

广西汉族人群20个常染色体STR基因座遗传多态性研究

王小娟^{1*}, 翟翌帆², 周嘉怡², 李格莹², 李志英², 梁书乐², 李明杰^{2,3#}

¹广西众邦司法鉴定中心, 广西 南宁

²广西医科大学第一临床医学院法医学系, 广西 南宁

³广西医科大学第一附属医院病理科, 广西 南宁

收稿日期: 2026年7月6日; 录用日期: 2026年7月29日; 发布日期: 2026年8月12日

摘要

目的: 研究广西汉族人群20个常染色体短串联重复序列(STR)基因座的遗传多态性, 探究其在法医学中的应用价值。方法: 选取实验室2021~2022年采集的广西南宁周边地市汉族无关个体样本516例, 进行STR复合扩增及等位基因分型。计算等位基因频率及群体遗传学参数。结果: 20个STR基因座的杂合度为0.610~0.903, 个体识别能力为0.780~0.981, 累积个人识别能力为0.99999999, 累积非父排除概率为0.99998910。20个基因座中检出等位基因和基因型种类分别为249和943。除D6S1043、D13S317外, 各基因座分布均符合Hardy-Weinberg定律。结论: 20个STR基因座在广西地区汉族人群中具有高遗传多态性和鉴别能力, 所获得的遗传数据可为该地区的法医学及群体遗传学研究提供参考。

关键词

广西, 汉族, STR基因座, 常染色体, 多态性

Genetic Polymorphisms of 20 Autosomal STR Loci in Guangxi Han Population

Xiaojuan Wang^{1*}, Yifan Zhai², Jiayi Zhou², Geying Li², Zhiying Li², Shule Liang², Mingjie Li^{2,3#}

¹Guangxi Zhongbang Judicial Appraisal Center, Nanning Guangxi

²Department of Forensic Medicine, The First Clinical Medical College of Guangxi Medical University, Nanning Guangxi

*第一作者。

#通讯作者。

文章引用: 王小娟, 翟翌帆, 周嘉怡, 李格莹, 李志英, 梁书乐, 李明杰. 广西汉族人群 20 个常染色体 STR 基因座遗传多态性研究[J]. 临床医学进展, 2026, 16(8): 667-673. DOI: 10.12677/acm.2026.1682839

Abstract

Objective: To investigate the genetic polymorphisms of 20 autosomal short tandem repeat (STR) loci in the Han population of Guangxi and to explore their application value in forensic science. **Methods:** A total of 516 unrelated Han individual samples collected from surrounding cities of Nanning, Guangxi, between 2021 and 2022 were selected. STR multiplex amplification and allele genotyping were performed. Allele frequencies and population genetic parameters were calculated. **Results:** The heterozygosity of the 20 STR loci ranged from 0.610 to 0.903. The power of discrimination (PD) for individual loci ranged from 0.780 to 0.981. The combined power of discrimination (CPD) was 0.99999999, and the combined probability of paternity exclusion (CPE) was 0.99998910. A total of 249 alleles and 943 genotypes were observed across the 20 loci. With the exception of D6S1043 and D13S317, the genotype distributions at all loci were in accordance with Hardy-Weinberg equilibrium. **Conclusion:** The 20 STR loci exhibit high genetic polymorphism and discriminative power in the Han population of Guangxi. The obtained genetic data can serve as a reference for forensic science and population genetics research in this region.

