

Article

Genetic Polymorphism Analysis of 24 Y-STRs in a Han Chinese Population in Luzhou, Southwest China

Jiewen Fu ^{1,2,3,†} , Binghui Song ^{1,3,†}, Jie Qian ¹ , Ting He ¹, Hanchun Chen ⁴, Jingliang Cheng ^{1,3}  and Junjiang Fu ^{1,3,*} 

- ¹ Key Laboratory of Epigenetics and Oncology, The Research Center for Preclinical Medicine, Southwest Medical University, Luzhou 646000, China; fujiewen@swmu.edu.cn (J.F.); songbinghui@stu.swmu.edu.cn (B.S.); 20210199120039@stu.swmu.edu.cn (J.Q.); 20220199120039@stu.swmu.edu.cn (T.H.); jingliangc@swmu.edu.cn (J.C.)
- ² School for Basic Medicine, Southwest Medical University, Luzhou 646000, China
- ³ Laboratory of Forensic DNA, The Judicial Authentication Center, Southwest Medical University, Luzhou 646000, China
- ⁴ Department of Biochemistry and Molecular Biology, School of Life Sciences, Central South University, Changsha 410013, China; chenhanhunchun@csu.edu.cn
- * Correspondence: fujunjiang@hotmail.com or fujunjiang@swmu.edu.cn
- † These authors contributed equally to this work.

Abstract: Han is the largest of China's 56 ethnic groups and the most populous ethnic group in the world. The Luzhou region is located in southwest China, at the junction of three provinces. The unique historical factors contribute to the genetic polymorphism information. Short tandem repeats (STRs) are highly polymorphic, but the polymorphism of the Y chromosomal STRs (Y-STRs) loci in the Luzhou region is still unclear. It is of great significance to provide Y-STRs genetic data for the Han population from the Luzhou areas of southwest China. A total of 910 unrelated male individuals of the Han population from the Luzhou area were recruited, and 24 Y-STRs were analyzed. The population structure and phylogenetic relationships were compared with those of another 11 related Han populations. A total of 893 different haplotypes were achieved from 910 samples, of which 877 (98.21%) haplotypes were unique. Haplotype diversity and discrimination were 0.999956 and 0.981319, respectively. The lowest genetic diversity of DYS437 is 0.4321, and the highest genetic diversity of DYS385a/b is 0.9642. Pair-to-pair genetic distance and relative probability values indicate that Luzhou Han people are close to Sichuan Han people, Guangdong Han people, and Hunan Han people, which is consistent with geographical distribution, historical influence, and economic development. The 24 Y-STR markers of the southwest Luzhou Han population were highly polymorphic, which provided us with genetic polymorphism information and enriched the population genetic database. Therefore, it is of great value to our forensic applications and population genetics research.

Keywords: Y-STRs; forensic genetics; genetic polymorphism; Han Chinese population



Citation: Fu, J.; Song, B.; Qian, J.; He, T.; Chen, H.; Cheng, J.; Fu, J. Genetic Polymorphism Analysis of 24 Y-STRs in a Han Chinese Population in Luzhou, Southwest China. *Genes* **2023**, *14*, 1904. <https://doi.org/10.3390/genes14101904>

Academic Editor: Nico M. Van Straalen

Received: 10 September 2023

Revised: 27 September 2023

Accepted: 30 September 2023

Published: 2 October 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Short tandem repeats (STRs) are highly polymorphic and widely used in forensic science, anthropology, and genetics. STR genetic markers can be categorized into autosomal short tandem repeats (A-STRs), Y chromosome short tandem repeats (Y-STRs), and X chromosome short tandem repeats (X-STRs), according to their distribution location on chromosomes. The unique genetic approach of sex chromosomes makes them extremely advantageous in the identification of victims of mass disaster events, the investigation of family lineages needed for missing person cases, and the testing of mixed samples in sexual assault crimes. Y-STRs have the characteristics of paternal inheritance, and the results of Y-STR typing of male individuals in the same family are consistent. Y-STR

polymorphism analysis is of great forensic practical value, which plays a critical role in the investigation and solving of cases for forensic casework [1,2]. Their usability and purpose for human identification are limited but can be helpful to enhance the evidence and completely differentiate, for example, in sexual assault cases [1]. Moreover, because Y-STRs are paternally inherited by male offspring and the Y-STR haplotypes within the same paternal lineage are consistent, they have great application value in the study of human origin, evolution, migration, and clan relationships [3]. Through the construction of the Y-STR library, the power of DNA technology can be more effectively exerted. It would be a useful addition to the A-STR repertoire. For example, a violent murder occurred in 2009. Through increasing the number of detected Y-STRs and the combined calculation of multiple A-STRs and the kinship index, Yang et al. [4] successfully identified the perpetrator as a previously unknown illegitimate child of a large family and solved the case in 2021. Their usability and purpose are helpful in enhancing the evidence for sexual assault cases [5]. In particular, they play a role in the personal identification of mixed stains and paternity identification in paternity testing of gang rape cases. Therefore, Y-STR markers have an important application value in forensic genetics.

Y-STR haplotypes are a combination of Y-STR loci alleles. It is necessary to combine multiple Y-STR loci to obtain more haplotypes for improving individual identification ability. At present, there are commonly used commercial Y-STR kits containing multiple markers including PowerPlex[®] Y kits, PowerPlex[®] Y23 kits, Yfiler kits, Yfiler TM Plus kits, etc. [6,7]. The PowerPlex[®] Y23 Kit has been widely used in the analysis of population data and the construction of Y-STR databases, including different ethnic populations in many countries, such as China [8], The United Kingdom [9], Brazil [10], Serbia [11], and so on. The MicroreaderTM24Y Direct ID System includes all 23 Y-STR markers of the PowerPlex Y23 system and another DYS460 locus, i.e., 24 Y-STR loci in total. Therefore, markers in this kit are highly polymorphic [12].

