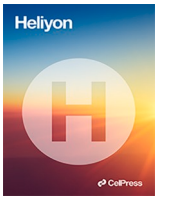
Heliyon 10 (2024) e38501

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Forensic characteristics of 38 Y-STR loci in the Hui population from Shaanxi Province, Northwest China

Li Zhang [a,b,1], Fei Zhan [c,1], Wenfang Zhang [d,1], Sanping Song [c], Shisheng Zhu [e,*]

[a] School of Basic Medical Sciences, Chongqing University of Chinese Medicine, Chongqing, 402760, China
[b] Chongqing Traditional Chinese Medicine Hospital, Chongqing, 400021, China
[c] Institute of Forensic Science, Public Security Department of Shaanxi, Shaanxi, 710000, China
[d] Chongqing Hongzheng Judicial Appraisal Institute, Chongqing, 400013, China
[e] Faculty of Basic Medical Sciences, Chongqing Medical and Pharmaceutical College, Chongqing, 401331, China

ARTICLE INFO

Keywords:
Y-chromosomal STR
Yfiler™ Platinum
Shaanxi Hui
Haplotype diversity
Genetic structure

ABSTRACT

In this study, we genotyped 939 unrelated healthy individuals to investigate the genetic polymorphisms of 38 Y chromosome short tandem repeats(Y-STRs) loci included in the Yfiler™ Platinum kit in the Hui population of Shaanxi Province. Comprehensive population comparisons were performed using analysis of molecular variance(AMOVA), multidimensional scaling(MDS), and molecular evolutionary genetics analysis(MEGA) to explore population relationships. The allele frequencies of the single-copy loci ranged from 0.001 to 0.940 (DYS645), while those of the multi-copy loci ranged from 0.001 to 0.138 (DYS527). The highest genetic diversity (GD) value was observed for DYS385 (0.9723), and the lowest for DYS645 (0.1138). A total of 849 distinct haplotypes were identified among the 939 Hui individuals, with 808 (95.2 %) being unique. Haplotype diversity (HD), haplotype match probability, and discriminatory power (DC) in the Shaanxi Hui population were 0.9996, 0.0015, and 0.9041, respectively. Interpopulation comparisons based on Rst genetic distance values, phylogenetic trees, and MDS indicated that the Shaanxi Hui population has a close genetic relationship with the Ningxia Hui, Qinghai Hui, Xinjiang Hui, and Yunnan Hui populations. The population genetic stratification largely coincides with geographic distribution and language families. Our study revealed that the 38 Y-STR loci exhibit high genetic polymorphisms in the Hui population from Shaanxi Province.

1. Introduction

The Hui group, the most widely distributed ethnic minority group in China, has a population of approximately 10.5 million, mainly settled in Ningxia(34.4 %), Qinghai(14.7 %), Gansu(4.8 %), Xinjiang(4.4 %), and other regions, according to the sixth National Census (http://www.phbang.cn/population/260194.html). Based on previous reports of ancestry history and origin, the Hui nationality primarily comprises Han individuals who integrated with Middle Eastern Arab and Persian ethnic groups, as well as Mongolian, Uyghur, and other Muslim groups [1–3]. The Hui language belongs to the Indo-European language family [4,5] and the group has its own traditional culture, with most members practising Islam. Shaanxi Province, located in northwestern China (Supplementary Fig. 1) and home to nearly 38 million people, is the cradle of Chinese civilization, nurturing the famous city of Chang'an (now Xi'an), the

* Corresponding author.
E-mail address: 10807@cqmpc.edu.cn (S. Zhu).
[1] These authors contributed equally to this work: Li Zhang, Fei Zhan and Wenfang Zhang.

https://doi.org/10.1016/j.heliyon.2024.e38501
Received 22 March 2024; Received in revised form 24 September 2024; Accepted 25 September 2024
Available online 29 September 2024

origin of the Ancient Silk Road. The Hui are the most populous ethnic minority group in Shaanxi, accounting for approximately 0.4 % of the total population.

