

·论 著·

71个Y-SNP位点在西北汉族人群的多态性及法医学应用价值

张立男^{1,2}, 宋雨桐^{1,2}, 姜磊¹, 朱如心¹, 李莉¹

(1. 司法鉴定科学研究院 上海市法医学重点实验室 上海市司法鉴定专业技术服务平台, 上海 200063; 2. 华东政法大学, 上海 200063)

摘要: 目的 分析Y染色体上71个SNP位点在西北地区汉族人群的多态性, 评价其法医学应用价值, 并整合前期研究成果, 筛选出适用于华东、华南、西北地区汉族人群法医学检验的Y-SNP位点。方法 采用多重聚合酶链反应(polymerase chain reaction, PCR)和基质辅助激光解吸/电离-飞行时间质谱(matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, MALDI-TOF-MS)技术对西北地区202名汉族男性无关个体进行71个Y-SNP位点的分型, 计算等位基因频率和基因多样性(gene diversity, GD)、单倍型多样性(haplotype diversity, HD)值, 同时结合前期研究所取得的数据筛选出适用于西北、华南和华东汉族群体的Y-SNP位点。结果 在所检测的71个位点中, 有67个位点在西北汉族男性人群中呈多态性分布, GD值在0.010 0~0.502 2, 中度信息量($0.2 \leq GD < 0.3$)的位点有22个, 高度信息量($GD \geq 0.3$)的位点有25个, 同时适用于西北、华南和华东汉族群体的位点有26个。结论 Y-SNP位点在亲权鉴定和个体识别中具有应用价值, 在判断人群分布及迁移信息方面亦具有潜在的应用价值。

关键词: 法医遗传学; Y染色体; 多态性; 单核苷酸; 等位基因; 汉族

中图分类号: DF795.2 文献标志码: A doi: 10.12116/j.issn.1004-5619.2019.01.005

文章编号: 1004-5619(2019)01-0023-07

Polymorphism and Forensic Application Value of 71 Y-SNP Loci in Han Population of Northwest China

ZHANG Li-nan^{1,2}, SONG Yu-tong^{1,2}, JIANG Lei¹, ZHU Ru-xin¹, LI Li¹

(1. Shanghai Key Laboratory of Forensic Medicine, Shanghai Forensic Service Platform, Academy of Forensic Science, Shanghai 200063, China; 2. East China University of Political Science and Law, Shanghai 200063, China)

Abstract: Objective To analyze the polymorphism of 71 SNP loci on Y chromosome in Han population of Northwest China, to assess its forensic application value, and to screen out Y-SNP loci for forensic examination of Han population in East, South, and Northwest China based on the integration of previous research results. **Methods** Multiplex polymerase chain reaction (PCR) and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) were performed on 71 Y-SNP loci of 202 unrelated Han male individuals in Northwest China. Gene diversity (GD) and haplotype diversity (HD) values were calculated, and then Y-SNP loci of Han population in East, South, and Northwest China were screened with the combination of data from previous research. **Results** Among the detected 71 loci, 67 loci were polymorphic in the Northwest Han male population, with GD values 0.010 0~0.502 2. There were 22 and 25 loci with a moderate ($0.2 \leq GD < 0.3$) and high ($GD \geq 0.3$) amount of genetic information, respectively. There were 26 loci for the Han communities in Northwest, South, and East China. **Conclusion** Y-SNP loci are potential in paternity testing and individual identification, as well as the judgement of population distribution and migration.

Keywords: forensic genetics; Y chromosome; polymorphism, single nucleotide; allele frequency; Han nationality

基金项目:“十三五”国家重点研发计划资助项目(2016YFC0800703);上海市法医学重点实验室资助项目(17DZ2273200);上海市司法鉴定专业技术服务平台资助项目(16DZ2290900)

作者简介:张立男(1992—),男,硕士研究生,主要从事法医物证学研究;E-mail:HitboyZLN@163.com

通信作者:李莉,女,研究员,硕士研究生导师,主要从事法医物证学科研、检案和教学;E-mail:lil@ssfjd.cn

单核苷酸多态性(single nucleotide polymorphism, SNP)是指在基因组水平上由单个核苷酸变异引起的DNA序列多态性,是人类可遗传变异中最常见的一种,占有已知多态性的90%以上。SNP具有突变率低、可稳定遗传、部分位点与个体外观表现有关^[1]、有利于降解检材分析^[2]、便于高通量自动化分型检测等优点^[3]。