Y-STRs were assigned names by the gene nomenclature committee (HGNC), a committee of the Human Genome Organization (HUGO) that sets the standards for the nomenclature of human genes including STRs. Y-STRs are taken specifically from the Y chromosome in males, which may provide a weaker analysis than A-STRs. The traditional Y-STR system can be used to identify male pedigrees, but it is more difficult to use it to distinguish related male individuals. Thus, the rapidly mutating Y-chromosomal STRs (RM Y-STRs) have relatively high mutation rates ($>1 \times 10^2$ mutation rates than common Y-STRs) and may play an important role in identifying related males. Until now, a large number of studies have been conducted on RM Y-STRs [13,14], which can provide a powerful tool for Y chromosome analysis in forensic genetics, including migratory events in the Luzhou area that would have affected the Y-STR haplotype diversity. In a previous study, our group investigated the 23 Y-STR mutation rates of this system in the southwest Han population [15]. In 175 father-son sample pairs, 16 repeat mutations at 10 loci including three mutations at DYS570, two mutations at DYS549, DYS458, DYS460, and DYS576, and one mutation at another five Y-STRs were found. All the observed repeat mutations showed single repeat changes such as repeat insertions or repeat deletions. Thus, RM Y-STRs are important for kinship analysis, paternal lineage identification, and family relationship reconstruction.

Luzhou, a city in southwest China, is situated upstream of the Yangtze River in southeast Sichuan Province and is at the border between Sichuan Province, Guizhou Province, Yunnan Province, and Chongqing City (please see Figure 1). Yangtze River, or “Long River”, is the longest in the world to flow entirely within one country and the third-longest river in the world. The geographical coordinates in Luzhou are between 27°39′ to 29°20′ north latitude and 105°8′ to 106°28′ east longitude. Based on China’s Seventh National Census Bulletin, the permanent population of Luzhou is 4,254,149, which was published online on 27 May 2021 (http://tjj.luzhou.gov.cn/tjsj/tjgb/content_802318) (accessed on 29 December 2021). Compared with China’s Sixth National Census ten years ago, the increase in the population is 35,723, its growth rate is 0.85% and the average annual growth rate is 0.08%. Although all ethnic groups are rich in population resources,

and the Han nationality is the main group in this region, its relevant population structure and genetic characteristics have not been deeply studied. Moreover, Luzhou has gone through several population migrations since the Qin Dynasty, the most famous being the “Huguang Tian Sichuan” Immigration Movement in the Qing Dynasty. The Han population has integrated with different ethnic groups here, forming a unique and complex group relationship in the region. Thus, it is interesting and important to analyze Y-STRs in the Han population in the Luzhou area and the historical factors that contributed to male lineage diversity. However, studies on the genetic polymorphism of Y-STR in Luzhou, southwest China are scarce, and the number of Y-STR markers and the number of samples are not enough [16]. Herein, we used 24 Y-STR loci (Microreader™ 24Y Direct ID System) to explore genetic polymorphism in the Han Chinese population from the Luzhou areas, and the suspected STR genotyping was verified using the Microreader™ 29Y Direct ID System or the Yfiler™ Plus PCR Amplification Kit.

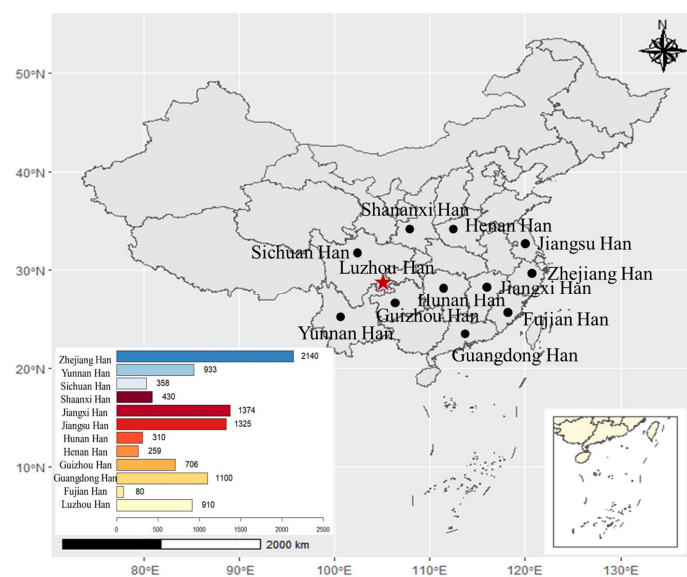


Figure 1. Map showing the approximate geographic positions of the studied Han population in Luzhou areas of southwest and other 11 reference populations in China. The number of haplotypes (the number of samples) in different populations is presented at the bottom left of the figure. The R project software (version 4.0.5) (<https://www.r-project.org/>) (accessed on 29 December 2021) was used to build this map. The star in red represents Luzhou City. The circle “•” in black represents the other 11 areas.

2. Materials and Methods

2.1. Regents

The Microreader™ 24Y Direct ID System was purchased at Microread Genetics Co., Ltd., Suzhou, China, including 24 Y-STR loci, namely DYS19, DYS385ab, DYS391, DYS392, DYS393, DYS570, DYS549, DYS448, DYS458, DYS481, DYS460, DYS635, DYS533, DYS456, DYS389I, DYS390, DYS389II, DYS438, DYS576, DYS439, DYS437, DYS643, and Y GATA H4. The Microreader™ 29Y Direct ID System (Cat. #: 10401207, Microread Genetics Co., Ltd., Suzhou, China) or the Yfiler™ Plus PCR Amplification Kit (Cat. #: 4482730, Thermo Fisher Scientific, Waltham, MA, USA) was also applied to verify the suspected STR genotyping. The Yfiler™ Plus PCR Amplification Kit covers 27 Y-STRs including 7 rapidly mutating (RM) Y-STRs with $>1 \times 10^2$ mutation rates, while the Microreader™ 29Y Direct ID System covers 29 Y-STRs including 7 RM Y-STRs. FTA cards were purchased from Shandong Chengwu Ronghua Biotechnology Co., Ltd., Heze, China (Cat. #: 20201127). Chelex-100 was purchased at Bio-Rad Laboratories Co., Ltd., Hercules, CA, USA (Cat. #: 143-2832). Proteinase K and phenol/chloroform were purchased from Merck KGaA,

Darmstadt, Germany (Cat. #: 1245680100) and Beijing Solarbio Science & Technology Co., Ltd., Beijing, China (Cat. #: T0250), respectively.