Keywords

Guangxi, Han Nationality, STR Loci, Autosome, Polymorphism

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

短串联重复序列(short tandem repeat, STR)是一段包含重复 DNA 基序的区域, 这些基序的长度从 1 到 6 个碱基对不等, 且会在特定区域内重复出现多次。STR 基因组在法医学 DNA 分型研究中发挥着至关重要的作用[1]。基于分子生物学和遗传学的飞速发展, STR 以其结果的高准确性及易于标准化等特点成为目前 DNA 检验的主流技术[1][2]。常染色体荧光标记复合扩增的 STR 分型技术现已广泛应用于法医物证亲权鉴定和个体识别研究, STR 也随之成为国内外法医 DNA 实验室运用最广泛的研究标记之一[2]。STR 技术具有较高的鉴别能力和种属特异性, 通过对多个 STR 基因座进行分型检测, 并计算累计匹配概率, 可判断生物检材间是否具有同一性[3]。同时, STR 引物设计通常针对人类基因组 DNA, 因此在大多数非人类 DNA 样本中不会产生扩增产物或仅有极少数非正常扩增产物, 这使得 STR 技术在人类样本分析中具有较高的特异性[4]。STR 基因组在法医学 DNA 分型研究中具有广泛的应用前景和重要的实践价值。

广西作为一个多民族聚居的自治区, 不同地域和民族之间的遗传学数据差异十分显著[5][6]。通过深入研究 STR 基因组在不同民族之间的遗传学数据, 可以为法医实践提供更准确、可靠的 DNA 鉴定技术支持。此外, 在医学领域也可以利用不同民族之间 STR 数据的差异性来开展相关疾病风险评估和预防工作[7]。针对特定种群常见的遗传性疾病或易感基因变异情况进行调查和分析, 有助于制定个性化治疗方案。

案和健康管理策略[5]。本文基于 20 个常染色体 STR 基因座，对广西地区汉族人群共 516 例样本进行检测，探讨其遗传多态性，为法医学及群体遗传学研究丰富可靠数据。

2. 材料与方法

2.1. 材料

根据知情同意原则，选取广西众邦司法鉴定中心法医物证实验室 2021~2022 年采集的广西南宁周边地市汉族无关个体样本 516 例，个体年龄在 14~90 岁，中位年龄 36 岁；其中男性 420 名，女性 96 名。人群地理主要分布在南宁、玉林、贵港、钦州等地区。无关个体要求：1. 籍贯为广西壮族自治区，民族为汉族；2. 三联体亲权鉴定中只保留父母数据，二联体亲权鉴定中只保留一人数据。

2.2. 方法

采用 Chelex-100 法提取样本 DNA；用 Power Plex21 System 试剂盒对 20 个常染色体 STR 基因座 (D19S433 D5S818 D21S11 D18S51 D6S1043 D3S1358 D13S317 D7S820 D16S539 CSF1P0 PentaD D2S441 vWA D8S1179 TPOX PentaE TH01 D12S391 D2S1338 FGA)进行复合扩增。利用 ABI 3500 遗传分析仪进行毛细管电泳检测，采用 Gene Mapper ID-X 1.5 软件进行基因型分析[2]。并用 Modified-powerstates 软件分析各基因座的等位基因频率(AF)、杂合度(H)、个体识别概率(DP)、非父排除率(PE)、匹配概率(PM)、多态性信息量(PIC)及累积个体识别能力(TDP)，并对各基因座进行 Hardy-Weinberg 平衡检验[4]，采用 Bonferroni 多重校正，以 $P < 0.0025$ 为差异有统计学意义。计算累积个人识别概率(cumulative discrimination power, CDP)和累积非父排除率(cumulative probability of exclusion, CPE)。

3. 结果

3.1. 20 个 STR 基因座等位基因频率

共检出 249 种等位基因，除 D6S1043、D13S317 外，各基因座分布均符合 Hardy-Weinberg 平衡($P < 0.05$)。其中检出等位基因最多的基因座是 PentaE 和 FGA，检出最少的基因座是 D16S539 和 TH01，详见表 1。

Table 1. Allele frequencies of 20 STR loci (n = 516)
表 1. 20 个 STR 基因座等位基因频率(n = 516)