Y染色体为男性特有的性染色体,除非发生突变,否则同一父系的男性后代均具有相同的单倍型,故可以用来揭示不同人群的起源、迁徙及亲缘关系等。Y染色体上SNP位点(Y-SNP)分型检测方法的建立,可以辅助现有的Y-STR检测体系,协助进行疑难案件的鉴定。本研究采用多重PCR联合基质辅助激光解吸/电离-飞行时间质谱(matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, MALDI-TOF-MS)技术对西北汉族人群71个Y-SNP位点进行多态性调查,并结合前期研究所得,筛选同时适用于西北、华南和华东汉族群体的位点,旨在为法医学应用提供基础数据。

1 材料与方法

1.1 样本

在知情同意原则下,收集西北新疆地区汉族男性

无关个体血样202份,每份样本取200 μL。

1.2 主要试剂和仪器

DNA抽提试剂盒QIAamp DNA Blood Mini试剂盒(德国Qiagen公司),基因分型系统Complete iPLEX[®] Gold Genotyping Reagent Sets(96 Format,美国Agena公司),9700型PCR仪(美国AB公司),MassARRAY[®]飞行时间质谱仪、MassARRAY[®]纳升点样仪RS1000(美国Agena公司),MassARRAY Typer 4.0软件(美国Agena公司)。

1.3 位点的选择、优化及引物设计

通过Y染色体系统树、Hapmap数据库^[4]、美国国立生物技术信息中心(National Center for Biotechnology Information,NCBI)数据库^[5]进行检索,结合《法医SNP分型与应用规范》(SF/Z JD0105003—2015)的要求,从Y染色体上筛选出在东亚人群中具有多态性的71个SNP位点,用Agena Bioscience公司的在线工具^[6]设计用于PCR扩增的引物和单碱基延伸反应的引物。

依据各个位点之间的距离并避免引物之间的相互干扰,将71个位点分成W1、W2、W3复合体系(表1)进行检测,PCR扩增子的大小为80~120 bp,单碱基延伸产物的相对分子质量为4429~8710。

表1 3个复合扩增体系所包含的Y-SNP位点

体系	位点
W1	rs13447361、rs52812045、rs2075181、SRY8299、M134、rs9786247、rs2267801、rs9786707、rs3848982、rs4589047、M119、rs17250121、rs16980426、rs3906、rs16980363、rs2032650、rs3898、rs17269816、rs17269928、rs2032674、rs16980601、rs9786502、rs2032665、rs2032652、P31、rs16981340、rs17269396、rs9306841、M117、rs17276358、rs17276338、rs3900
W2	rs2032678、rs11575897、rs9786479、rs9786401、rs2075640、rs16980391、rs17316592、rs9341278、rs17276345、rs2032602、rs16980396、rs11096432、rs17316543、rs17842387、rs2032624、rs35284970、rs17323322、rs9785908、M148、rs16981290、rs9306848、rs917759、rs17316007、rs7067458、rs9786095、rs13447354、rs16980641、rs2032645、rs9786394、rs11096433、rs16980610
W3	M122、rs2196155、rs17174528、rs2032631、rs3853054、rs9306845、M111、rs17276777

1.4 PCR扩增和产物纯化

用QIAamp DNA Blood Mini试剂盒进行DNA抽提。扩增体系为5 μL,含10×PCR缓冲液0.5 μL、25 mmol/L MgCl₂ 0.4 μL、PCR引物混合液0.5 μL、25 mmol/L dNTP混合液0.1 μL、5 U/μL PCR酶0.1 μL、纯水1.4 μL、模板DNA溶液2 μL。PCR循环参数:95℃ 2 min;95℃ 30 s,56℃ 30 s,72℃ 60 s,45个循环;72℃ 5 min。PCR扩增结束后,加入1.7 U/μL 虾碱性磷酸酶(shrimp alkaline phosphatase, SAP)0.30 μL、SAP缓冲液0.17 μL和超纯水1.53 μL,按下述条件进行酶切反应:37℃ 40 min,85℃ 5 min。

1.5 单碱基延伸反应

在纯化后的产物中加入10×iPLEX[®]缓冲液

0.20 μL、单碱基延伸引物混合液0.94 μL、10×iPLEX[®]终止混合物0.20 μL、iPLEX[®]酶0.04 μL、超纯水0.62 μL,按下列条件进行单碱基延伸反应:94℃ 5 s;52℃ 5 s,80℃ 5 s,5次小循环;40次大循环。