2.2. Sample Recruitment and Genomic DNA Extraction

Under informed consent, we recruited blood samples from 910 unrelated male individuals of the Han population in the Luzhou areas, including Sichuan Province, Guizhou Province, Yunnan Province, and Chongqing City in southwest China (Figure 1). The Chelex-100 method [17–19] was used to extract genomic DNA (gDNA) from FTA cards, and the protease K/phenol/chloroform method was used to extract gDNA from EDTA-coated blood samples [20–22]. For details regarding the chelex-100 method: a hole punch was used to collect two 1 mm diameter blood spot samples from the FTA card. It was then placed in a 600 μ L centrifuge tube. A total of 400 μ L double distilled water (ddH₂O) was supplied into a 600 μ L Eppendorf (EP) tube, mixed upside down for 2–3 s, and left at room temperature for 15 min. After centrifuging at 13,200 \times g for 1 min, the blood spot sample was submerged at the bottom of the EP tube. Then the supernatant was removed, 35 μ L 5% chelex-100 was supplied, and the sample was bathed at 56 °C for 30 min. Then the samples were boiled directly in boiling water for 8 min, and finally, the centrifuge tube was turned into a high-speed centrifuge (Thermo Fisher, Waltham, MA, USA) at 13,200 \times g for 1 min. The 20 μ L supernatant was transferred to a new 600 μ L EP tube, which can then be used directly for polymerase chain reaction (PCR) amplification. Samples were stored in the refrigerator at –20 °C for future use.

For whole blood samples, we used the protease K/phenol/chloroform method to extract genomic DNA. DNA extraction from whole blood samples was performed because we have a DNA library for normal samples which can store larger amounts of DNA of better quality for further use. For details, we first took 2 mL of whole blood and placed it in a 10 mL EP tube, added 2.5 times the volume of 1 \times red blood cell lysate (1.55 mol/L NH₄Cl, 0.1 mol/L KHCO₃, 0.001 mol/L EDTA (pH 8.0)), mixed by inversion to make it fully lysate, and then placed it into an ice bath for 15 min before centrifuging for 10 min in a low-speed centrifuge (825 \times g). The supernatant was removed, 700 μ L of 1 \times leukocyte lysate (0.001 mol/L Tris, 0.04 mol/L NaCl, 0.0002 mol/L EDTA (pH 8.0)) was added, mixed well, and then transferred to a new 1.5 ml EP tube with 80 μ L of 20% SDS and 7 μ L of protease K (20 mg/mL). Next, it was mixed and bathed in water at 56 °C for 4 h. Then, an equal volume of balanced phenol was added, and it was shaken and mixed. Next, an equal volume of balanced phenol was added, shaken, and mixed again, before being high-speed centrifuged at 13,200 \times g for 5 min. Approximately 400 μ L of supernatant was transferred to a new 1.5 ml EP tube, and an equal volume of phenol/chloroform was supplied. After shaking and mixing, high-speed centrifugation was performed at 13,200 \times g for 5 min. Next, we transferred 250 μ L of supernatant to a new 1.5 mL EP tube, added 4 times the volume of anhydrous ethanol, mixed upside down, and centrifuged at 13,200 \times g for 1 min. The supernatant was removed, washed twice with 75% ethanol, then sterilized water was added to dissolve the DNA, which can be directly used for polymerase chain reaction (PCR) amplification. Samples were also stored in the refrigerator at –20 °C for future use.

2.3. PCR Amplification and STR Genotyping

The 24 Y chromosome STR loci were amplified using the Microreader 24Y™ Direct ID system (Shandong Chengwu Ronghua Biotechnology Co., Ltd., China, Lot: 20210115) according to the manufacturer's instructions. The Y-STR loci were amplified in an Applied Biosystems Veriti® 96-Well Thermal Cycler (Applied Biosystems, Life Technology, USA). For details, each sample used 10 μ L PCR reaction volume, including DNA sample 1 μ L, ddH₂O 2.8 μ L, Microreader™ 24Y-D 5 \times Primer Mix 2 μ L, Microreader™ Taq DNA Polymerase II 0.2 μ L, and Microreader™ 2.5 \times Buffer B 4 μ L. The steps of PCR amplification were performed as previously described [22,23]. For details, The PCR reaction was set to the following program: pre-denaturation at 96 °C for 2 min; denaturation at 94 °C for 5 s,

annealing and extension at 60 °C for 70 s, cycle number was 28, and final extension at 60 °C for 30 min, maintenance at 16 °C.

After PCR amplification, the 10 µL mixture system was genotyped by capillary electrophoresis, including 1 µL PCR products, deionized formamide 8.5 µL, and Microreader™ Size Standard Org 500 0.5 µL. For quality, control DNA M308 was also added to this kit. Blank control (no DNA template) was used as a negative control to ensure that there was no contamination during the experiment. Genotyping and analysis were performed using the ladder that came with the kit. Our center gets the certification from the China National Accreditation Service for Conformity Assessment (CNAS) and Calibration Laboratories Accreditation Criteria for the Competence of Testing and Calibration Laboratories. Genotype analysis for Y-STRs was conducted by referencing the Specification of Y-STR Testing for Forensic Purpose by China (SF/Z JD0105007-2018) and the International Society of Forensic Genetics' (ISFG) recommendations for analyzing the Y-STR markers [24]. Then the amplified PCR products were used to genotype by capillary electrophoresis on the Genetic Analyzer (3500 DX, Applied Biosystems, Life Technology, USA), and the results were analyzed by the GeneMapper ID-X software (Thermo Fisher, Waltham, MA, USA) [23]. For details, open the analysis software "GeneMapper-X 1.5" and click the "Add sample to project" button to import the electrophoresis file to be analyzed. After import, select "MR24Y_STR_ID" in "Analysis method", "MR24Y Kits V1.3" in "Panel", and "Org 500" in "Size Standard". The STR typing of each sample can be obtained by clicking the "Analysis" button after selecting "Positive Control", "Negative Control", and "Allelic Ladder" for M308, Negative, and Ladder samples, respectively.

