

· 研究简报 ·

云南回族人17个常染色体STR基因座的遗传多态性

张幼芳¹, 张巍², 侯思如³

1. 浙江警察学院, 浙江 杭州 310053; 2. 中国刑事警察学院, 辽宁 沈阳 110854; 3. 中国医科大学, 辽宁 沈阳 110122

关键词: 法医遗传学; 多态现象; 遗传; 短串联重复; 遗传距离; 系统发生树; 回族; 云南

中图分类号: DF795.2 文献标志码: B doi: 10.12116/j.issn.1004-5619.2020.501001

文章编号: 1004-5619(2021)04-0574-04



STR是当前法医DNA分析的主流遗传标记,具有较高的突变率,这虽然会影响法医分型判断和鉴定意见的准确性,但为群体遗传学提供了研究基础。因此,常染色体STR基因座被广泛运用于人类群体遗传学研究。本研究对云南回族人17个常染色体STR基因座的遗传多态性进行调查,以期为法医学个体识别、亲权鉴定以及群体遗传学研究提供基础数据。

1 材料与方法

1.1 样本

根据知情同意原则收集725例云南回族无关个体的血滤纸样本。

1.2 DNA提取、扩增及检测

采用打孔器取1.2 mm直径血滤纸进行检验。使用PowerPlex® 18D试剂盒(美国Promega公司)在9700型PCR仪(美国Applied Biosystems公司)上进行直接扩增,扩增体系为10 μ L。采用3130xl基因分析仪(美国Applied Biosystems公司)进行电泳分离。使用GeneMapper® ID v3.2软件(美国Applied Biosystems公司)进行基因分型,等位基因峰高阈值设置为50相对荧光单位(relative fluorescence units, RFU)。以男性标准品2800M(美国Thermo Fisher Scientific公司)作为质量控制的阳性对照,双蒸水作为阴性对照。对于off-ladder(OL)等位基因,采用ForenSeq DNA Signature Prep试剂盒(美国Verogen公司)进行二代测序,使用ForenSeq™ Universal Analysis Software(UAS) v1.3(美国Verogen公司)和STRait Razor 3.0^[1]软件进行序列分析。

1.3 统计学分析

采用Modified Powerstats软件^[2]计算17个常染色

体STR基因座的等位基因频率、匹配概率(probability of match, P_m)、个体识别率(discrimination power, DP)、多态信息含量(polymorphic information content, PIC)、非父排除率(probability of paternity excluding, PE)、亲权指数(parentage index, PI)、观察杂合度(observed heterozygosity, H_o)以及Hardy-Weinberg平衡检验等。

使用Arlequin 3.5软件^[3]进行基因座间连锁不平衡检验,并比较云南回族与其他7个云南群体(白族^[4]、藏族^[5]、拉祜族^[6]、哈尼族^[7]、苗族^[8]、佤族^[9]、汉族^[10])。在除Penta D和Penta E外的15个常染色体STR基因座的群体遗传距离(F_{st})。利用MEGA软件^[11]基于 F_{st} 平均值构建系统发生树。

2 结果

2.1 等位基因频率分布及群体遗传学参数

本研究在云南回族人17个常染色体STR基因座中共检出190个等位基因,等位基因频率分布在0.000 7~0.504 1(表1)。其中,发现两个OL等位基因,分别为FGA上的等位基因26.2(频率为0.000 7)和Penta E上的等位基因25(频率为0.003 4)。尽管这两个OL等位基因在STRBase网站(<https://strbase.nist.gov/>)上均有报道,但所报道的Penta E等位基因25并非使用PowerPlex® 18D试剂盒检测所得的基因分型。二代测序结果显示,该等位基因命名为Penta E_25_[TCTTT]25,侧翼序列分别为上游序列AGAAACTCCTTACAATTT(经UAS分析)和下游序列GAGACTGAGTCTTGCTCAGTCGCCAGGCTGGAGTGCAATGGCGTGATCTCGGCTCACTTCAATCTCCACCTCCT(经STRait Razor分析)。

表 1 17个常染色体STR基因座在云南回族人群中的等位基因频率分布

Tab. 1 Allele frequency distribution of 17 autosome STR loci in Yunnan Hui population (n=725)

