

中国海南黎族人群 15 个 STR 位点的遗传多态性

邓文国^{1,2}, 何蕴韶², 李 钢³, 金仁晶², 江梅华², 王一鸣¹, 曾瑞萍¹

(1. 中山大学医学遗传学教研室, 广东 广州 510080; 2. 中山大学达安基因诊断中心, 广东 广州 510089;
3. 海口市人民医院, 海南 海口 400038)

摘 要:【目的】研究海南黎族人群 15 个 STR 位点 D3S1358、TH01、D21S11、D18S51、VWA、CSF1PO、D8S1179、TPOX、FGA、D5S818、D13S317、D7S820、D16S539、D19S433、D2S1338 的遗传多态性。【方法】通过人类短串联重复序列(short tandem repeat, STR)复合扩增、基因扫描、基因分型调查了 110 名黎族无关个体 15 个 STR 位点等位基因分布情况。【结果】15 个 STR 位点均符合 Hardy-Weinberg 平衡, 共检出 149 个 STR 等位基因, 其频率分布在 0.0045~0.5045 之间, 杂合度(heterozygosity, H)为 0.6716~0.8754, 个体识别力(discrimination power, DP)为 0.8770~0.9673, 非父排除率(probabilities of paternity exclusion, EPP)为 0.4383~0.7396, 多态信息含量(polymorphic information content, PIC)为 0.6265~0.8574。【结论】选用的 15 个 STR 位点均能检出等位基因遗传多态性, 并具有较丰富的信息含量, 为进一步研究中华民族 STR 遗传结构提供了基础资料。

关键词:人类短串联重复序列; 遗传多态性; 基因扫描

中图分类号: Q987

文献标识码: A

文章编号: 1672-3554(2005)06-0651-03

Polymorphism of Fifteen Short Tandem Repeat Loci in Hainan Li Ethnic Population

DENG Wen-guo^{1,2}, HE Yun-shao², LI Gang³, JIN Ren-jing², JIANG Mei-hua², WANG Yi-ming¹, ZENG Rui-ping¹

(1. Department of Medical Genetics, 2. Da-an Genetic Center, SUN Yat-sen University, Guangzhou 510089, China;
3. Haikou Municipal People's Hospital, Haikou 400038, China)

Abstract: 【Objective】To investigate genetic polymorphism of fifteen short tandem repeat (STR) loci (D3S1358, TH01, D21S11, D18S51, VWA, CSF1PO, D8S1179, TPOX, FGA, D5S818, D13S317, D7S820, D16S539, D19S433, D2S1338) in Hainan Li ethnic population. 【Methods】The allelic distributions of fifteen STR loci from 110 unrelated individuals were examined through coamplification, gene scan and genotyping. 【Results】All the 15 loci met Hardy-Weinberg equilibrium. There were 149 STR alleles in Li ethnic group, with allelic frequencies ranging from 0.0045 to 0.5045, heterozygosity (H) from 0.6716 to 0.8754, discrimination power (DP) from 0.8770 to 0.9673, probabilities of paternity exclusion (EPP) from 0.4383 to 0.7396, and polymorphic information content (PIC) from 0.6265 to 0.8574. 【Conclusion】Our results provides a basis for studying the genetic structure of Chinese ethnic groups, the 15 STR loci selected were suitable to detect polymorphisms, having good enough information contents.

Key words: short tandem repeat; genetic polymorphism; gene scan

[J SUN Yat-sen Univ(Med Sci), 2005, 26(6): 651-653]

人类短串联重复序列(short tandem repeat, STR)是存在于人类基因组中的一类具有高度多态性的遗传学标记, 因此人类短串联重复序列已广泛应用于人类遗传学、法医学等许多研究领域。我们应用 STR 复合扩增技术对 15 个 STR 位点进行扩增, 其产物经 ABI 3100 型遗传分析仪检测, 对 110 名海

口黎族无关个体进行基因频率分布调查, 为人类遗传学研究提供有关数据。

1 材料和方法

1.1 材 料

收稿日期: 2005-08-01

基金项目: 国家博士点科研基金资助项目(200044)

作者简介: 邓文国 (1968-), 男, 湖南郴州人, 在职博士生, 助理研究员. E-mail: dengwg418@163.com

标本来源:采自海南黎族 110 名无关个体外周血, EDTA 抗凝。主要试剂、仪器 和软件: 五色荧光标记的 15 个 STR 位点复合扩增试剂 AmpFISTR Identifiler™ 试剂盒、GeneScan - 500 内标、ABI 3100 型遗传分析仪、9700 型 PCR 扩增仪、数据收集软件 (Data Collection Software v1.0.1)、基因扫描软件 (GeneScan® Software v3.7) 和基因分型软件 (GeneTyper® Software v3.7) 均为美国 ABI 公司产品。