2.4. Data Analysis

The indices to evaluate the differential ability of a Y-STR typing system are haplotype diversity (HD) and genetic diversity (GD), calculated by a direct counting method using the formulas $HD = N(1 - \sum p_i^2)/(N - 1)$ and $GD = N(1 - \sum a_i^2)/(N - 1)$, where N is the total number of the target population, p_i is the frequency of the i th haplotype, and a_i is the frequency of the i th allele. The discrimination capacity (DC) was calculated as $DC = M/N$, where M is the number of different haplotypes, and N is the total number of the target population. The haplotype match probability (HMP) was measured using the formula $HMP = \sum p_i^2$, where p_i is the frequency for the i th haplotype. We used the online tools for the Y chromosome haplotype reference database (YHRD) [25,26] (<https://www.yhrd.org/>) (accessed on 29 December 2021) to obtain the pairwise genetic distances (Rst) through Analysis of Molecular Variance (AMOVA), and the multidimensional scaling (MDS) plots based on the Rst values to reveal the population relationships. For details, firstly we selected "AMOVA&MDS" for the Y-STR data analysis under the "Tools" option of the website, and secondly, based on the haplotype data of Luzhou Han and databases of relevant Han Chinese populations in the YHRD, we calculated the Rst and built the MDS plot. Through this method, we can download the analysis results from this website and obtain references of relevant Han Chinese populations. Then the heatmap of Rst between these populations was plotted by OriginPro 2023b software (version 10.0.5.153) [27]. In addition, MEGA-X software (version 10.0.5) [28] was used to construct the neighbor-joining phylogenetic tree based on the Rst values.

3. Results

3.1. Genetic Polymorphisms and Forensic Characteristics

The haplotype data for 24 Y-STRs in the Luzhou Han population are listed in Supplementary Table S1. A total of 893 different haplotypes were obtained from 910 samples, among which 877 (98.21%) haplotypes were unique, 15 (1.68%) haplotypes occurred twice, and one (0.11%) haplotype occurred thrice. There was a total of three microvariants at one Y-STR locus: DYS458. It had three different microvariants, including 13.1, 14.1, and 15.1. Furthermore, a total of five samples were observed copy-number variants. DYS385a/b had a tri-allelic genotype, three tetra-allelic genotypes, and a pentagenic genotype. The mi-

crovariants of locus DYS458 and the copy-number variants of locus DYS385a/b have been frequently observed in previous research on the Han or global population [15,23,29,30]. As shown in Supplementary Figure S1, the eight samples were also detected using the Yfiler™ Plus PCR Amplification Kit (Supplementary Figure S1B) or Microreader™ 29Y Direct ID System (Supplementary Figure S1C), and the result is consistent with the genotyping obtained using the Microreader™24Y Direct ID System (Supplementary Figure S1A).

Then, the allele frequencies and GD values of 24 Y-STR loci including 22 single-copy and one multi-copy Y-STR loci are shown in Supplementary Table S2. Among 22 single-copy Y-STR loci, the number of alleles varied from four in locus DYS437 to 17 in locus DYS481, and allele frequencies varied from 0.0011 to 0.7033. The lowest GD value was 0.4321 in locus DYS437 and the highest was 0.8274 in locus DYS481, followed by 0.8198 in locus DYS458. A multi-copy Y-STR locus DYS385a/b owned more genetic information than all of the 22 single-copy Y-STR loci, with a GD value of 0.9642. Furthermore, twenty of 24 Y-STR loci exhibited GD values greater than 0.6, which demonstrated that the Luzhou Han population had high genetic diversity.

We further evaluated the forensic parameters in four YHRD separate datasets (Minimal, PowerPlex Y12, Yfiler, and PowerPlex Y23) and the Microreader 24Y System, and the results are shown in Table 1. It is found that the increasing number of Y-STR markers improved HD and DC values and decreased HMP values. In the nine minimal loci systems, the discrimination capacity was only 0.7703, including 701 different haplotypes. After the addition of 15 Y-STR markers, the DC values were significantly increased to 0.9813 in the Microreader 24Y System. The number of different haplotypes increased by 27.39% to 893. In addition, the Microreader 24Y System produced the highest HD value of 0.999956.

Table 1. The forensic parameters were evaluated in the five datasets of Y-STR loci in the Han population from the Luzhou area (n = 910).

Count of Observed Haplotype	Minimal	PowerPlex Y	Yfiler	PowerPlex Y23	Microreader 24Y
1	590	674	813	867	877
2	74	70	40	20	15
3	16	12	2	1	1
4	9	6	1	-	-
5	3	3	-	-	-
6	5	1	-	-	-
7	-	1	1	-	-
8	-	1	-	-	-
9	2	-	-	-	-
10	1	-	-	-	-
11	-	-	-	-	-
12	-	-	-	-	-
13	-	-	-	-	-
14	-	-	-	-	-
15	1	-	-	-	-
Number of different haplotypes	701	768	857	888	893
Haplotype match probability	2.31×10^3	1.67×10^3	1.28×10^3	1.15×10^3	1.14×10^3
Haplotype diversity	0.998784	0.999429	0.999823	0.999944	0.999956
Discrimination capacity	0.770330	0.843956	0.941758	0.975824	0.981319

3.2. Population Structure and Population Genetic Relationship

In order to study population structure and phylogenetic relationships between Luzhou Han and relevant Han Chinese populations, we selected previously reported population data in YHRD. As a result, a total of 11 Han populations were chosen as comparison populations, including Henan Han [31], Fujian Han [32], Jiangsu Han [33], Guangdong Han [34], Guizhou Han [35], Hunan Han [36], Jiangxi Han [32], Shanxi Han [37], Sichuan Han [38], Yunnan Han [39], and Zhejiang Han [40], and Figure 1 depicts the approximate geographical location of the Han population in 17 different regions in China. The pairwise Rst values and

associated p -values between the Luzhou Han population and the other 11 Han populations are shown in Supplementary Table S3. We didn't find significant differences in genetic distances between Luzhou Han and Fujian Han, Guangdong Han, Guizhou Han, Hunan Han, and Sichuan Han, according to p -values after Bonferroni correction ($p > 0.00076$). The results demonstrated that the Luzhou Han had the nearest genetic distance with Sichuan Han ($R_{st} = 0.0008$), Guangdong Han ($R_{st} = 0.0008$), and Hunan Han ($R_{st} = 0.0026$). The heatmap of R_{st} is shown in Figure 2. From Figure 2, we can observe that the R_{st} between these populations is more detailed. From Figure 3, Luzhou Han, Sichuan Han, Hunan Han, Guangdong Han, Jiangxi Han, and Zhejiang Han were distributed in the middle part of the MDS plot and were far apart from other Han populations. Moreover, in the neighbor-joining phylogenetic tree (Figure 4), Luzhou Han and Sichuan Han clustered at the same branch. Notably, Luzhou Han was clustered closely with Hunan Han and Guizhou Han. These results were closely associated with the geographical distribution of Han populations, suggesting the influence of social environments and geographical features.