Y chromosome short tandem repeats (Y-STRs) are widely used in paternity identification, genealogical investigations, and anthropology. As a specific part of the male Y chromosome, Y-STRs are crucial in cases that cannot be resolved using standard autosomal DNA profiling. Additionally, Y-STR haplotype analysis provides a reliable approach for pedigree investigation, population traceability, and male medical genetics. An enhanced Y-STR detection system with high identification efficiency is indispensable for forensic identification [4,6]. The Yfiler™ Platinum System (Thermo Fisher Scientific, Waltham, MA, USA), which includes 38 Y-STR loci (DYS576, DYS389I, DYS635, DYS389II, DYS627, DYS549, DYS645, DYS460, DYS458, DYS19, YGATAH4, DYS448, DYS391, DYS557, DYS593, DYS522, DYS456, DYS390, DYS438, DYS392, DYS518, DYS444, DYS596, DYS570, DYS437, DYS449, DYS643, DYS393, DYS439, DYS481, DYS533, DYS447, DYS385, DYF387S1, and DYS527) and 3 Y-InDels (rs771783753, rs759551978, is a newly available detection system for forensic practice [7–9]. To evaluate the performance of the Yfiler™ Platinum system in the Shaanxi Hui population, we calculated the allele/haplotype distribution and other genetic parameters of the 38 Y-STR loci in 939 Hui individuals. Furthermore, the genetic relationships between the investigated population and 19 previously reported Chinese populations [2,9–22] (Supplementary Fig. 1 and Table S1) were explored.

2. Material and method

2.1. Sample

A total of 939 male blood samples from unrelated individuals in the Shaanxi Hui population were collected for genotype analysis. Informed consent was obtained for participation in the study. The inclusion criterion for sample collection was at least three generations of local residency. The sample collection steps and experimental procedures were approved by the Ethics Committee of Chongqing Medical University (reference number: IACUC-CQMU-2023-07023) and adhered to the Declaration of Helsinki.

2.2. Data collection

Genomic DNA from the samples was directly amplified without extraction. Coamplification was performed on a ProFlex™ PCR System (Thermo Fisher Scientific) using a Yfiler™ Platinum multiple system containing 10 μL of master mix, 5 μL of primer set, and 10 μL of low TE buffer. The thermal cycling parameters were as follows: a initial incubation step at 95 °C for 1 min; 4 cycles of denaturation at 94 °C for 4 s, annealing at 62 °C for 90 s; 24 cycles of denaturation at 94 °C for 4 s, annealing at 60 °C for 90 s; final extension at 60 °C for 30 min and hold at 4°Cfor 24 h. A 3730xl DNA Analyser (Thermo Fisher Scientific) was used for electrophoresis, and raw data were analysed using GeneMapper ID-X 1.5 software (Thermo Fisher Scientific). Internal standard(GS600_LIZ_(60–580).xml) were referenced to conducted genotyping. All experiments were conducted following the relevant instructions, guidelines, and regulations. The recommendations of the DNA Commission of the International Society of Forensic Genetics (ISFG) were strictly followed for the analysis of Y-STRs.

2.3. Data management and statistical analysis

Allele frequencies were calculated using the online software Genepop (https://genepop.curtin.edu.au/), and haplotype frequencies were counted directly. Genetic parameters, including genetic diversity (GD), haplotype diversity (HD), discriminatory power (DP), and haplotype match probability (HMP), were computed using previously reported formulas [23]. The discrimination capacity (DC) was computed as the ratio of the number of distinct haplotypes to the total sample size. To explore the genetic relationship between the research population and other ethnic groups, Y chromosome haplotype reference database(YHRD) (https://yhrd.org/amova), analysis of molecular variance (AMOVA), and multidimensional scaling (MDS) tools were used to evaluate the Rst genetic distance values between the Shaanxi Hui population and 19 other Chinese populations based on the 27 shared Y-STR loci [22]. Then, a

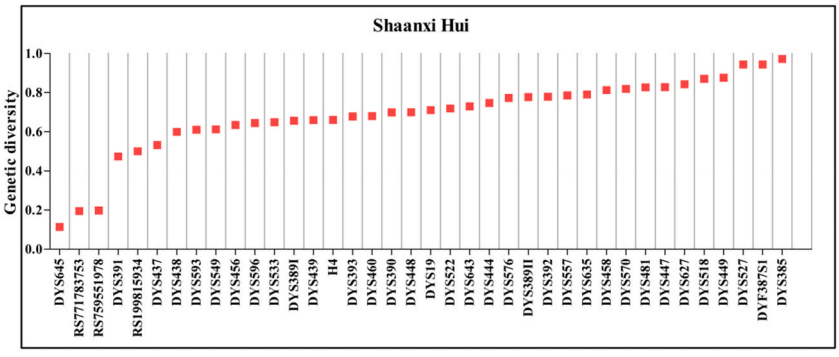


Fig. 1. Genetic diversity (GD) of 38 Y-STR loci included Yfiler™ Platinum system in the Shaanxi Hui population.

neighbour-joining (N-J) tree was generated using molecular evolutionary genetics analysis(MEGA) 7.0 and statistical product and service solutions(SPSS) 25.0, based on the Rst value.