<i>D18S51</i>		<i>FGA</i>		<i>D21S11</i>		<i>D7S820</i>		<i>Penta D</i>	
等位基因	频率	等位基因	频率	等位基因	频率	等位基因	频率	等位基因	频率
9	0.0007	16	0.0021	28	0.0400	15.2	0.1441	6	0.0055
11	0.0048	17	0.0034	28.2	0.0041	16	0.0200	7	0.0103
12	0.0469	18	0.0407	29	0.2262	16.2	0.0414	8	0.0379
13	0.2145	19	0.0579	29.2	0.0034	17	0.0014	9	0.3076
14	0.1910	20	0.0552	30	0.3241	10	0.1703	10	0.1276
15	0.1745	20.2	0.0007	30.2	0.0207	11	0.3221	11	0.1469
16	0.1338	21	0.1041	31	0.1124	12	0.2359	12	0.1828
17	0.0766	21.2	0.0117	31.2	0.0621	13	0.0366	13	0.1345
18	0.0414	22	0.1828	32	0.0297	14	0.0034	14	0.0379
19	0.0345	22.2	0.0062	32.2	0.1317	<i>D5S818</i>		15	0.0069
20	0.0338	23	0.2131	33	0.0028	等位基因	频率	16	0.0021
21	0.0193	23.2	0.0076	33.2	0.0407	7	0.0193	<i>D2S1338</i>	
22	0.0186	24	0.1648	34.2	0.0021	8	0.0014	等位基因	频率
23	0.0048	24.2	0.0145	<i>vWA</i>		9	0.0579	16	0.0083
24	0.0041	25	0.0931	等位基因	频率	10	0.2007	17	0.0634
25	0.0007	25.2	0.0021	13	0.0014	11	0.3110	18	0.1193
<i>TH01</i>		26	0.0269	14	0.2152	12	0.2779	19	0.1607
等位基因	频率	26.2 ¹⁾	0.0007	15	0.0393	13	0.1214	20	0.1290
6	0.1103	27	0.0110	16	0.1710	14	0.0090	21	0.0172
7	0.2517	28	0.0007	17	0.2614	15	0.0014	22	0.0531
8	0.0614	30	0.0007	18	0.1855	<i>Penta E</i>		23	0.2262
9	0.4924	<i>D8S1179</i>		19	0.1014	等位基因	频率	24	0.1379
9.3	0.0510	等位基因	频率	20	0.0221	5	0.0345	25	0.0683
10	0.0331	8	0.0014	21	0.0028	7	0.0041	26	0.0076
<i>D3S1358</i>		9	0.0007	<i>D16S539</i>		8	0.0083	27	0.0083
等位基因	频率	10	0.1172	等位基因	频率	9	0.0076	28	0.0007
13	0.0028	11	0.0752	5	0.0007	10	0.0483	<i>D19S433</i>	
14	0.0448	12	0.1269	8	0.0131	11	0.1655	等位基因	频率
15	0.3434	13	0.2076	9	0.2510	12	0.1310	9	0.0007
16	0.3152	14	0.1993	10	0.1366	13	0.0428	10	0.0007
17	0.2290	15	0.1890	11	0.2572	14	0.0731	11	0.0048
18	0.0634	16	0.0662	12	0.2014	15	0.1138	12	0.0421
20	0.0014	17	0.0110	13	0.1110	16	0.0972	12.2	0.0062
<i>TPOX</i>		18	0.0055	14	0.0276	17	0.0703	13	0.2497
等位基因	频率	<i>CSFIPO</i>		15	0.0014	18	0.0628	13.2	0.0600
7	0.0014	等位基因	频率	<i>D13S317</i>		19	0.0434	14	0.2600
8	0.5041	7	0.0034	等位基因	频率	20	0.0310	14.2	0.0986
9	0.1159	9	0.0455	8	0.2469	21	0.0352	15	0.0655
10	0.0303	10	0.2407	9	0.1476	22	0.0159	15.2	0.1441
11	0.3166	11	0.2310	10	0.1359	23	0.0090	16	0.0200
12	0.0290	12	0.3938	11	0.2303	24	0.0028	16.2	0.0414
13	0.0028	13	0.0710	12	0.1807	25 ¹⁾	0.0034	17	0.0014
		14	0.0117	13	0.0434			17.2	0.0048
		15	0.0028	14	0.0152				