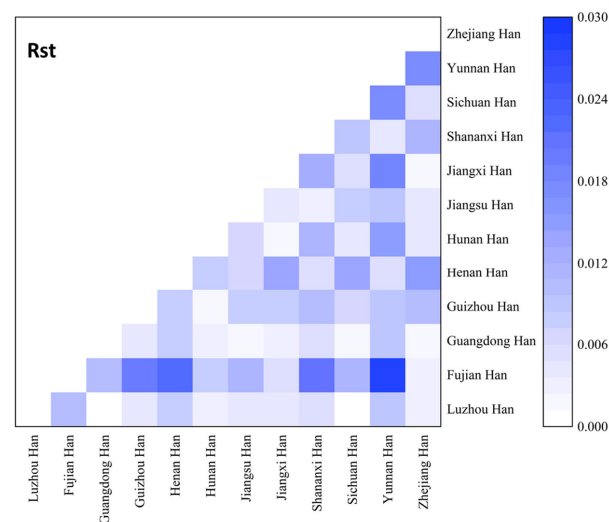


Figure 2. Heatmap of the pairwise genetic distances (R_{st}) of the Luzhou Han population and the other 11 reference populations.

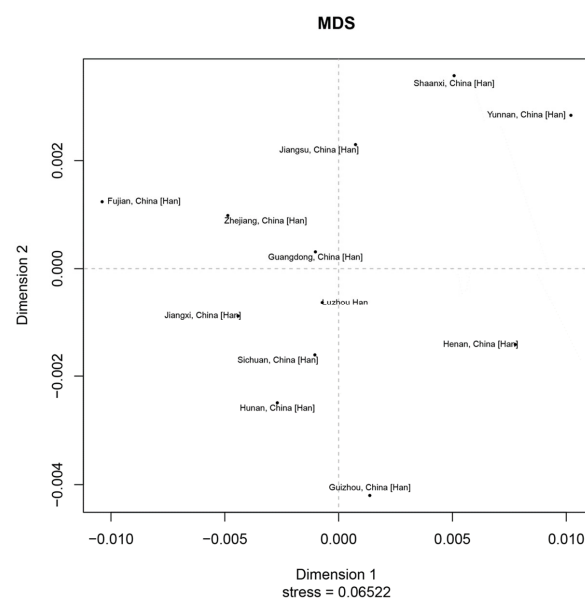


Figure 3. Multidimensional Scaling plots displayed the genetic relationships based on R_{st} values between the Luzhou Han and the other 11 reference populations.

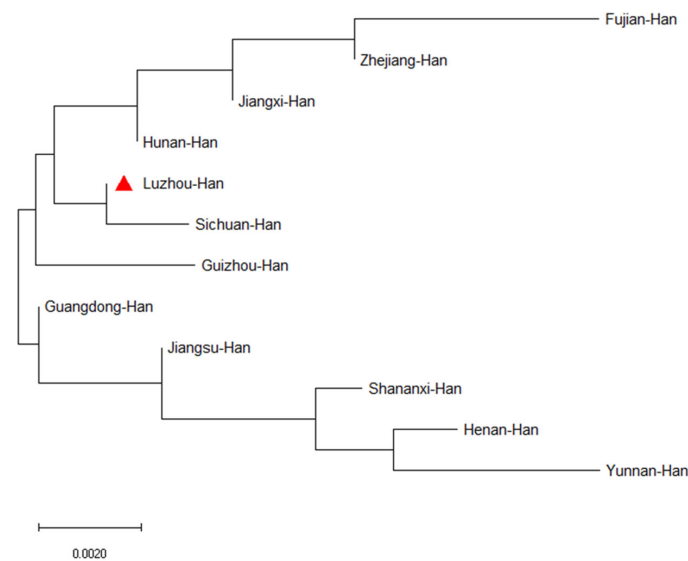


Figure 4. The phylogenetic tree displayed the genetic relationships between the Luzhou population and the other 11 reference populations. The phylogenetic tree was constructed using the neighbor-joining method based on 23 Y-STR loci with MEGA-X software. The red triangle indicates the work in this study.

4. Discussion

China currently has a population of 1,443,497,378. It is the world's most populous country, among which the Han population numbers 12,863,113,334, accounting for 91.11%. Compared with China's Sixth National Census, the increase in the Han population was 60,378,693, with a growth rate of 4.93% (http://www.stats.gov.cn/zt_18555/zdtjgz/zgrkpc/dqcrkpc/ggl/202302/t20230215_1903998.html, accessed on 10 September 2023). Han nationality is the most populous nation in the world, and it has different living customs and cultural characteristics in different regions of China. Located in the southwest of China, Luzhou is the junction of Sichuan Province, Yunnan Province, Guizhou Province, and Chongqing City. In these areas, there are many ethnic minorities, which together with the Han nationality form the unique group structure and genetic characteristics of this region. Due to the high mutation rate of Y-STR loci, it can reflect population structure, population differentiation, and phylogenetic relationships of different populations in a relatively short time. However, the Y-STR data of the Han population in Luzhou is scarce and dated.

In the current study, a total of 893 different haplotypes were obtained from 910 samples, of which 877 (98.21%) haplotypes were unique. The same haplotypes might reflect that these individuals came from the same patrilineal inheritance or shared ancestry. Increasing the number of Y-STR loci, these individuals might also be distinguished. Moreover, the haplotype diversity and discrimination capacity were 0.999956 and 0.981319, respectively. The lowest genetic diversity value was 0.4321 in locus DYS437 and the highest was 0.9642 in locus DYS385a/b. Of course, it is no surprise that increasing the number of Y-STR markers improved HD and DC values and decreased HMP values. It is proven that the 24 Y-STR markers are polymorphic and discriminative in the Luzhou Han population and are appropriate for mass disaster events, identifying paternal lineages, and detecting crime samples.