3. Results

3.1. Genetic polymorphisms of 38 Y-STR loci in the Shaanxi Hui population

Allele frequencies and GD values are listed in Table S2. In our study, the allele frequencies of the single-copy loci ranged from 0.001 to 0.940 (DYS645), and the frequencies of the multicopy loci ranged from 0.001 to 0.138 (DYS527). The highest GD value was observed for DYS385 (0.9723), and the lowest was found for DYS645 (0.1138) (Fig. 1). Each locus included in the Yfiler™ Platinum Kit displayed high genetic polymorphisms and was suitable for forensic genetic testing.

Haplotype distribution results are shown in Table 1. A total of 849 distinct haplotypes were identified among 939 Hui individuals based on 38 Y-STR loci, of which 808 (95.2 %) were unique. The genetic parameters, including HD, HMP, and DC, were 0.9996, 0.0015, and 0.9041, respectively, in the Shaanxi Hui population. Allelic mutations, including copy number variants and microvariants, were observed in the present study, as shown in Table S3. The single-copy loci DYS19, DYS390, DYS449, DYS460, DYS518, DYS570, DYS576, and DYS627 exhibited double alleles, and triplicate alleles were observed at the multicopy loci DYS527, DYF387S1, and DYS385. DYF387S1 and DYS385 appeared as quadruple- and five-fold alleles, respectively. Microvariants were detected at loci DYS458, DYS481, DYS518, DYS527, DYS627, and DYS645. Due to the instability of the Y chromosome genome, Y-STRs loci can appear 2–3 copies in the genetic process.The above results have also been reported in previous studies [7,8].

3.2. The genetic differentiation between the Hui group and other Chinese populations

To further explore the genetic relationships between the Shaanxi Hui population and 19 previously reported Chinese populations, Rst and p-values were calculated based on the 27 shared Y-STR loci included on the YHRD website.

Pairwise genetic distances (Rst) between distinct populations were calculated to gain information on genetic traits (Supplementary Table S4). The nearest genetic distance was observed between the Shaanxi Hui and Ningxia Hui populations, and the largest distance was observed between the Shaanxi Hui and Sichuan Tibetan populations. Heatmap visualization of the genetic distance (Fig. 2A and B) showed that the genetic distance between the Han and Hui populations was relatively small, indicating a close genetic relationship between these populations.

MDS plots were drawn to analyse the genetic structure of the 27 Y-STRs, and the results are shown in Fig. 3. All Han and Hui (Sino-Tibetan and Indo-European language groups) groups were located in the first quadrant, the northern minority groups (Altaic language group) were located in the second quadrant (except for the Xinjiang Kazakh group), and the southern minority groups (Sino-Tibetan language groups) were located in the third and fourth quadrants. In this study, the Shaanxi Hui population was closely clustered with the Ningxia Hui, Yunnan Hui, Qinghai Hui, and Xinjiang Hui populations, which is consistent with previous reports [18].

The results of the phylogenetic tree (Fig. 4) showed that the 20 Chinese populations were divided into two branches: Tibetan groups formed one branch, and the remaining 18 ethnic groups formed another branch. The remaining 18 ethnic groups were further divided into three groups: Altaic, Indo-European, and Sino-Tibetan. The Shaanxi Hui population was grouped with the Ningxia Hui population, then with the Yunnan Hui, Qinghai Hui, and Xinjiang Hui populations, and finally with the Yunnan Yi and Han populations within the Sino-Tibetan language family.

There were significant genetic differences between ethnic groups in northern and southern China, the genetic relationships between ethnic groups of the same language family are closer, and the distribution of geography and historical data present relationships that coincide with the genetic relationships between ethnic groups. In fact, since various factors can have a certain influence on ethnic evolution, and the relationship between ethnic origins is quite complex. In order to fully reveal the actual source-flow relationship between ethnic groups, we need to increase the number of research loci and expand the scope of research groups.