PentaE		D21S11		D19S433		D2S1338		D5S818		D7S820		FGA		D6S1043	
A	F	A	F	A	F	A	F	A	F	A	F	A	F	A	F
5	0.0484	27	0.0029	11	0.0029	12	0.0010	7	0.0349	7	0.0059	13	0.0010	7	0.0010
7	0.0010	28	0.0620	12	0.0379	16	0.0165	8	0.0029	8	0.1292	14	0.0010	8	0.0284
8	0.0039	28.2	0.0010	12.2	0.0058	17	0.0796	9	0.0494	9	0.0568	16	0.0048	9	0.0147
9	0.0136	29	0.2771	13	0.2796	18	0.0718	10	0.2287	9.1	0.0088	17	0.0019	9.1	0.0010
10	0.0504	30	0.2548	13.2	0.0515	19	0.2204	11	0.2917	9.2	0.0010	18	0.0378	10	0.0656
11	0.2267	30.2	0.0087	14	0.2437	20	0.1087	12	0.2413	10	0.1556	19	0.0766	11	0.1869
12	0.1017	30.3	0.0010	14.2	0.1146	21	0.0505	13	0.1463	11	0.3904	20	0.0359	12	0.1595
13	0.0766	31	0.0659	15	0.0631	22	0.0524	14	0.0039	11.1	0.0010	21	0.1328	12.3	0.0020

续表

14	0.0901	31.2	0.0950	15.2	0.1485	23	0.1864	15	0.0010	12	0.2035	21.2	0.0019	13	0.1204
15	0.0640	32	0.0281	16	0.0087	24	0.1379	D13S317		13	0.0450	22	0.1773	14	0.1194
16	0.0572	32.2	0.1463	16.2	0.0369	25	0.0670	A	F	14	0.0029	22.2	0.0058	15	0.0108
17	0.0562	33	0.0019	17	0.0010	26	0.0068	7	0.0039	CSF1P0		23	0.1705	16	0.0039
18	0.0591	33.1	0.0010	17.2	0.0039	27	0.0010	8	0.2290	A	F	23.2	0.0078	17	0.0215
18.4	0.0010	33.2	0.0446	18.2	0.0019	D8S1179		9	0.1037	7	0.0155	24	0.1705	18	0.1282
19	0.0552	34.2	0.0087	PentaD		A	F	10	0.1008	8	0.0010	24.2	0.0107	18.2	0.0029
19.4	0.0019	35.2	0.0010	A	F	8	0.0010	11	0.1800	9	0.0320	25	0.0921	19	0.1027
20	0.0378	D18S51		7	0.0174	10	0.1851	12	0.1047	10	0.2345	25.2	0.0097	20	0.0254
21	0.0194	A	F	8	0.0610	11	0.1502	13	0.0196	11	0.2587	26	0.0388	20.3	0.0020
22	0.0203	10	0.0010	9	0.3769	12	0.1366	14	0.0088	12	0.3798	26.2	0.0039	21	0.0039
23	0.0078	11	0.0058	10	0.1357	13	0.1570	15	0.0949	13	0.0659	27	0.0155	D2S441	
24	0.0078	12	0.0465	11	0.1202	14	0.1269	16	0.0763	14	0.0126	28	0.0039	A	F
D12S391		13	0.1483	12	0.1541	15	0.1647	17	0.0675	vWA		D3S1358		9	0.0010
A	F	14	0.1967	13	0.0814	16	0.0591	18	0.0098	A	F	A	F	9.1	0.0165
15	0.0116	15	0.2103	14	0.0436	17	0.0165	19	0.0010	10	0.0019	8	0.0010	10	0.2277
16	0.0068	16	0.1589	15	0.0087	18	0.0029	TPOX		11	0.0010	11	0.0010	10.1	0.0010
17	0.0514	17	0.0795	19	0.0010	TH01		A	F	13	0.0010	13	0.0049	11	0.3091
18	0.1977	18	0.0407	D16S539		A	F	7	0.0019	14	0.2994	14	0.0313	11.3	0.0901
19	0.1938	19	0.0359	A	F	6	0.1047	8	0.5281	15	0.0252	15	0.3023	12	0.1628
20	0.1957	20	0.0271	8	0.0019	7	0.3062	9	0.0998	16	0.1298	16	0.2965	13	0.0213
21	0.1328	21	0.0213	9	0.2548	8	0.0494	10	0.0271	17	0.2190	17	0.2935	14	0.1628
22	0.1105	22	0.0087	10	0.1085	9	0.4516	11	0.3140	18	0.2093	18	0.0636	14.1	0.0010
23	0.0572	23	0.0097	11	0.3081	9.3	0.0262	12	0.0242	19	0.0959	19	0.0049	15	0.0058
24	0.0300	24	0.0097	12	0.2364	10	0.0601	13	0.0029	20	0.0165	20	0.0010	16	0.0010
25	0.0107			13	0.0804	11	0.0019	15	0.0010	21	0.0010				
26	0.0019			14	0.0097			16	0.0010						