1.6 产物检测

按参考文献[7]中的方法对产物脱盐后,使用MassARRAY[®]纳升点样仪RS1000将反应产物点样至芯片上(点样量控制在8~12 nL),将点样完毕的芯片放入MassARRAY[®]飞行时间质谱仪进行质谱检测,用MassARRAY Typer 4.0软件查看分型结果。

1.7 统计学分析

用直接计数法计算各Y-SNP位点在西北汉族无关男性人群中的等位基因频率,基因多样性(gene di-

iversity, GD)计算公式:

$$GD=[n(1-\sum f_i^2)]/(n-1), \quad (1)$$

式中, n 为样本数, f_i 为第 i 个等位基因的分布频率。单倍型多样性(haplotype diversity, HD)^[8]计算公式:

$$HD=[n(1-\sum P_i^2)]/(n-1), \quad (2)$$

式中, n 为单倍型数, P_i 为第 i 个单倍型的频率。用 Arlequin v3.5 软件^[9]计算单倍型频率、 F_{st} 值和 P 值, 并进行群体遗传学比较。 F_{st} 是种群分化和遗传距离的一种衡量方法, WRIGHT^[10]建议: F_{st} 为 0~0.05, 群体

间遗传分化可以不考虑; F_{st} 为 0.05~0.15, 群体间存在中等程度的遗传分化; F_{st} 为 0.15~0.25, 群体间遗传分化较大; $F_{st}>0.25$, 则群体间有很大的遗传分化。

2 结 果

2.1 西北汉族群体 Y-SNP 位点多态性分析

经过实验和统计, 得到西北地区汉族无关男性人群 71 个 Y-SNP 位点的等位基因频率数据和 GD 值, 有 67 个位点在西北汉族人群中呈多态性分布, 详见表 2。

表 2 西北汉族无关男性人群 71 个 Y-SNP 位点的等位基因频率和 GD 值 (n=202)

位点	等位基因	频率	GD 值	位点	等位基因	频率	GD 值
M111	TT/-	0.985 1/0.014 9	0.029 6	rs2032624	A/C	0.980 1/0.019 9	0.039 2
M117	ATCT/-	0.875 6/0.124 4	0.218 9	rs2032631	G/A	0.855 7/0.144 3	0.248 2
M119	A/C	0.850 7/0.149 3	0.255 2	rs2032645	A/G	1.000 0/0.000 0	0.000 0
M122	T/C	0.543 2/0.457 7	0.498 9	rs2032650	C/T	0.990 0/0.010 0	0.019 8
M134	G/-	0.801 0/0.199 0	0.320 4	rs2032652	T/C	0.865 7/0.134 3	0.233 7
M148	A/-	1.000 0/0.000 0	0.000 0	rs2032665	C/T	0.928 6/0.071 4	0.133 3
P31	T/C	0.970 1/0.029 9	0.058 2	rs2032674	C/T	0.781 1/0.218 9	0.343 7
rs11096432	C/T	0.756 2/0.243 8	0.370 5	rs2032678	-/TTCTC	0.860 7/0.139 3	0.241 0
rs11096433	C/T	0.910 4/0.089 6	0.163 9	rs2075181	C/T	0.845 8/0.154 2	0.262 2
rs11575897	T/C	0.005 0/0.995 0	0.010 0	rs2075640	A/G	0.900 5/0.099 5	0.180 1
rs13447354	G/A	0.910 4/0.089 6	0.163 9	rs2196155	G/A	0.696 5/0.303 5	0.424 9
rs13447361	G/C	0.572 1/0.427 9	0.492 0	rs2267801	T/C	0.761 2/0.238 8	0.365 4
rs16980363	G/C	0.771 1/0.228 9	0.354 7	rs35284970	C/T	0.885 6/0.114 4	0.203 7
rs16980391	A/G	0.870 6/0.129 4	0.226 4	rs3848982	G/A	0.980 1/0.019 9	0.039 2
rs16980396	T/C	0.855 7/0.144 3	0.248 2	rs3853054	A/C	0.801 0/0.199 0	0.320 4
rs16980426	G/T	0.801 0/0.199 0	0.320 4	rs3898	C/G	0.990 0/0.010 0	0.019 8
rs16980601	G/A	0.741 3/0.258 7	0.385 5	rs3900	G/C	0.850 7/0.149 3	0.255 2
rs16980610	A/G	0.766 2/0.233 8	0.360 1	rs3906	GG/TTGTACAGAGA	0.990 0/0.010 0	0.019 8
rs16980641	G/A	0.781 1/0.218 9	0.343 7	rs4589047	T/A	0.845 8/0.154 2	0.262 2
rs16981290	A/C	0.761 2/0.238 8	0.365 4	rs52812045	C/T	0.845 8/0.154 2	0.262 2
rs16981340	T/C	0.890 5/0.109 5	0.195 9	rs7067458	A/G	0.878 8/0.121 2	0.214 1
rs17174528	T/G	0.850 7/0.149 3	0.255 2	rs917759	A/C	0.746 3/0.253 7	0.380 6
rs17250121	T/C	0.855 7/0.144 3	0.248 2	rs9306841	C/G	1.000 0/0.000 0	0.000 0
rs17269396	G/A	0.744 2/0.255 8	0.382 6	rs9306845	A/G	0.840 8/0.159 2	0.269 1
rs17269816	C/T	0.532 3/0.497 7	0.500 4	rs9306848	C/A	0.850 7/0.149 3	0.255 2
rs17269928	A/G	0.547 3/0.452 7	0.498 0	rs9341278	G/A	0.995 0/0.005 0	0.010 0
rs17286338	G/T	0.711 4/0.288 6	0.412 6	rs9785908	G/T	0.865 7/0.134 3	0.233 7
rs17276345	C/G	0.935 3/0.064 7	0.121 6	rs9786095	C/T	0.890 5/0.109 5	0.195 9
rs17276358	G/T	0.512 4/0.487 9	0.502 2	rs9786247	G/A	0.895 5/0.104 5	0.188 1
rs17276777	G/A	0.870 6/0.129 4	0.226 4	rs9786394	A/G	0.764 7/0.235 3	0.361 7
rs17316007	A/G	0.801 0/0.199 0	0.320 4	rs9786401	C/G	0.905 5/0.094 5	0.172 0
rs17316543	G/C	0.761 2/0.238 8	0.365 4	rs9786479	T/G	0.975 1/0.024 9	0.048 8
rs17316592	G/T	0.552 2/0.447 8	0.497 0	rs9786502	T/C	0.840 8/0.159 2	0.269 1
rs17323322	T/C	0.746 3/0.253 7	0.380 6	rs9786707	C/T	0.865 7/0.134 3	0.233 7
rs17842387	A/G	0.870 6/0.129 4	0.226 4	SRY8299	C/A	1.000 0/0.000 0	0.000 0
rs2032602	T/C	0.980 1/0.019 9	0.039 2				