According to the statistics, the Hui group is the second largest ethnic minority in China. Based on previous reports [24], the Hui

Table 1
The forensic parameters based on 38 Y-STR loci for the Shaanxi Hui population (n=939).

Time (s) of apperead Haplotypes	Haplotypes size
1 (unique)	808
2	19
3	6
4	10
5	1
6	5
No. of haplotypes	849
Proportion of unique haplotypes	95.20 %
HD	0.9996
MP	0.0015
DC	0.9041

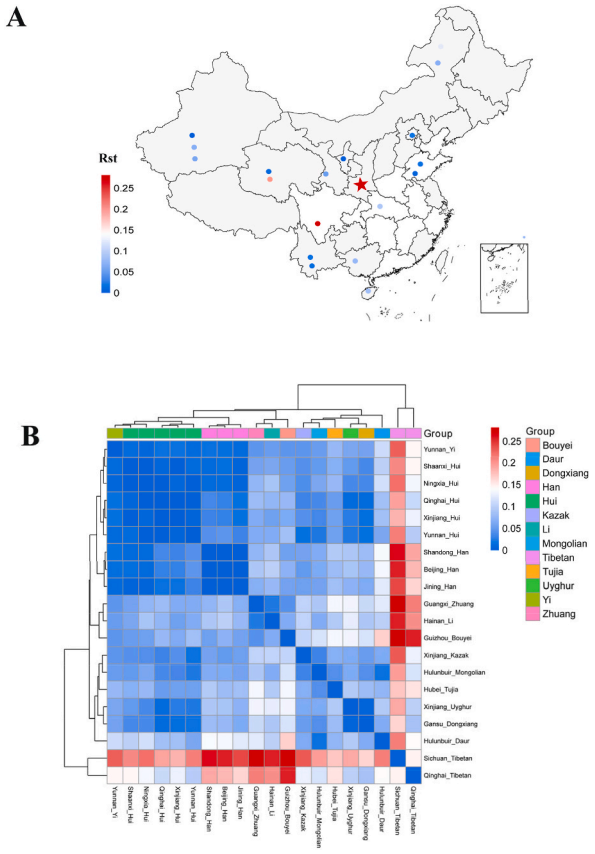


Fig. 2. Heatmap of pairwise genetic distances between Shaanxi Hui and referenced 19 Chinese populations (**A** and **B**).

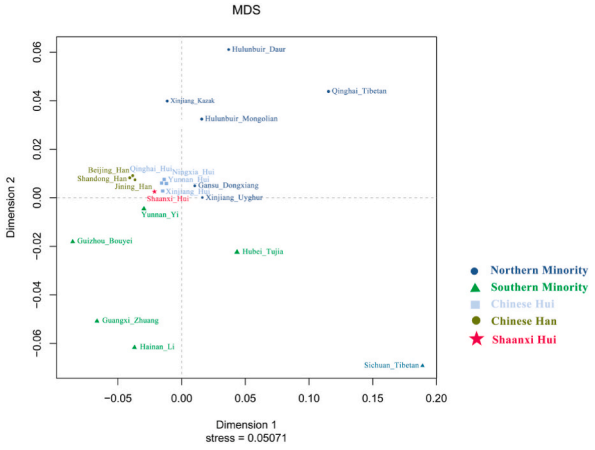


Fig. 3. Multidimensional scale (MDS) analysis map of 20 Chinese groups based on the shared 27 Y-STR loci.

people are genetically close to the Han population, but more closely related to the East Asian population, which consistent with our findings. The genetic distance between Hui nationality and Han nationality is small, and the different genetic structure is directly caused by migration, mixing and intermarriage. The phylogenetic tree was drawn by NJ method, and the results showed that Shaanxi Hui grouped with Ningxia Hui and then formed a branch with people in Yunnan Hui, Qinghai Hui and Xinjiang Hui. The origin of the Hui nationality in Shaanxi and Ningxia can be traced back to the Tang Dynasty, and the Hui nationality is a fusion of many nationalities. This study confirms the evolutionary relationships among different populations in China from the perspective of molecular genetics, which can provide valuable data background for related historical and anthropological research.