注:1)OL等位基因。

本研究中,17个常染色体STR基因座基因型的观察值与期望值经过Bonferroni校正,基因型分布均符合Hardy-Weinberg平衡($P>0.0029$),不存在连锁

不平衡现象($P>0.0004$)。Ho值分布在0.6083~0.8966。*Penta E*和*FGA*为高多态性基因座($PIC>0.85$);*D2S1338*、*D5S818*、*D7S820*等11个基因座为中高度多态性基

因座 (PIC>0.70) ; *CSF1PO*、*D3S1358*、*TH01* 和 *TPOX* 4 个基因座识别能力稍差。Pm 值在 0.017 5~0.196 0, 累积匹配概率 (cumulative match probability, CMP) 达到 7.02×10^{-21} 。DP 值在 0.804 0~0.982 5, 累积个体识别率 (cumulative discrimination power, CDP) 达到 0.999 999 999 999 999 999 993。PE 值在 0.300 9~0.788 4, 三联体累积非父排除率 (cumulative probability of exclusion, CPE) 达到 0.999 999 801。详见表 2。

2.2 遗传距离及系统发生树

云南回族人群与白族、藏族、拉祜族等 7 个人群

15 个常染色体 STR 基因座遗传距离的比较结果见表 3。结果显示, 经 Bonferroni 矫正 ($P < 0.003 3$), 云南回族人群与白族在 12 个基因座上差异具有统计学意义, 与藏族、拉祜族和哈尼族在 11 个基因座上差异具有统计学意义, 与苗族在 8 个基因座上差异具有统计学意义, 与佤族和汉族在 2 个基因座上差异具有统计学意义。

基于云南 8 个人群遗传距离构建的系统发生树见图 1, 其中哈尼族和藏族聚为一类, 回族与佤族、汉族、苗族、拉祜族和白族相对聚集成树。

表 2 17 个常染色体 STR 基因座在云南回族人群中的群体遗传学参数

Tab. 2 Population genetic parameters of 17 autosome STR loci in Yunnan Hui population (n=725)

基因座	Ho	PIC	Pm	DP	PE	PI
<i>D18S51</i>	0.834 5	0.840 4	0.036 9	0.963 1	0.664 5	3.020 8
<i>D21S11</i>	0.822 1	0.781 6	0.067 3	0.932 7	0.640 6	2.810 1
<i>TH01</i>	0.652 4	0.631 6	0.149 0	0.851 0	0.358 5	1.438 5
<i>D3S1358</i>	0.691 0	0.674 7	0.121 0	0.879 0	0.414 5	1.618 3
<i>FGA</i>	0.864 8	0.851 1	0.033 5	0.966 5	0.724 3	3.699 0
<i>TPOX</i>	0.608 3	0.568 7	0.196 0	0.804 0	0.300 9	1.276 4
<i>D8S1179</i>	0.820 7	0.821 5	0.045 0	0.955 0	0.638 0	2.788 5
<i>vWA</i>	0.787 6	0.782 0	0.062 7	0.937 3	0.576 2	2.353 9
<i>CSF1PO</i>	0.717 2	0.681 7	0.122 8	0.877 2	0.455 4	1.768 3
<i>D16S539</i>	0.775 2	0.768 2	0.070 1	0.929 9	0.553 8	2.223 9
<i>D7S820</i>	0.813 8	0.747 3	0.087 8	0.912 2	0.624 9	2.685 2
<i>D13S317</i>	0.791 7	0.783 7	0.062 8	0.937 2	0.583 8	2.400 7
<i>D5S818</i>	0.784 8	0.730 2	0.092 9	0.907 1	0.571 2	2.323 7
<i>Penta E</i>	0.896 6	0.901 9	0.017 5	0.982 5	0.788 4	4.833 3
<i>Penta D</i>	0.798 6	0.789 1	0.059 3	0.940 7	0.596 5	2.482 9
<i>D2S1338</i>	0.855 2	0.846 0	0.035 4	0.964 6	0.705 1	3.452 4
<i>D19S433</i>	0.820 7	0.807 0	0.052 1	0.947 9	0.638 0	2.788 5