注：A 为等位基因，F 为频率。

3.2. 群体遗传学参数

20 个 STR 基因座的杂合度(He)为 0.610~0.903，个体识别率(DP)为 0.780~0.981，匹配概率(Pm)为 0.019~0.220，非父排除率(PE)为 0.304~0.802，多态性信息含量(PIC)为 0.55~0.89，同时经过计算 20 个 STR 基因座的累计个体识别力(TDP)为 0.9999999999999999，详见表 2。

4. 讨论

本研究检测的 20 个 STR 基因座中，18 个基因座分布符合 Hardy-Weinberg 定律，表现出高度多态性 [6]，说明所调查的群体数据可信度高。本研究共检测出 240 种等位基因和 883 种基因型，其中等位基因

Table 2. Population genetic parameters (n = 516)
表 2. 群体遗传学参数(n = 516)

基因座	Pm	DP	PIC	PE	He	基因座	Pm	DP	PIC	PE	He
D19S433	0.057	0.943	0.79	0.654	0.829	D3S1358	0.128	0.872	0.68	0.483	0.734
D5S818	0.083	0.917	0.74	0.561	0.779	D13S317	0.053	0.947	0.85	0.533	0.763
D21S11	0.056	0.944	0.79	0.651	0.828	D7S820	0.092	0.908	0.73	0.516	0.753
D18S51	0.038	0.962	0.84	0.716	0.860	D16S539	0.094	0.906	0.73	0.573	0.766
D6S1043	0.030	0.970	0.86	0.663	0.834	CSF1P0	0.126	0.874	0.68	0.471	0.727
D2S1338	0.033	0.967	0.85	0.692	0.849	PentaD	0.069	0.931	0.77	0.530	0.762
FGA	0.030	0.970	0.86	0.708	0.857	D2S441	0.073	0.927	0.76	0.568	0.783
PentaE	0.019	0.981	0.89	0.802	0.903	vWA	0.075	0.925	0.76	0.544	0.769
TH01	0.152	0.848	0.64	0.431	0.702	D8S1179	0.043	0.957	0.83	0.739	0.872
D12S391	0.044	0.956	0.83	0.704	0.855	TP0X	0.220	0.780	0.55	0.304	0.610

最多的基因座是 PentaE 和 FGA，均有 21 个等位基因，等位基因最少的基因座是 D16S539 和 TH01，有 7 个等位基因，与刘学军等[6]报道的结果相符。

依据《法医学个体识别技术规范》¹(GB/T 43642-2024)及《常染色体 STR 基因座的法医学参数计算规范》²(SF/Z JD0105010-2018)计算,其个体识别能力达 0.9999999999999999, 累积非父排除概率大于 0.9999, 满足个体识别和亲权鉴定的需求。

本研究所检的 20 个 STR 基因座杂合度(He)介于 0.610~0.903，较高的杂合度意味着该基因座具有更强的个人识别能力[8]。对比广西壮族人群，本研究涉及的 STR 基因座杂合子比例整体较高，其中 D16S539 [7]、vWA [1]基因座杂合度差异显著。此外，D21S11、D18S51、D3S1358、D7S820、D8S1179、TP0X、PentaE、TH01、FGA 基因座也显示较大差异[9] [10]，提示这些基因座在广西汉族人群中具备较高的个体识别效能。