Population structure and phylogenetic relationships between the Luzhou Han Chinese group and the other 11 relevant Han Chinese groups from the Y chromosome haplotype reference database were compared. The pairwise R_{st} values and associated p -values of AMOVA between the Luzhou Han group and the other 11 Han groups revealed that the Luzhou Han are close with Sichuan Han, Guangdong Han, and Hunan Han populations in genetic relationships, but distantly related to the Han population of Fujian. The MDS plot showed that Luzhou Han, Sichuan Han, and Guangdong Han are close to each other, but

the Han populations in Guizhou province, Shaanxi, and Yunnan are far away. In addition, the neighbor-joining phylogenetic tree showed that Luzhou Han was clustered closely with Sichuan Han, Hunan Han, and Guizhou Han. Among these Han populations, the Fujian Han and Yunnan Han have more distant genetic relationships with Luzhou Han. Although belonging to the same ethnic group, the Han population showed different genetic distances in various places. These also demonstrated that Han Chinese groups from different geographical regions have genetic differences. Distinct cultures, geographic isolation, and differences in intermarriage may have developed in different geographic areas as a result of geographical distance. The closer the geographical distance, the more homogeneous it tends to be, while the farther it is, the more different it is, which explains the diversity of genetic polymorphic information in the Han population. These results proved that geographical distribution, historical influences, economic development, and cultural integration exerted an influence on the population structure and phylogenetic relationships.

In our previous study, we were the first to examine allele frequencies and forensic parameters of 21 A-STR markers in Luzhou Han and the population genetic structure between this Luzhou population and other Chinese populations [19]. In addition, 10 Y-STR loci, including DYS570, DYS549, DYS460, DYS458, DYS576, Y-GATA-H4, DYS635, DYS389II, DYS438, and DYS385ab, were discovered to possess mutations in 175 Chinese father-son pairs [15]. In this study, we first explored allele frequencies and forensic parameters of 24 Y-STR markers in Luzhou Han and the population genetic structure between this studied population and other Chinese populations. The results are similar to previous studies in that the Luzhou Han are significantly closely associated with the Sichuan Han, Guangdong Han, and Hunan Han genetically. The results are also consistent with historical events, reflecting the impact of population migration on genetic diversity. Combining both markers could provide more useful information for population genetics and practice in forensic medicine in Luzhou.

5. Conclusions

In conclusion, our study reveals that the Luzhou Han population is highly genetically polymorphic, including in the allele frequencies, GD values, and forensic characteristic data, in the 24 Y-STR loci, thus could provide useful information for population genetics and practice in forensic medicine in these areas. The research might not only enrich the Y-STR research of population structure, population genetic relationships, and population genetic differentiation in the Luzhou areas of southwest China but also reveal diversity and admixture in these Han populations. This study also ensures the scientific accuracy of the identification results and provides a useful tool for judicial expertise in southwest China.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes14101904/s1>, Figure S1: A. Electropherograms showing the copy-number variants of locus DYS385a/b in five samples by Microreader™24Y Direct ID System. B. The copy-number variants of locus DYS385a/b in five samples were checked by the Yfiler™ Plus PCR Amplification Kit. C. The microvariants of locus DYS458 in three samples were checked by the Microreader™ 29Y Direct ID System; Table S1: 24 Y-STR haplotypes of Han population in Luzhou, China (n = 910); Table S2: Allele frequencies and genetic diversities (GD) of 24 Y-STR loci in Han population from Luzhou, China; Table S3: The pairwise genetic distances (Rst) and associated probability values (p-values) between the Luzhou Han population and the other 11 populations.

Author Contributions: Design of project by J.F. (Junjiang Fu). Analysis of data by B.S., T.H. and J.C. Y-STR analysis by J.F. (Jiewen Fu) and J.C. Draft of manuscript by B.S., J.F. (Jiewen Fu), J.Q., H.C. and J.F. (Junjiang Fu). All authors have read and agreed to the published version of the manuscript.

Funding: National Natural Science Foundation of China, Grant/Award Number: 82073263, 81672887, and 30371493; Research Foundation of Luzhou City, Grant/Award Number: 2021-SYF-37; Translational Medicine Foundation of Southwest Medical University, Grant/Award Number: 00031477.

Institutional Review Board Statement: This study was performed in accordance with the principles stated in the Declaration of Helsinki. This study was approved by the Ethics Committee of Southwest

Medical University, Sichuan Province, China (KY2021237). The informed consent was conducted from male individuals.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: All data used for the analyses in this report are available from the corresponding author on reasonable request.

Acknowledgments: We'd like to thank all the people from the Key Laboratory of Epigenetics and Oncology, the Research Center for Preclinical Medicine, Southwest Medical University. We also thank Shiqiang Yin from the Judicial Authentication Institute, Luzhou Public Security Bureau, Luzhou City, for the technical help.

Conflicts of Interest: The authors declare no competing interests.