表2显示,在所选择的位点中除M148、rs2032645、rs9306841、SRY8299外,其余67个位点都具有多态性,GD值在0.0100(rs11575897和rs9341278)~0.5022(rs17276358)。中度信息量(0.2≤GD<0.3)的位点有22个(M117、M119、rs2032631、rs2032652、rs2032678、rs2075181、rs35284970、rs16980391、rs16980396、rs3900、rs4589047、rs52812045、rs7067458、rs17174528、rs17250121、rs9306845、rs9306848、rs9785908、rs17276777、rs9786502、rs9786707、rs17842387),高度信息量(GD≥0.3)的位点有25个(M122、M134、rs2032674、rs11096432、rs2196155、rs13447361、rs2267801、rs16980363、rs3853054、rs16980426、

rs16980601、rs16980610、rs16980641、rs16981290、rs917759、rs17269396、rs17269816、rs17269928、rs17286338、rs17276358、rs9786394、rs17316007、rs17316543、rs17316592、rs17323322)。

使用Arlequin v3.5统计软件分析71个Y-SNP位点的单倍型发现,本群体共有170种单倍型,其中频率最高的为0.0348(7/202),频率最低的只有0.0050(1/202)。经计算,HD为0.9930。

2.2 华东、华南和西北汉族群体比较

结合前期的实验结果^[7-11],用Arlequin v3.5软件计算 F_{st} 值和 P 值,对华东、华南和西北汉族无关男性人群的等位基因分布进行比对分析(表3)。