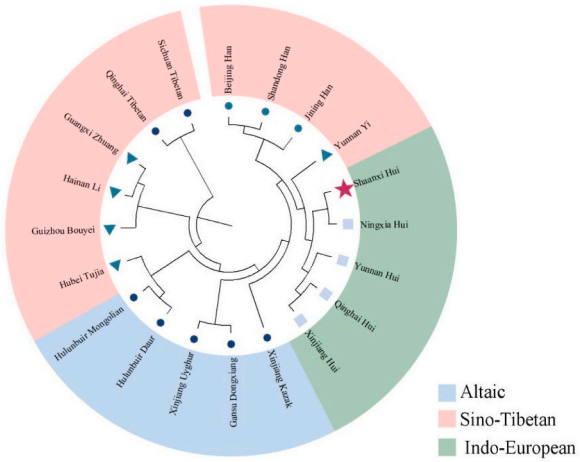


Fig. 4. The phylogenetic tree was rebuilt based on the 27 shared Y-STR loci between the Shaanxi Hui population and the other 19 populations.

4. Conclusion

In the current study, 38 Y-STR loci and three Y-InDels in the Yfiler™ Platinum system demonstrated high genetic polymorphisms in the Shaanxi Hui population. The present data will further enrich the genetic databases of ethnic minority groups and provide valuable information for forensic DNA case detection and population genetics research. Additionally, the results of population comparisons suggest that the Shaanxi Hui population maintains a close genetic relationship with the Ningxia Hui, Qinghai Hui, Xinjiang Hui, and Yunnan Hui populations, which may contribute insights into paternal inheritance and population genetics.

Data accessibility

The datasets generated during and/or analysed during the current study are available from the corresponding author. Part of data were uploaded as supplementary.

Funding information

This work was supported by the Scientific and Technological Research Program of Chongqing Municipal Education Commission (KJON202315134) and the Natural Science Foundation of Chongqing Science and Technology Bureau (CSTB2023NSCQ-BHX0052).

Ethical statement

All samples were collected by the Criminal Investigation Bureau of Shaanxi Public Security Department, and the experimental operations were completed in the genetic laboratory certified by the China National Accreditation Service for Conformity Assessment (CNAS). This work described has been carried out in accordance with the Code of Ethics of the World Medical Association Declaration of Helsinki for experiments involving humans.

CRediT authorship contribution statement

Li Zhang: Writing – original draft, Methodology, Funding acquisition, Conceptualization. **Fei Zhan:** Writing – original draft, Software, Methodology. **Wenfang Zhang:** Methodology, Data curation. **Sanping Song:** Writing – review & editing, Resources, Investigation. **Shisheng Zhu:** Formal analysis, Conceptualization.

Declaration of competing interest

All authors declare no conflicts of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e38501.