References

- Ge, J.; Crysip, B.; Peters, D.; Franco, R.; Liu, M.; Wang, X.; Huang, M.; Budowle, B. MPKin-YSTR: Interpretation of Y chromosome STR haplotypes for missing persons cases. *Electrophoresis* **2023**, *44*, 1080–1087. [[CrossRef](#)] [[PubMed](#)]
- Huffman, K.; Hanson, E.; Ballantyne, J. Y-STR mixture deconvolution by single-cell analysis. *J. Forensic Sci.* **2023**, *68*, 275–288. [[CrossRef](#)] [[PubMed](#)]
- Clairhout, S.; Van der Haegen, M.; Vangeel, L.; Larmuseau, M.H.D.; Decorte, R. A game of hide and seq: Identification of parallel Y-STR evolution in deep-rooting pedigrees. *Eur. J. Hum. Genet.* **2019**, *27*, 637–646. [[CrossRef](#)] [[PubMed](#)]
- Yang, X.; Xu, Q.; Liu, Q.; Liu, C.; Hu, H.; Han, X.; Chen, L.; Liu, C. Combined application of multiple autosomal and Y-chromosomal STR loci in solving a homicide case in 2009. *Int. J. Leg. Med.* **2021**, *135*, 755–760. [[CrossRef](#)] [[PubMed](#)]
- Ferreira-Silva, B.; Fonseca-Cardoso, M.; Porto, M.J.; Magalhaes, T.; Caine, L. A Comparison Among Three Multiplex Y-STR Profiling Kits for Sexual Assault Cases. *J. Forensic Sci.* **2018**, *63*, 1836–1840. [[CrossRef](#)] [[PubMed](#)]
- Diegoli, T.M. Forensic typing of short tandem repeat markers on the X and Y chromosomes. *Forensic Sci. Int. Genet.* **2015**, *18*, 140–151. [[CrossRef](#)] [[PubMed](#)]
- Kaysner, M. Forensic use of Y-chromosome DNA: A general overview. *Hum. Genet.* **2017**, *136*, 621–635. [[CrossRef](#)] [[PubMed](#)]
- Atif, A.; Kasim, K.; Rakha, A.; Noor, A.; Cheema, A.S.; Hadi, S.; Xing, J. Population data of 23 Y STRs from Manchu population of Liaoning Province, Northeast China. *Int. J. Leg. Med.* **2019**, *133*, 785–788. [[CrossRef](#)]
- Aliferi, A.; Thomson, J.; McDonald, A.; Paynter, V.M.; Ferguson, S.; Vanhinsbergh, D.; Syndercombe Court, D.; Ballard, D. UK and Irish Y-STR population data-A catalogue of variant alleles. *Forensic Sci. Int. Genet.* **2018**, *34*, e1–e6. [[CrossRef](#)]
- Ambrosio, I.B.; Braganholi, D.F.; Orlando, L.B.M.; Andrekenas, N.C.; da Mota Pontes, I.; da Silva, D.A.; Astolfi-Filho, S.; de Carvalho, E.F.; Cicarelli, R.M.B.; Gusmao, L. Mutational data and population profiling of 23 Y-STRs in three Brazilian populations. *Forensic Sci. Int. Genet.* **2020**, *48*, 102348. [[CrossRef](#)]
- Kacar, T.; Stamenkovic, G.; Blagojevic, J.; Krtnic, J.; Mijovic, D.; Marjanovic, D. Y chromosome genetic data defined by 23 short tandem repeats in a Serbian population on the Balkan Peninsula. *Ann. Hum. Biol.* **2019**, *46*, 77–83. [[CrossRef](#)]
- Liu, Y.; Guo, Y.; Jin, X.; Mei, S.; Xie, T.; Lan, Q.; Fang, Y.; Zhu, B. Developmental validation study of a 24-plex Y-STR direct amplification system for forensic application. *Int. J. Leg. Med.* **2020**, *134*, 945–951. [[CrossRef](#)] [[PubMed](#)]
- Della Rocca, C.; Trombetta, B.; Barni, F.; D'Atanasio, E.; Hajiesmaeil, M.; Berti, A.; Hadi, S.; Cruciani, F. Improving discrimination capacity through rapidly mutating Y-STRs in structured populations from the African continent. *Forensic Sci. Int. Genet.* **2022**, *61*, 102755. [[CrossRef](#)] [[PubMed](#)]
- Albastaki, A.; Najj, M.; Lootah, R.; Almheiri, R.; Almulla, H.; Alreymi, A.; Almarri, I.; Yu, Z.; Zhang, Y.; Shu, C.; et al. Design and development of novel single multiplex system incorporating 26 rapidly mutating Y-STRs; 26 RM Yplex. *Electrophoresis* **2022**, *43*, 1911–1919. [[CrossRef](#)] [[PubMed](#)]
- Fu, J.; Cheng, J.; Wei, C.; Khan, M.A.; Jin, Z.; Fu, J. Assessing 23 Y-STR loci mutation rates in Chinese Han father-son pairs from southwestern China. *Mol. Biol. Rep.* **2020**, *47*, 7755–7760. [[CrossRef](#)] [[PubMed](#)]
- Bing, L.; Liang, W.; Pi, J.; Zhang, D.; Yong, D.; Luo, H.; Zhang, L.; Lin, Z. Population genetics for 17 Y-STR loci(AmpFISTR®Y-filerTM) in Luzhou Han ethnic group. *Forensic Sci. Int. Genet.* **2013**, *7*, e23–e26. [[CrossRef](#)] [[PubMed](#)]
- Walsh, P.S.; Metzger, D.A.; Higushi, R. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *BioTechniques* **2013**, *54*, 134–139. [[CrossRef](#)] [[PubMed](#)]
- Cheng, J.; Song, B.; Fu, J.; Zheng, X.; He, T.; Fu, J. Genetic polymorphism of 19 autosomal STR loci in the Yi ethnic minority of Liangshan Yi autonomous prefecture from Sichuan province in China. *Sci. Rep.* **2021**, *11*, 16327. [[CrossRef](#)]
- Song, B.; Fu, J.; Qian, J.; Yang, L.; Cheng, J.; Fu, J. Genetic Polymorphism and Population Genetic Structure Analysis of 21 Autosomal STR Loci for a Han-Chinese Population from Luzhou of Southwest China. *Genes* **2023**, *14*, 1419. [[CrossRef](#)]
- Cheng, J.; Peng, J.; Fu, J.; Khan, M.A.; Tan, P.; Wei, C.; Deng, X.; Chen, H.; Fu, J. Identification of a novel germline BRCA2 variant in a Chinese breast cancer family. *J. Cell. Mol. Med.* **2020**, *24*, 1676–1683. [[CrossRef](#)]
- Ghatak, S.; Muthukumar, R.B.; Nachimuthu, S.K. A simple method of genomic DNA extraction from human samples for PCR-RFLP analysis. *J. Biomol. Tech.* **2013**, *24*, 224–231. [[CrossRef](#)] [[PubMed](#)]