表3 71个Y-SNP位点在华东、华南和西北汉族无关男性人群的等位基因分布差异

位点	华东-西北		华东-华南		西北-华南	
	F_{st} 值	P 值	F_{st} 值	P 值	F_{st} 值	P 值
M111	0.0043	1.0000	0.0027	0.5405	0.0155	0.5159
M117	0.0042	0.7667	0.0207	0.0264	0.0419	0.0124
M119	0.0043	1.0000	0.0037	0.6828	0.0422	0.7836
M122	0.0067	0.1385	0.0002	0.3169	0.0235	0.0156
M134	0.0011	0.4051	0.0108	0.0762	0.0330	0.0119
M148	0.0000	1.0000	0.0000	1.0000	0.0230	1.0000
P31	0.0412	0.0028	0.0015	0.4467	0.0045	0.0002
rs11096432	0.0018	0.4769	0.0097	0.0920	0.0266	0.0161
rs11096433	0.0107	0.0905	0.0037	0.6823	0.0234	0.0353
rs11575897	0.0049	1.0000	0.0032	0.6198	0.0257	1.0000
rs13447354	0.0063	0.1566	0.0049	1.0000	0.0157	0.1573
rs13447361	0.0154	0.0473	0.0031	0.5491	0.0220	0.0121
rs16980363	0.0141	0.0625	0.0129	0.0683	0.0305	0.0003
rs16980391	0.0047	0.8812	0.0049	0.1889	0.0032	0.1436
rs16980396	0.0029	0.6469	0.0049	0.1871	0.0085	0.0637
rs16980426	0.0027	0.2386	0.0129	0.0658	0.0009	0.0027
rs16980601	0.0014	0.2888	0.0149	0.0472	0.0296	0.0021
rs16980610	0.0040	0.7230	0.0144	0.0253	0.0143	0.0091
rs16980641	0.0056	0.1612	0.0042	0.1095	0.0034	0.0025
rs16981290	0.0018	0.4717	0.0078	0.1170	0.2450	0.0239
rs16981340	0.0035	0.6051	0.0014	0.4701	0.0079	0.1974
rs17174528	0.0029	0.5682	0.0134	0.0647	0.0043	0.0157
rs17250121	0.0029	0.5626	0.0108	0.0924	0.0149	0.0241
rs17269396	0.4800	0.0937	0.3402	0.0000	0.0923	0.0000
rs17269816	0.0066	0.1615	0.0036	0.6206	0.0301	0.0584
rs17269928	0.0039	0.1992	0.0001	0.3170	0.0506	0.0284
rs17276338	0.0612	0.0003	0.0004	0.3383	0.0173	0.0092
rs17276345	0.0146	0.0613	0.0049	0.1874	0.6960	0.7011
rs17276358	0.0022	0.5462	0.0008	0.3715	0.0566	0.1356
rs17276777	0.0007	0.4847	0.0147	0.4017	0.6466	1.0000
rs17316007	0.0126	0.0628	0.0019	0.4497	0.0016	0.0115
rs17316543	0.0134	0.0596	0.0049	1.0000	0.0027	0.0592
rs17316592	0.0049	0.9204	0.0040	0.6891	0.1439	0.6118
rs17323322	0.0204	0.0267	0.0048	0.8910	0.0596	0.0174
rs17842387	0.0043	0.7624	0.0030	0.2459	0.0067	0.1407

