhou, Journal of Guizhou Minzu University (Philosophy and Social Sciences), 2015, pp. 20–24.

[15] Q. Li, Y. Deng, M. Wang, A study on the regional characteristics and causes of shui ethnic minority costumes in Guizhou, ZHUANGSHI (2022) 100–103, https://doi.org/10.16272/j.cnki.cn11-1392/j.2022.02.029.

[16] L. Gao, On the origin of the Shui ethnic group, J. Guangzhou Univ. 3 (2002) 11–15.

[17] X. Zhang, J. Zhang, A new theory of Shui's origin, Journal of Guangxi Normal University for Nationalitie 35 (2018) 44–48, https://doi.org/10.19488/j.cnki.45-1378/g4.2018.01.010.

[18] 1000 Genomes Project Consortium A. Auton, L.D. Brooks, R.M. Durbin, E.P. Garrison, H.M. Kang, J.O. Korbel, J.L. Marchini, S. McCarthy, G.A. McVean, G. R. Abecasis, A global reference for human genetic variation, Nature 526 (2015) 68–74, https://doi.org/10.1038/nature15393.

[19] Y. Liu, W. Cui, X. Jin, K. Wang, S. Mei, X. Zheng, B. Zhu, Forensic efficiency estimation of a homemade six-color fluorescence multiplex panel and in-depth anatomy of the population genetic architecture in two Tibetan groups, Front. Genet. 13 (2022), 880346, https://doi.org/10.3389/fgene.2022.880346.

[20] W. Wan, Z. Ren, H. Zhang, Q. Wang, T. Wang, Y. Yang, J. You, K. He, J. Huang, X. Jin, Insight into forensic efficiency and genetic structure of the Guizhou Dong group via a 64-plex panel, Frontiers in Ecology and Evolution 10 (2023). https://www.frontiersin.org/articles/10.3389/fevo.2022.988504. (Accessed 10 July 2023).

[21] Y. Zheng, T. Wang, K. He, Y. Yang, J. You, X. Huang, H. Zhang, Z. Ren, Q. Wang, J. Huang, X. Jin, Forensic efficiency evaluation of a novel multiplex panel of InDels and STRs in the Guizhou Han population and its phylogenetic relationships with other reference populations, Ann. Hum. Biol. 50 (2023) 42–47, https://doi.org/10.1080/03014460.2023.2168754.

[22] A. Gouy, M. Zieger, STRAF-A convenient online tool for STR data evaluation in forensic genetics, Forensic Sci Int Genet 30 (2017) 148–151, https://doi.org/10.1016/j.fsigen.2017.07.007.

[23] S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, X. Mega, Molecular evolutionary genetics analysis across computing platforms, Mol. Biol. Evol. 35 (2018) 1547–1549, https://doi.org/10.1093/molbev/msy096.

[24] A.J. Pakstis, W.C. Speed, R. Fang, F.C.L. Hyland, M.R. Furtado, J.R. Kidd, K.K. Kidd, SNPs for a universal individual identification panel, Hum. Genet. 127 (2010) 315–324, https://doi.org/10.1007/s00439-009-0771-1.

[25] D. Botstein, R.L. White, M. Skolnick, R.W. Davis, Construction of a genetic linkage map in man using restriction fragment length polymorphisms, Am. J. Hum. Genet. 32 (1980) 314–331.

[26] M. Yang, X. Yang, Z. Ren, G. He, H. Zhang, Q. Wang, Y. Liu, H. Zhang, J. Ji, J. Chen, J. Guo, J. Huang, C.-C. Wang, Genetic admixture history and forensic characteristics of Guizhou Sui people inferred from autosomal insertion/deletion and genome-wide single-nucleotide polymorphisms, Frontiers in Ecology and Evolution 10 (2022). https://www.frontiersin.org/articles/10.3389/fevo.2022.844761. (Accessed 10 July 2023).

[27] H. G, R. Z, G. J, Z. F, Z. X, Z. H, W. Q, J. J, Y. M, Z. Z, Z. J, N. Y, H.J. W. Cc, Population genetics, diversity and forensic characteristics of Tai-Kadai-speaking Bouyei revealed by insertion/deletions markers, Mol. Genet. Genom. : MGG 294 (2019), https://doi.org/10.1007/s00438-019-01584-6.

[28] Y. Liu, H. Zhang, G. He, Z. Ren, H. Zhang, Q. Wang, J. Ji, M. Yang, J. Guo, X. Yang, J. Sun, J. Ba, D. Peng, R. Hu, L.-H. Wei, C.-C. Wang, J. Huang, Forensic features and population genetic structure of Dong, yi, han, and chuanqing human populations in southwest China inferred from insertion/deletion markers, Front. Genet. 11 (2020) 360, https://doi.org/10.3389/fgene.2020.00360.

[29] L. Yang, Y. Zhao, C. Liu, D.W.T. Chan, M. Chan, M. He, Allele frequencies of 15 STRs in five ethnic groups (han, gelao, jing, shui and Zhuang) in South China, Forensic Sci Int Genet 7 (2013) e9–e14, https://doi.org/10.1016/j.fsigen.2012.10.009.

[30] L. Zhang, Population data for 15 autosomal STR loci in the Bouyei ethnic minority from Guizhou Province, Southwest China, Forensic Sci Int Genet 17 (2015) 108–109, https://doi.org/10.1016/j.fsigen.2015.04.006.

[31] L. Zhang, Population data for 15 autosomal STR loci in the Dong ethnic minority from Guizhou Province, Southwest China, Forensic Sci Int Genet 16 (2015) 237–238, https://doi.org/10.1016/j.fsigen.2015.02.005.

[32] B. Brinkmann, A. Junge, E. Meyer, P. Wiegand, Population genetic diversity in relation to microsatellite heterogeneity, Hum. Mutat. 11 (1998) 135–144, https://doi.org/10.1002/(SICI)1098-1004 (1998)11:2<135::AID-HUMU6>3.0.CO;2-I.

[33] E. Alladio, C. Della Rocca, F. Barni, J.-M. Dugoujon, P. Garofano, O. Semino, A. Berti, A. Novelletto, M. Vincenti, F. Cruciani, A multivariate statistical approach for the estimation of the ethnic origin of unknown genetic profiles in forensic genetics, Forensic Sci Int Genet 45 (2020), 102209, https://doi.org/10.1016/j.fsigen.2019.102209.

[34] T. Tvedebrink, Review of the forensic applicability of biostatistical methods for inferring ancestry from autosomal genetic markers, Genes 13 (2022) 141, https://doi.org/10.3390/genes13010141.

[35] J. Ji, Z. Ren, H. Zhang, Q. Wang, J. Wang, Z. Kong, C. Xu, M. Tian, J. Huang, Genetic profile of 23 Y chromosomal STR loci in Guizhou Shui population, southwest China, Forensic Sci Int Genet 28 (2017), https://doi.org/10.1016/j.fsigen.2017.01.010 e16–e17.