1.2 方法

1.2.1 DNA 模板的制备 采用饱和酚/氯仿抽提法或 Chelex-100 快速 DNA 抽提法^[1]。

1.2.2 PCR 扩增 15 个 STR 基因位点的特征见表 1。反应总体积为 10 μ L, 其反应成分为: PCR 反应混合液 4 μ L、引物混合液 2.5 μ L、Taq Gold DNA 聚合酶 1.0 U、模板 DNA 20 ng。扩增条件为: 95 $^{\circ}$ C 预变性 10 min, 94 $^{\circ}$ C 30 s, 60 $^{\circ}$ C 30 s, 70 $^{\circ}$ C 45 s, 循环 10 次; 再 94 $^{\circ}$ C 30 s, 60 $^{\circ}$ C 30 s, 70 $^{\circ}$ C 45 s, 进行 20 个循环, 最后 60 $^{\circ}$ C 延伸 30 min。

表 1 15 个 STR 基因位点的特性

Table 1 Characteristics of STR loci

Loci	Chromosome site	Repeated sequence	Polymorphic fragment length (bp)	Alleles
D8S1179	8	TCTR	128-172	8-19
D21S11	21q11.2-q21	TCTA	187-243	24-38
D7S820	7q11.21-22	GATA	258-294	8-14
CSF1PO	5q33.3-34	AGTA	281-317	7-14
D3S1358	3p	TCTA	114-142	13-18
TH01	11p15.5	AATG	168-189	5-10
D13S317	13q22-31	GATA	206-234	8-14
D16S539	16q24-qter	AGAT	234-274	8-15
D2S1338	2q35-37.1	TGCC	289-341	15-28
D19S433	19q12-13.1	AAGG	106-140	9-18
vWA	12p12-pter	TCTA	157-197	14-20
TPOX	2p23-pter	AACT	218-242	8-12
D18S51	18q21.3	AGAA	265-345	9-25
D5S818	5q21-31	AGAT	135-171	7-14
FGA	4q28	CTTT	219-267	18-28

1.2.3 扩增产物的分离与检测 取 1.5 μ L PCR 扩增产物, 0.3 μ L GeneScan500 内标, 24 μ L 甲酰胺于 0.5 mL 的离心管中, 95 $^{\circ}$ C 变性 3 min 后迅速放入冰水中冰浴, 上 ABI 3100 型基因分析仪进行分离并检测, 用 POP4 液态胶毛细管电泳, 电压 15 kV, 温

度 60 $^{\circ}$ C。用 GeneScan 分析软件进行自动基因型分型。

1.3 统计学处理

用 χ^2 检验判断各 STR 位点的基因型分布是否符合 Hardy-Weinberg 平衡定律, 按 Jamieson 等^[2,3] 人的方法计算期望杂合度 (H)、多态信息含量 (PIC)、个体识别力 (DP) 和非父排除率 (EPP)。

2 结果

2.1 海南黎族 STR 等位基因频率

15 个 STR 基因位点共检出等位基因 149 个, 其中 D21S11 位点 16 个, FGA 位点 15 个, D18S51 位点 14 个, D19S433、D2S1338 位点各 12 个, D8S1179 位点 10 个, vWA 位点、D5S818 位点各 9 个, CSF1PO、D7S820 和 D13S317 位点各 8 个, D16S539、D3S1358、TPOX 和 TH01 各 7 个, 其频率分布在 0.0045~0.5045 之间。

2.2 海南黎族群体遗传学

对海南黎族 15 个 STR 位点基因型的观察值和期望值进行 χ^2 检验, 各基因位点的基因型分布符合 Hardy-Weinberg 平衡定律 ($P > 0.05$)。根据实验所得各民族 STR 等位基因及基因型频率, 按文献方法, 分别计算出各基因位点在海南黎族中的杂合度 (H)、偶合率 (Pm)、个体识别力 (DP)、非父排除率 (EPP) 及多态信息量 (PIC), 结果见表 2。

3 讨论

STR 位点具有高度信息度的特点。DP、EPP、H 和 PIC 是衡量一个遗传标记是否具有高度信息度的主要指标。一般来说, 当多态信息量 (PIC) > 0.50 时, 标记具有高度的可提供信息性; $0.50 > \text{PIC} > 0.25$, 标记能够较合理的提供信息; 当 $\text{PIC} < 0.25$, 标记可提供的信息性较差^[5]。15 个 STR 位点杂合度在 0.6716~0.8754 之间, PIC 大于 0.6265, 累计 PIC 为 0.999999, 适用于基因扫描, 遗传连锁分析等研究。