续表3

位点	华东-西北		华东-华南		西北-华南	
	F_{st} 值	P 值	F_{st} 值	P 值	F_{st} 值	P 值
rs2032602	0.0049	0.6944	0.0043	1.0000	0.0043	1.0000
rs2032624	0.0151	0.0591	0.0000	1.0000	0.0149	0.1238
rs2032631	0.0594	0.0002	0.0014	0.3807	0.0923	0.0000
rs2032645	0.0048	0.4980	0.0095	0.1019	0.0301	0.0150
rs2032650	0.0127	0.1053	0.0098	0.0959	0.0506	0.0012
rs2032652	0.0043	0.7697	0.0108	0.0912	0.0173	0.0462
rs2032665	0.5090	0.0000	0.1166	0.0000	0.6960	0.0000
rs2032674	0.0098	0.0943	0.0129	0.0661	0.0566	0.0005
rs2032678	0.5960	0.0679	0.0004	0.3131	0.6466	0.4891
rs2075181	0.0047	0.8922	0.0042	0.1937	0.0016	0.3077
rs2075640	0.0023	0.5298	0.0049	0.1861	0.0027	0.6016
rs2196155	0.0539	0.0006	0.0229	0.0217	0.1439	0.0000
rs2267801	0.0283	0.0138	0.0008	0.3060	0.0596	0.0001
rs35284970	0.0048	0.1745	0.0049	1.0000	0.0067	0.1659
rs3848982	0.0049	1.0000	0.0043	1.0000	0.0043	1.0000
rs3853054	0.0165	0.0419	0.0143	0.0611	0.0703	0.0001
rs3898	0.0013	0.4480	0.0045	1.0000	0.0016	0.6857
rs3900	0.0019	0.4684	0.0108	0.0918	0.0266	0.0161
rs3906	0.0018	0.6850	0.0043	1.0000	0.0040	1.0000
rs4589047	0.0007	0.4492	0.0108	0.0908	0.0300	0.0114
rs52812045	0.0049	1.0000	0.0772	0.0001	0.0796	0.0001
rs7067458	0.0267	0.0122	0.0385	0.0029	0.0039	0.7647
rs917759	0.0002	0.3449	0.0097	0.0905	0.0313	0.0081
rs9306841	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
rs9306845	0.0023	0.2884	0.0153	0.0539	0.0441	0.0022
rs9306848	0.0006	0.3803	0.0030	0.2440	0.0183	0.0409
rs9341278	0.0040	0.1211	0.0051	0.4420	0.0274	0.0202
rs9785908	0.0043	0.7674	0.0049	0.1892	0.0103	0.1047
rs9786095	0.0046	0.8756	0.0030	0.2439	0.0001	0.3945
rs9786247	0.0040	0.7528	0.0083	0.1239	0.0024	0.2884
rs9786394	0.1156	0.0000	0.2318	0.0000	0.0268	0.0113
rs9786401	0.0005	0.3674	0.0049	1.0000	0.0009	0.3564
rs9786479	0.0043	0.7484	0.0043	1.0000	0.0025	0.7243
rs9786502	0.0024	0.5125	0.0092	0.1026	0.0002	0.3920
rs9786707	0.0043	0.7671	0.0108	0.0900	0.0173	0.0495
SRY8299	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000

注:粗体表示 $P < 0.05$ 或 $F_{st} > 0.05$

表3数据显示:华东与西北汉族无关男性群体之间有7个位点存在中等程度以上的分化,华东与华南之间只有4个位点存在中等程度以上的分化,西北与华南之间有15个位点存在中等程度以上的分化。

71个位点中,分别有11个位点在华东与西北汉族群体的差异具有统计学意义($P < 0.05$)、8个位点在华东与华南汉族群体的差异具有统计学意义($P < 0.05$)、36个位点在华南与西北汉族群体的差异具有

统计学意义($P < 0.05$)。

2.3 筛选在三个地域汉族群体均具有应用价值的位点

结合本实验结果与前期研究取得的华南、华东汉族群体的遗传学数据^[7,11],通过评价GD值,筛选出在西北、华南和华东群体具有较高信息量的Y-SNP位点。统计结果表明,在西北、华东和华南3个地域群体均具有应用价值的Y-SNP位点共有26个(表4)。

表4 在西北、华东和华南汉族群体中具有中、高度信息量的位点

位点	适用于西北群体的位点	适用于华南群体的位点	适用于华东群体的位点	同时适用于西北、华南和华东群体的位点
M117	√	√√	√	√
M119	√	√	√	√
M122	√√	√√	√√	√
M134	√√	√√	√√	√
P31	-	√	-	-
rs11096432	√√	√	√√	√
rs11096433	-	√	√	-
rs13447354	√	√	√	√
rs13447361	√√	√√	√√	√
rs16980363	√√	-	√	-
rs16980391	√	-	√	-
rs16980396	√	-	√	-
rs16980426	√√	-	√	-
rs16980601	√√	√	√√	√
rs16980610	√√	√	√√	√
rs16980641	√√	√	√	√
rs16981290	√√	√	√√	√
rs16981340	-	-	-	-
rs17174528	√	-	√	-
rs17250121	√	-	√	-
rs17269396	√√	√√	√√	√
rs17269816	√√	√√	√√	√
rs17269928	√√	√√	√√	√
rs17286338	√√	√	√	√
rs17276345	-	-	√	-
rs17276358	√√	√√	√√	√
rs17276777	√	-	√	-
rs17316007	√√	√√	√√	√
rs17316543	√√	√	√	√
rs17316592	√√	√√	√√	√
rs17323322	√√	√	√	√
rs17842387	√	-	√	-
rs2032631	√	-	-	-
rs2032650	-	√	-	-
rs2032652	√	-	√	-
rs2032665	-	-	√√	-
rs2032674	√√	-	√	-
rs2032678	√	√	√√	√
rs2075181	√	√	√	√
rs2075640	-	-	√	-
rs2196155	√√	-	√	-
rs2267801	√√	√√	√√	√
rs35284970	√	-	-	-
rs3853054	√√	-	√	-
rs3900	√	-	√	-
rs4589047	√	-	√	-
rs52812045	√	√√	√	√
rs7067458	√	√	-	-
rs917759	√√	√	√√	√
rs9306845	√	-	√	-