表 3 云南回族人群与其他人群的遗传距离比较结果 (P 值)

Tab. 3 Comparison of genetic distance between the Hui population and other populations in Yunnan (P value)

基因座	白族 (n=1 158) ^[4]	藏族 (n=331) ^[5]	拉祜族 (n=1 083) ^[6]	哈尼族 (n=706) ^[7]	苗族 (n=748) ^[8]	佤族 (n=1 542) ^[9]	汉族 (n=2 490) ^[10]
<i>D18S51</i>	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.018 0	0.360 4
<i>D21S11</i>	0.000 0	0.000 0	0.000 0	0.000 0	0.009 0	0.027 0	0.000 0
<i>TH01</i>	0.000 0	0.000 0	0.000 0	0.099 1	0.000 0	0.351 4	0.270 3
<i>D3S1358</i>	0.000 0	0.000 0	0.036 0	0.000 0	0.099 1	0.531 5	0.891 9
<i>FGA</i>	0.009 0	0.000 0	0.000 0	0.045 1	0.000 0	0.315 3	0.774 8
<i>TPOX</i>	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.864 9
<i>D8S1179</i>	0.000 0	0.000 0	0.000 0	0.000 0	0.018 0	0.036 0	0.135 1
<i>vWA</i>	0.000 0	0.054 1	0.009 0	0.000 0	0.018 0	0.018 0	0.000 0
<i>CSF1PO</i>	0.000 0	0.324 3	0.432 4	0.252 3	0.045 1	0.729 7	0.955 0
<i>D16S539</i>	0.000 0	0.000 0	0.000 0	0.000 0	0.009 0	0.000 0	0.144 1
<i>D7S820</i>	0.072 1	0.036 0	0.000 0	0.000 0	0.000 0	0.801 8	0.711 7
<i>D13S317</i>	0.000 0	0.000 0	0.000 0	0.000 0	0.009 0	0.027 0	0.045 1
<i>D5S818</i>	0.000 0	0.009 0	0.000 0	0.000 0	0.000 0	0.495 5	0.144 1
<i>D2S1338</i>	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.216 2	0.468 5
<i>D19S433</i>	0.018 0	0.000 0	0.009 0	0.018 0	0.000 0	0.036 0	0.090 1

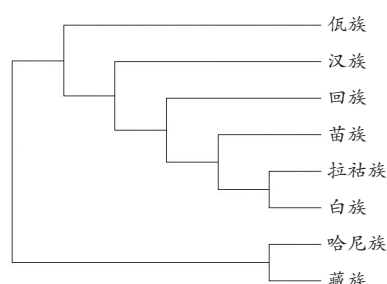


图1 云南8个人群的系统发生树

Fig. 1 Phylogenetic tree of 8 populations in Yunnan

3 讨论

本研究结果显示,在云南回族群体17个常染色体STR基因座中,除*CSF1PO*、*D3S1358*、*TH01*和*TPOX*外,其余13个基因座均具有高多态性和高识别能力($H_o > 0.7$ 、 $PIC > 0.7$ 、 $DP > 0.9$)。所有基因座的基因型分布均符合Hardy-Weinberg平衡且不存在连锁不平衡现象,说明采样随机且各基因座相互独立。根据乘积原则,17个常染色体STR基因座构成的复合扩增体系在云南回族人群中具有较高的个体识别和亲权鉴定的系统效能($CDP > 0.999\ 999\ 999$ 、 $CPE > 0.999\ 999$)。