多态信息含量(PIC)均大于 0.5，个体识别能力(DP)平均值为 0.924，非父排除率(PE)平均值为 0.592，均符合高遗传特性指标(DP > 0.8、PE > 0.5)，本研究的 20 个 STR 基因座具有较高遗传特性。基因座 DP ≥ 0.9，He ≥ 0.7 时，表明其具有较高鉴别力[11]，本研究显示 D19S433、D5S818、D21S11 等 16 个 STR 基因座属于高鉴别力、高杂合度基因座，可为个体识别与亲权鉴定提供丰富遗传信息[10]；TH01、D3S1358、CSF1P0、TPOX 基因座在人群中多态性稍低，结果与其他人群[2] [5] [10] [12]-[15]报道一致。

值得注意的是，D6S1043 和 D13S317 两个基因座偏离 Hardy-Weinberg 平衡，其原因可能为样本量较少导致的偏差或群体内亚结构的影响。作为南方汉族的重要分支，广西汉族人群在形成过程中经历了古百越土著与北方移民的深度混合[16]，相互杂居和相互通婚导致基因频率漂移。因此，上述两个基因座的偏离或许从侧面印证了广西汉族群体复杂的微进化历史。

现有多地汉族常染色体 STR 基因座遗传多态性研究已有报道。与南方汉族[17]多态性数据对比发现，D6S1043、D13S317 为华南汉族高多态位点，但不同地域亚群等位基因分布存在小幅差异；广东茂名汉族 [2]、梅州汉族[18]的等位基因频率整体与本研究广西汉族更为接近。基于全国 32 个地区汉族 STR 数据聚类分析[19]证实，两广汉族遗传亲缘关系最近，得益于历史上持续的人口往来；而广西长期存在汉、壮、瑶等多民族混居通婚，少数民族基因流入可造成部分基因座打破随机婚配平衡，与本研究 D6S1043、

¹<https://openstd.samr.gov.cn/bzgk/std/newGbInfo?hcno=F4F093B83244F10EFDCCDF09142E6BB1>

²https://www.moj.gov.cn/pub/sfbgw/zfw/zfwfbgxz/201908/t20190816_150945.html

D13S317 出现 Hardy-Weinberg 偏离的现象吻合。

综上,本研究体现了广西汉族人群 20 个 STR 基因座的遗传多态性特征,为该地区法医学实践与群体遗传学研究提供基础数据。建立适合案件所涉地区的群体遗传学数据库,可准确评估个体识别和亲权鉴定案件的 DNA 分型证据[7]。同时表明 Power Plex21 System 在广西汉族人群中具有高度的多态性和较好的鉴别能力,适用于该地区群体法医学个体识别和亲权鉴定。

本研究仅探究广西汉族单一群体 STR 遗传多态性,缺乏多群体对照数据支撑群体遗传分化分析。未来可进一步收集广西其他人群样本,获取多群体基因型数据,计算群体间 F_{st} 遗传分化指数并构建 PCA、MDS 聚类图,定量解析广西各民族人群遗传亲缘关系与群体分层特征,进一步完善广西地区人群遗传背景数据库。

声 明

本研究已通过广西医科大学第一附属医院伦理委员会批准(批准号:2022-KT-自筹-009)。

基金项目

广西壮族自治区卫生健康委员会自筹经费科研课题(NO: Z-A20220404)。广西医科大学大学生创新创业训练计划项目(NO: S202410598183, S202510598168)。