续表4

位点	适用于西北群体的位点	适用于华南群体的位点	适用于华东群体的位点	同时适用于西北、华南和华东群体的位点
rs9306848	√	-	√	-
rs9785908	√	-	√	-
rs9786095	-	-	√	-
rs9786247	-	-	√	-
rs9786394	√√	√√	-	-
rs9786401	-	-	-	-
rs9786502	√	√	√√	√
rs9786707	√	-	√	-

注：“√”表示中度信息量的位点(0.2≤GD<0.3)；“√√”表示具有高度信息量的位点(GD≥0.3)；“-”表示在该地区不具有适度信息量的位点

3 讨论

3.1 多态性分析

观察本实验结果,分析检测的71个Y-SNP位点在西北汉族无关男性人群的多态性及等位基因分布频率,可知有67个Y-SNP在西北汉族男性人群中呈多态性分布,GD值在0.010 0~0.502 2,中度信息量(0.2≤GD<0.3)的位点有22个,高度信息量(GD≥0.3)的位点有25个。本课题组前期研究结果^[7,11]显示,华南和华东汉族人群中分别有66、67个位点存在遗传多样性,中度信息量的分别有18、31个,高度信息量的分别有13、18个,整合本研究结果可知,同时适用于西北、华南和华东群体的中高度信息量的位点有26个。他们在亲权鉴定和个体识别中具有一定的应用价值,可以作为STR、Y-STR检测的补充。

3.2 群体分布差异分析

经比较分析P值,发现华南和西北汉族群体之间的差异最大,华东与西北汉族群体之间的差异次之,华东与华南汉族群体之间的差异最小。分析比较F_{st}值发现,西北与华南之间分化的程度最高,华东与西北汉族无关男性群体之间分化程度居中,华东与华南之间分化程度最低。由此可见,Y染色体上SNP标记的等位基因分布具有一定的地域差异,用于法医学鉴定时,需要使用相应区域的群体遗传学数据。

3.3 法医学应用价值评估

目前,SNP在法医学中的研究主要有两个方面:(1)对SNP位点进行筛选,获得不同人群的分布频率,建立数据库,用于个体识别和亲权鉴定;(2)对SNP位点的功能进行分析,通过对现场遗留的生物检材的部分位点进行分型,提供嫌疑人的一些特征,如毛发颜色、眼睛颜色等^[12]。

本研究对西北汉族男性无关个体的71个Y-SNP位点进行分析,共检测出170种单倍型,其中频率最高的为0.034 8(7/202),频率最低的只有0.005 0

(1/202)。HD为0.993 0,得到了较大的群体遗传信息,可应用于涉及男性的个体识别和亲权鉴定的案例中。例如,在进行父系关系检验的时候,某些情况下仅仅凭借Y-STR分型的结果难以出具鉴定意见,不能因为几个STR的分型结果不同而否定父系关系^[13],针对这种案例,对高信息量的Y-SNP位点进行分型检测可以视为一种补充鉴定的手段。

同时,本实验也比对了华南、华东地区汉族男性Y-SNP位点的分布特点,筛选出26个在西北、华南和华东汉族具有普适性的位点,这将有助于完善法医DNA检验手段。

参考文献:

- [1] 罗少堂,胡利平,聂胜洁,等. SNPs在个体识别与表型预测中的研究进展[J]. 中国法医学杂志,2014,29(6): 537-540.
- [2] BOUAKAZE C, KEYSER C, AMORY S, et al. First successful assay of Y-SNP typing by SNaP-shot minisequencing on ancient DNA[J]. Int J Legal Med,2007,121(6):493-499.
- [3] BUTLER J M. 法医DNA分型专论:方法学[M].3版. 侯一平,李成涛,译. 上海:科学出版社,2013:294.
- [4] International Hapmap Project[DB/OL]. (2005-10-26) [2015-4-20]. www.hapmap.org.
- [5] The National Center of Biotechnology Information[DB/OL]. (1992-10-01) [2016-09-15]. http://www.ncbi.nlm.nih.gov/pubmed.
- [6] Agena Bioscience's online design tools[DB/OL]. (2007-08-24) [2016-08-20]. http://www.mysequenom.com/Home.
- [7] 宋雨桐,李莉,张立男,等. MALDI-TOF-MS对Y染色体上66个二等位基因的检测[J]. 法医学杂志,2017,33(3):239-243.
- [8] 侯一平. 法医物证学[M].4版. 北京:人民卫生出版社,2016:92.
- [9] EXCOFFIER L, LICHER H E. 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. (下转第38页)