黎族是我国海南岛主要的少数民族, 历史悠久。由于地域的缘故, 黎族人处于与外界相对隔离的状态。我们对 110 名黎族无关个体的 15 个 STR 位点的多态性分布进行检验, 结果符合 Hardy-Weinberg 平衡定律 ($P > 0.05$), 检出的等位基因数最多的为 16 个, 最少的为 7 个。在 15 个 STR 位点

表2 Hardy-Weinberg 平衡检验及遗传统计学指标
Table 2 Hardy-Weinberg test and statistical indexes of polymorphism

Locus	H	P _{nt}	DP	EPP	PIC	df	χ^2	P
CSFIPO	0.7394	0.1122	0.8878	0.4999	0.6910	7	8.7696	>0.25
D5S818	0.7599	0.0893	0.9107	0.5391	0.7188	11	18.7880	>0.05
D13S317	0.8227	0.0715	0.9285	0.6373	0.7933	13	15.3299	>0.25
D7S820	0.7873	0.0504	0.9496	0.5771	0.8036	20	14.3124	>0.75
TPOX	0.6716	0.0664	0.9336	0.4383	0.6265	6	12.0991	>0.05
vWA	0.8091	0.0656	0.9344	0.6192	0.8165	13	19.7500	>0.10
FGA	0.8555	0.0440	0.9560	0.7031	0.8343	17	20.1831	>0.25
TH01	0.7778	0.0879	0.9121	0.5611	0.7392	9	16.6401	>0.05
D3S1358	0.7253	0.1230	0.8770	0.4734	0.6715	7	5.8194	>0.50
D8S1179	0.8293	0.0504	0.9496	0.6571	0.8036	20	14.3124	>0.75
D21S11	0.7615	0.0883	0.9117	0.5584	0.7268	14	17.4327	>0.10
D19S433	0.8138	0.0661	0.9339	0.6301	0.7848	23	18.4458	>0.50
D2S1338	0.8754	0.0327	0.9673	0.7396	0.8574	23	25.2453	>0.25
D16S539	0.7817	0.0833	0.9167	0.5701	0.7446	11	12.3223	>0.25
D18S51	0.8268	0.0524	0.9476	0.6587	0.8024	18	19.6344	>0.25

中,黎族群体中最常见的等位基因是 TPOX*8,频率是 0.5045,美国高加索人等位基因最高频率也是 TPOX*8,频率是 0.5330,美国黑人最常见的等位基因是 D13S317*12,频率是 0.4622,而广东汉族人和中国独龙族最高频率的等位基因分别是 TH01*9 和 D16S539*9^[6]。15 个位点的累积非父排除率为 0.9999 以上,累积个体识别率为 0.9999 以上,累积多态信息含量为 0.9985,符合人类高识别能力 STR 位点的选择标准。我国民族众多,已有关于少数民族遗传多态性分析的报道^[6-9]为我国建立不同民族 STR 基因数据库奠定了一定基础,海南黎族的调查检验结果为人类学分析、遗传学研究及亲权关系鉴定提供了宝贵的群体资料。

参考文献:

- [1] Walsh PS, Metzgar DA, Higuchi R, *et al.* Chelex-100 as a medium for simple extraction of DNA for PCR-based typing from forensic material[J]. *Bio Techniques*, 1991, 10(8):506-13.
- [2] Jamieson A, Taylor SC. Comparisons of three probability formulae for parentage exclusion[J]. *Anim Genet*, 1997, 28(6): 397-400.
- [3] Lareu MV, Phillips CP, Carracedo A, *et al.*

Investigation of the STR locus HUMTH01 using PCR and Caucasian population surveys and usefulness in paternity investigations[J]. *Forensic Sci Int*, 1994, 66(1): 41-52.

- [4] 杜志淳,李莉,林源,等.中国“罪犯 DNA 数据库”STR 基因座研究[J].*中国法医学杂志*,2000, 15(2): 65-8.
- [5] Gill P. A new method of STR interpretation using inferential logic-development of a criminal intelligence database[J].*Int J Leg Med*, 1996, 109(1):14-22.
- [6] 于亮,黄小琴,史荔,等.中国四个少数民族九个 Y STR 位点基因频率和单倍型研究[J].*中华医学遗传学杂志*,2005, 22(3):337-40.
- [7] 何吉,许先国,傅启华,等.浙江畲族和汉族人群六个短串联重复序列位点遗传多态性的研究[J].*中华医学遗传学杂志*,2003, 20(3):250-2
- [8] 郭辰虹,王玉炯,舒畅,等.宁夏地区回族和汉族人群 4 个 STR 位点的遗传多态性分析[J].*山东大学学报(医学版)*,2005, 43(3):192-4
- [9] 陈雪玲,黄辰,袁育康,等.D16S539,D7S820 和 D13S317 位点在新疆哈萨克族中的遗传多态性[J].*中华医学遗传学杂志*,2002, 19(1):55-7.

(编辑 黄小延)