参考文献

- [1] Tong, D., Chen, Y., Ou, X., Chen, W., Liu, S., Zhang, Y., *et al.* (2013) Polymorphism Analysis and Evaluation of 19 STR Loci in the Han Population of Southern China. *Annals of Human Biology*, **40**, 191-196. <https://doi.org/10.3109/03014460.2012.750685>
- [2] 罗结珊, 黄颖, 邝文健, 等. 粤西茂名地区汉族人群 20 个 STR 基因座遗传多态性[J]. 法医学杂志, 2022, 38(5): 657-661.
- [3] 徐冲冲, 王上, 魏巍, 等. 文山红河地区苗族人群 20 个常染色体 STR 基因座遗传多态性研究[J]. 昆明医科大学学报, 2022, 43(10): 153-162.
- [4] Excoffier, L. and Lischer, H.E. (2010) Arlequin Suite Ver 3.5: A New Senes of Programs to Perform Population Genetics Analyses under Linux and Windows. *Molecular Ecology ResourcesPaperpile*, **10**, 564-567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x>
- [5] 何保仁, 刘学军, 黄惠妮, 等. 广西地区壮族人群 17 个 STR 基因座遗传多态性研究[J]. 中国司法鉴定, 2020(1): 58-60.
- [6] 刘学军, 何保仁, 杨亚丽, 等. 南宁壮族人群 17 个 STR 基因座遗传多态性研究[J]. 广西医学, 2014, 36(10): 1360-1362.
- [7] Chakraborty, R. (1992) Sample Size Requirements for Addressing the Population Genetic Issues of Forensic Use of DNA Typing. *Human Biology*, **64**, 141-159.
- [8] 郭银水, 郭灵军, 何军, 等. 洛阳地区汉族群体 CSF1PO、TH01 和 TPOX 遗传多态性分析[J]. 河南科技大学学报(医学版), 2009, 27(4): 303-304.
- [9] 马浩. 海南黎族、四川汉族群体 57 个常染色体 InDel 位点的遗传多态性及群体遗传结构分析[D]: [硕士学位论文]. 南充: 川北医学院, 2023.
- [10] 张梹, 王斌, 张公赢, 等. 浦东地区汉族人群 23 个常染色体 STR 基因座的遗传多态性[J]. 法医学杂志, 2022, 38(4): 545-550.
- [11] Gill, P., Urquhart, A., Millican, E., Oldroyd, N., Watson, S., Sparkes, R., *et al.* (1996) A New Method of STR Interpretation Using Inferential Logic-Development of a Criminal Intelligence Database. *International Journal of Legal Medicine*, **109**, 14-22. <https://doi.org/10.1007/bf01369596>
- [12] 魏巍, 张寿勋, 徐冲冲, 等. 云南白族民家支系 18 个常染色体 STR 位点的遗传多态性[J]. 昆明医科大学学报, 2022, 43(5): 33-43.
- [13] 杨秀, 牟睿, 曹雨薇, 等. 新疆石河子地区哈萨克族、回族、汉族人群 20 个 STR 基因座的遗传多态性研究[J].

激光生物学报, 2021, 30(4): 339-347.

- [14] 张幼芳, 张巍, 侯思如. 云南回族人群 17 个常染色体 STR 基因座的遗传多态性[J]. 法医学杂志, 2021, 37(4): 574-577.
- [15] 高红艳, 余舰, 冯小丹, 等. 山西运城汉族群体 23 个常染色体 STR 基因座遗传多态性[J]. 中南大学学报(医学版), 2021, 46(4): 351-360.
- [16] 顾有识. 试论历史上的壮汉互为同化[J]. 广西民族研究, 1999(3): 46-51, 75.
- [17] Chen, L., Lu, H.J., Du, W.A., *et al.* (2017) Polymorphism Analysis of 20 Autosomal Short-Tandem Repeat Loci in Southern Chinese Han Population. *International Journal of Legal Medicine*, **131**, 987-994.
- [18] 吴瑛瑕, 张芷清, 李庆昌, 等. 广东梅州客家汉族人群 20 个 STR 基因座遗传多态性[J]. 中国司法鉴定, 2024, 32(3): 78-84.
- [19] 张蒙, 饶健安, 赵艳超, 等. 基于 STR 基因频率探究我国 32 个行政区域汉族亚群的遗传特征[J]. 中山大学学报(自然科学版), 2014, 53(1): 89-95.