本研究所调查的回族群体是我国第二大少数民族,主要集中在宁夏回族自治区(20.53%)、甘肃(11.89%)、新疆维吾尔自治区(9.29%)和青海(7.88%)等西北地区,但在全国各地亦存在集中分布,尤其是在河南(9.05%)、云南(6.60%)、河北(5.39%)、山东(5.06%)、安徽(3.10%)、北京(2.35%)、辽宁(2.32%)等地^[12]。同时,云南又是我国拥有少数民族最多的省份,明确各民族人群的遗传多态性和群体遗传关系有助于科学地进行法医学鉴定。通过观察,笔者认为,云南8个民族人群遗传距离的差异与语系有一定的关联:云南回族和汉族的遗传距离比较接近,两者同属于汉藏语系汉语族;云南白族、哈尼族、拉祜族和藏族四者均属于汉藏语系藏缅语族;云南苗族属于汉藏语系苗瑶语族;云南佤族属于南亚语系孟高棉语族。比较特别的是,藏缅语族中,白族和拉祜族比哈尼族和藏族有更多的机会与其他语系或语族的人群进行基因交换,分析可能与民族混居有关。综上,云南回族与其他各民族之间的关系基本符合地域及语系的分布。

综上所述,17个常染色体STR基因座在云南回族人群中具有较高的多态性和较好的识别能力,可满足法医学个体识别和亲权鉴定应用,本研究亦可为群体遗传学和语言学等研究提供基础数据。

参考文献:

[1] WOERNER A E, KING J L, BUDOWLE B. Fast STR allele identification with STRait Razor 3.0[J].

Forensic Sci Int Genet, 2017, 30: 18-23. doi:10.1016/j.fsigen.2017.05.008.

- [2] GUO F, LI J, WEI T, et al. Genetic variation of 17 autosomal STR loci in the Zhuang ethnic minority from Guangxi Zhuang Autonomous Region in the south of China[J]. Forensic Sci Int Genet, 2017, 28: e51-e52. doi:10.1016/j.fsigen.2017.03.015.
- [3] EXCOFFIER L, LISCHER H E L. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows[J]. Mol Ecol Resour, 2010, 10(3): 564-567. doi:10.1111/j.1755-0998.2010.02847.x.
- [4] LI Y, HONG Y, LI X, et al. Allele frequency of 19 autosomal STR loci in the Bai population from the southwestern region of mainland China[J]. Electrophoresis, 2015, 36(19): 2498-2503. doi:10.1002/elps.201500129.
- [5] ZHAI D, YANG J, HUANG Y, et al. The allele frequency of 15 STRs among three Tibeto-Burman-speaking populations from the southwest region of mainland China[J]. Forensic Sci Int Genet, 2014, 13: e22-e24. doi:10.1016/j.fsigen.2014.09.001.
- [6] GUO F. Genetic polymorphism of 17 autosomal STR loci in the Lahu ethnic minority from Yunnan Province, Southwest China[J]. Forensic Sci Int Genet, 2017, 31: e52-e53. doi:10.1016/j.fsigen.2017.08.002.
- [7] HUANG Y, YAO J, LI J, et al. Population genetic data for 17 autosomal STR markers in the Hani population from China[J]. Int J Legal Med, 2015, 129(5): 995-996. doi:10.1007/s00414-015-1176-4.
- [8] ZHANG X, HU L, DU L, et al. Genetic analysis of 20 autosomal STR loci in the Miao ethnic group from Yunnan Province, Southwest China[J]. Forensic Sci Int Genet, 2017, 28: e28-e29. doi:10.1016/j.fsigen.2017.02.003.
- [9] GUO F. Allele frequencies of 17 autosomal STR loci in the Va ethnic minority from Yunnan Province, Southwest China[J]. Int J Legal Med, 2017, 131(5): 1251-1252. doi:10.1007/s00414-017-1620-8.
- [10] XIANG C, HUANG L, SU S, et al. Genetic variation of the 20 STRs in Han population from Yunnan-Kweichow plateau of China[J]. Forensic Sci Int Genet, 2016, 23: e10-e11. doi:10.1016/j.fsigen.2016.03.003.
- [11] KUMAR S, STECHER G, TAMURA K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets[J]. Mol Biol Evol, 2016, 33(7): 1870-1874. doi:10.1093/molbev/msw054.
- [12] GUO F. Population genetics for 17 Y-STR loci in Hui ethnic minority from Liaoning Province, Northeast China[J]. Forensic Sci Int Genet, 2017, 28: e36-e37. doi:10.1016/j.fsigen.2017.02.011.

(收稿日期:2020-09-30)

(本文编辑:张素华)