22. Fu, J.; Li, L.; Lu, G. Relationship between microdeletion on Y chromosome and patients with idiopathic azoospermia and severe oligozoospermia in the Chinese. *Chin. Med. J. Eng.* **2002**, *115*, 72–75.
23. Fu, J.; Fu, S.; Yin, S.; Cheng, J.; Liu, X.; Jin, Z.; He, T.; Fu, J. Technical note: Multi-alleles at the DYS385ab locus with high frequency in a Han Chinese population from southwestern China. *Int. J. Leg. Med.* **2021**, *135*, 1737–1741. [[CrossRef](#)] [[PubMed](#)]
24. Roewer, L.; Andersen, M.M.; Ballantyne, J.; Butler, J.M.; Caliebe, A.; Corach, D.; D'Amato, M.E.; Gusmao, L.; Hou, Y.; de Knijff, P.; et al. DNA commission of the International Society of Forensic Genetics (ISFG): Recommendations on the interpretation of Y-STR results in forensic analysis. *Forensic Sci. Int. Genet.* **2020**, *48*, 102308. [[CrossRef](#)] [[PubMed](#)]
25. Willuweit, S.; Roewer, L.; International Forensic Y Chromosome User Group. Y chromosome haplotype reference database (YHRD): Update. *Forensic Sci. Int. Genet.* **2007**, *1*, 83–87. [[CrossRef](#)] [[PubMed](#)]
26. Willuweit, S.; Roewer, L. The new Y Chromosome Haplotype Reference Database. *Forensic Sci. Int. Genet.* **2015**, *15*, 43–48. [[CrossRef](#)] [[PubMed](#)]
27. Moberly, J.G.; Bernards, M.T.; Waynant, K.V. Key features and updates for origin 2018. *J. Cheminform.* **2018**, *10*, 5. [[CrossRef](#)] [[PubMed](#)]
28. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [[CrossRef](#)]
29. Purps, J.; Siegert, S.; Willuweit, S.; Nagy, M.; Alves, C.; Salazar, R.; Angustia, S.M.; Santos, L.H.; Anslinger, K.; Bayer, B.; et al. A global analysis of Y-chromosomal haplotype diversity for 23 STR loci. *Forensic Sci. Int. Genet.* **2014**, *12*, 12–23. [[CrossRef](#)]
30. Yang, Y.; Wang, W.; Cheng, F.; Chen, M.; Chen, T.; Zhao, J.; Chen, C.; Shi, Y.; Li, C.; Chen, C.; et al. Haplotypic polymorphisms and mutation rate estimates of 22 Y-chromosome STRs in the Northern Chinese Han father-son pairs. *Sci. Rep.* **2018**, *8*, 7135. [[CrossRef](#)]
31. Wang, L.; Chen, F.; Kang, B.; Zheng, H.; Zhao, Y.; Li, L.; Zeng, Z. Genetic population data of Yfiler Plus kit from 1434 unrelated Hans in Henan Province (Central China). *Forensic Sci. Int. Genet.* **2016**, *22*, e25–e27. [[CrossRef](#)] [[PubMed](#)]
32. Lang, M.; Liu, H.; Song, F.; Qiao, X.; Ye, Y.; Ren, H.; Li, J.; Huang, J.; Xie, M.; Chen, S.; et al. Forensic characteristics and genetic analysis of both 27 Y-STRs and 143 Y-SNPs in Eastern Han Chinese population. *Forensic Sci. Int. Genet.* **2019**, *42*, e13–e20. [[CrossRef](#)] [[PubMed](#)]
33. Wang, H.; Ba, H.; Yang, C.; Zhang, J.; Tai, Y. Inner and inter population structure construction of Chinese Jiangsu Han population based on Y23 STR system. *PLoS ONE* **2017**, *12*, e0180921. [[CrossRef](#)] [[PubMed](#)]
34. Wang, Y.; Zhang, Y.J.; Zhang, C.C.; Li, R.; Yang, Y.; Ou, X.L.; Tong, D.Y.; Sun, H.Y. Genetic polymorphisms and mutation rates of 27 Y-chromosomal STRs in a Han population from Guangdong Province, Southern China. *Forensic Sci. Int. Genet.* **2016**, *21*, 5–9. [[CrossRef](#)] [[PubMed](#)]
35. Sun, H.; Su, K.; Fan, C.; Long, F.; Liu, Y.; Sun, J.; Mo, X.; Ge, Y.; Zhang, L.; Zhai, L.; et al. Y-STRs' genetic profiling of 1953 individuals from two Chinese Han populations (Guizhou and Shanxi). *Forensic Sci. Int. Genet.* **2019**, *38*, e8–e10. [[CrossRef](#)] [[PubMed](#)]
36. Jiang, W.; Gong, Z.; Rong, H.; Guan, H.; Zhang, T.; Zhao, Y.; Fu, X.; Zha, L.; Jin, C.; Ding, Y. Population genetics of 26 Y-STR loci for the Han ethnic in Hunan province, China. *Int. J. Leg. Med.* **2017**, *131*, 115–117. [[CrossRef](#)] [[PubMed](#)]
37. Li, L.; Yao, L.; Gong, H.; Luo, L.; Liang, H.; Zhang, J.; Wang, Y.; He, G.; Jia, F.; Chen, P. Genetic characterisation for Yan'an Han population in Northern Shaanxi Province, China, via 38Y-STRs using Yfiler™ Platinum. *Ann. Hum. Biol.* **2021**, *48*, 327–334. [[CrossRef](#)]
38. Fan, G.Y.; Zhang, Z.Q.; Tang, P.Z.; Song, D.L.; Zheng, X.K.; Zhou, Y.J.; Liu, M.N. Forensic and phylogenetic analyses of populations in the Tibetan-Yi corridor using 41 Y-STRs. *Int. J. Leg. Med.* **2021**, *135*, 783–785. [[CrossRef](#)]
39. Yin, C.; Su, K.; He, Z.; Zhai, D.; Guo, K.; Chen, X.; Jin, L.; Li, S. Genetic Reconstruction and Forensic Analysis of Chinese Shandong and Yunnan Han Populations by Co-Analyzing Y Chromosomal STRs and SNPs. *Genes* **2020**, *11*, 743. [[CrossRef](#)]
40. Fan, G.; Pan, L.; Tang, P.; Zhou, Y.; Liu, M.; Luo, X. Technical note: Developmental validation of a novel 41-plex Y-STR system for the direct amplification of reference samples. *Int. J. Leg. Med.* **2021**, *135*, 409–419. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.