References

[1] Jonathan N. Lipman, Familiar Strangers: A History of Muslims in Northwest Chinese, the United States, Seattle, 1997. http://hdl.handle.net/1773/46905.
[2] F. Guo, et al., Population genetic data for 27 Y-STR loci in the Zhuang ethnic minority from Guangxi Zhuang Autonomous Region in the south of China, Forensic science international Genetics 27 (2017) 182–183, https://doi.org/10.1016/j.fsigen.2016.11.009.
[3] Qi, et al., Population genetics study using 26 Y-chromosomal STR loci in the Hui ethnic group in China, Forensic Science International Genetics 28 (2017) 26–27, https://doi.org/10.1016/j.fsigen.2017.01.018.
[4] H. Fan, et al., The Y-STR landscape of coastal southeastern Han: forensic characteristics, haplotype analyses, mutation rates, and population genetics, Electrophoresis 42 (2021) 1578–1593, https://doi.org/10.1002/elps.202100037.
[5] D.C. Gladney, Ethnic identity in China (1998).
[6] M. Kayser, Forensic use of Y-chromosome DNA: a general overview, Hum. Genet. 136 (2017) 621–635, https://doi.org/10.1007/s00439-017-1776-9.
[7] M. Song, et al., Developmental validation of the Yfiler Platinum PCR Amplification Kit for forensic genetic caseworks and databases, Electrophoresis 42 (2021) 126–133, https://doi.org/10.1002/elps.202000187.
[8] L. Li, et al., Haplotype diversity and phylogenetic characteristics for guanzhong han population from northwest China via 38 Y-STRs using Yfiler™ Platinum amplification system, Molecular genetics & genomic medicine 8 (2020) 1187, https://doi.org/10.1002/mgg3.1187.
[9] L. Li, et al., Genetic characterisation for yan'an han population in northern Shaanxi province, China, via 38Y-STRs using Yfiler™ Platinum, Ann. Hum. Biol. 48 (2021) 327–334, https://doi.org/10.1080/03014460.2021.1951350.
[10] S. Cao, et al., Genetic portrait of 27 Y-STR loci in the Tibetan ethnic population of the Qinghai province of China, Forensic science international, Genetics. 34 (2018) 18–19, https://doi.org/10.1016/j.fsigen.2018.02.005.
[11] H. Fan, et al., Population analysis of 27 Y-chromosomal STRs in the Li ethnic minority from Hainan province, southernmost China, Forensic science international, Genetics 34 (2018) 20–22, https://doi.org/10.1016/j.fsigen.2018.01.007.
[12] G.L. He, et al., Genetic polymorphism investigation of the Chinese Yi minority using PowerPlex Y23 STR amplification system, Int. J. Leg. Med. 131 (2017) 663–666, https://doi.org/10.1007/s00414-017-1537-2.
[13] Y. Liu, et al., Haplotype data of 27 Y-STRs analyzed in the Hui and Tujia ethnic minorities from China, Forensic science international, Genetics. 35 (2018) 7–9, https://doi.org/10.1016/j.fsigen.2018.04.006, 2018.
[14] J. Wang, et al., Haplotype structure of 27 YfilerPlus loci in Chinese Dongxiang ethnic group and its genetic relationships with other populations, Forensic science international, Genetics. 33 (2018) 13–16, https://doi.org/10.1016/j.fsigen.2017.12.014.
[15] M. Lang, et al., Forensic characteristics and genetic analysis of both 27 Y-STRs and 143 Y-SNPs in Eastern Han Chinese population, Forensic science international, Genetics 42 (2019) 13–20, https://doi.org/10.1016/j.fsigen.2019.07.011.
[16] Y. Luo, et al., Population genetic analysis of 36 Y-chromosomal STRs yields comprehensive insights into the forensic features and phylogenetic relationship of Chinese Tai-Kadai-speaking Bouyei, PLoS One 14 (2019) 0224601, https://doi.org/10.1371/journal.pone.0224601.
[17] Y. Wang, et al., Genetic diversity and haplotype structure of 27 Y-STR loci in a Han population from Jining, Shandong province, eastern China, Forensic Sci. Int.: Genetics 42 (2019) 25–26, https://doi.org/10.1016/j.fsigen.2019.06.008.
[18] M. Xie, et al., Genetic substructure and forensic characteristics of Chinese Hui populations using 157 Y-SNPs and 27 Y-STRs, Forensic science international, Genetics 41 (2019) 11–18, https://doi.org/10.1016/j.fsigen.2019.03.022.
[19] F. Song, et al., Genetic diversity and phylogenetic analysis of 29 Y-STR loci in the Tibetan population from Sichuan Province, Southwest China, Int. J. Leg. Med. 134 (2020) 513–516, https://doi.org/10.1007/s00414-019-02043-y.
[20] C. Yin, et al., Genetic reconstruction and forensic analysis of Chinese shandong and yunnan han populations by Co-analyzing Y chromosomal STRs and SNPs, Genes 11 (2020) 743, https://doi.org/10.3390/genes11070743.
[21] D. Zhang, et al., Retracted article: Y Chromosomal STR haplotypes in Chinese Uyghur, Kazakh and Hui ethnic groups and genetic features of DYS448 null allele and DYS19 duplicated allele, Int. J. Leg. Med. 135 (2021) 1119, https://doi.org/10.1007/s00414-019-02049-6.
[22] Y. Liu, et al., Forensic characteristics and genetic affinity analyses of Xinjiang Mongolian group using a novel six fluorescent dye-labeled typing system including 41 Y-STRs and 3 Y-InDels, Molecular genetics & genomic medicine 8 (2020) 1097, https://doi.org/10.1002/mgg3.1097.
[23] G. He, et al., Genetic structure and forensic characteristics of Tibeto-Burman-speaking Ü-Tsang and Kham Tibetan Highlanders revealed by 27 Y-chromosomal STRs, Sci. Rep. 9 (2019) 7739, https://doi.org/10.1038/s41598-019-44230-2.
[24] Q. Lan, et al., Genetic structure and polymorphism analysis of Xinjiang Hui ethnic minority based on 21 STRs, Mol. Biol. Rep. 45 (2018) 99–108, https://doi.org/10.1007/s11033-018-4143-6.