法确定。

综上所述,近10年法庭科学领域文献数量呈逐年递增趋势,2017年相比2016年中国大陆文献数量增速高于全球和排名前5位的国家;文献类型以论文为主,中国大陆在综述、会议录文献、社论等方面仍有待发力;文献的期刊来源与2017年JCR中所列期刊不一致;2017年法庭科学领域JCR中,FSIG的影响因子最高,FSIG、SJ和FSMP的影响因子较本研究纳入的最早影响因子增长较大,IJLM和LM连续两年呈下降趋势。科研成果在不同国家科研机构中的分布,揭示了各个国家法庭科学领域科研实力的分布状态,文献数量排名前20位的机构中,6个国家的科研集中在少数几家机构,60%的国家科研分布于多家机构,中国大陆属于后者;病理学专业的文献最多,但在高引文献中仅有1篇,中国大陆文献集中于病理学、药理学和药剂学、毒物学、遗传学和遗传性领域;所有文献中仅有28.17%得到基金机构资助,国家自然科学基金委员会资助发表的文献排在全球第1位。

参考文献:

[1] 宁笔. JCR与SCI、SCIE期刊的区别[EB/OL]. (2010-09-10) [2018-08-22]. <http://blog.sciencenet.cn/home.php?mod=space&uid=408109&do=blog&id=361700>.

[2] 文献类型说明[EB/OL]. (2018-01-31)[2018-07-14]. http://images.webofknowledge.com/images/help/zh_CN/WOS/hs_document_type.html#dsy2200-TRS_doc_type

_descriptions.

[3] 研究方向(类别/分类)[EB/OL]. (2018-01-31)[2018-07-19]. http://images.webofknowledge.com/images/help/zh_CN/WOS/hp_research_areas_easca.html.

[4] 中国合格评定国家认可委员会. 司法鉴定/法庭科学机构认可领域分类:CNAS—AL13[S].2015.

[5] 国家自然科学基金委员会. 科学基金共享服务网(科技成果信息系统)[EB/OL]. (2018-07-28)[2018-08-02]. <http://output.nsf.gov.cn/>.

[6] 中华人民共和国国务院. 国家中长期科学和技术发展规划纲要(2006—2020年)[EB/OL]. [2018-07-21]. <http://www.most.gov.cn/kjgh/kjghzqc/>.

[7] 国家自然科学基金委员会. 国家自然科学基金“十三五”发展规划[EB/OL]. (2015-09-17)[2018-07-19]. http://www.nsf.gov.cn/publish/portal0/xxgk/043/info72249.htm#_Toc442082163.

[8] 朱大明. 综述类论文的学术价值及其鉴审[J]. 科技与出版,2007(4):37-38.

[9] ANALYTICS C. InCites Journal Citation Reports[EB/OL]. (2018-09-20)[2018-11-12]. <http://jcr.incites.thomsonreuters.com/JCRJournalProfileAction.action?pg=JRNLPF&journalImpactFactor=n%20Fa&year=2017&journalTitle=MED%20LAW%20REV&edition=SCIE&journal=MED%20LAW%20REV>.

[10] 于笑天,梁璐,黄平,等. 法医学领域的SCI收录期刊与国内期刊比较[J]. 法医学杂志,2017,33(4):411-415.

(收稿日期:2018-07-30)

(本文编辑:张建华)

(上接第29页)

[10] WRIGHT. Evolution and the genetics of populations[M]. Chicago: University of Chicago Press, 1978:157.

[11] LI L, SONG Y, LIU Y, et al. Genetic diversity of 71 Y-chromosomal biallelic markers in the Han population living in Southern China[J]. Forensic Sci Int Genet Suppl Ser,2017,6:e180-e181.

[12] 孙元田,杨波,区彩莹,等. Y-SNP多态性揭示的海南岛黎族、回族和仡隆人群的起源[J]. 科技导报,2007,25(24):44-47.

[13] 李莉,柳燕,林源,等. 从实际案例探讨父系关系鉴定中单独检验Y-STR标记的局限性[J]. 中国司法鉴定,2015(3):112-113.

(收稿日期:2018-04-19)

(本文编辑:张素华)