

陕西汉族人群 2 号和 11 号染色体 15 个 STR 基因座的遗传多态性

史晓薇^{1,2}, 任峰玲¹, 耿东¹, 郭雄¹

(西安交通大学医学院 1. 公共卫生学系; 2. 第一附属医院妇幼中心, 西安 710061)

[摘要] 目的: 分析陕西汉族人群 2 号和 11 号染色体上 15 个短串联重复序列 (short tandem repeat, STR) 位点的多态性。方法: 采用荧光标记基因扫描技术对 15 个 STR 基因座在 175 例无血缘关系的陕西汉族人中的遗传多态性进行分析。结果: *D2S335*, *D2S396*, *D2S338*, *D2S2382*, *D2S305*, *D2S151*, *D2S2368*, *D2S391*, *D11S912*, *D11S4090*, *D11S4147*, *D11S4190*, *D11S4149*, *D11S4126*, *D11S4094* 分别检出 11, 11, 11, 10, 8, 8, 9, 12, 7, 11, 8, 10, 5, 5, 6 个等位基因, 各位点等位基因型分布符合 Hardy-Weinberg 平衡, 杂和度 (heterozygosity, H) 为 0.4216 ~ 0.8517, 个体识别力 (power of discrimination, DP) 为 0.6568 ~ 0.9598, 多态信息量 (polymorphism information content, PIC) 为 0.4078 ~ 0.8366, 非父排除率 (probability of paternity exclusion, EPP) 为 0.3135 ~ 0.8537。结论: 15 个 STR 基因座在陕西汉族人群中具有较好的多态性, 表明了陕西汉族人群 15 个 STR 基因座结构特征, 为人类学、法医学研究提供了数据。

[关键词] 短串联重复序列; 遗传多态性; 基因扫描; Hardy-Weinberg 平衡

[中图分类号] Q319.1 **[文献标识码]** A **[文章编号]** 1672-7347(2008)07-0587-05

Genetic polymorphism of 15 STR on chromosome 2 and 11 in Shaanxi Han people in China

SHI Xiao-wei^{1,2}, REN Feng-ling¹, GENG Dong¹, GUO Xiong¹

(1. Department of Public Health; 2. Center of Maternal and Child Health Care, First Affiliated Hospital, Medical School of Xi'an Jiaotong University, Xi'an 710061, China)

Abstract: **Objective** To investigate the genetic polymorphism of 15 short tandem repeat (STR) loci on chromosome 2 and chromosome 11 in Shaanxi Han people in China. **Methods** Fluorescence-based gene scan technique was used to examine the genetic polymorphism of 15 STR loci in 175 unrelated individuals from Chinese Han population in Shaanxi province. **Results** The number of alleles *D2S335*, *D2S396*, *D2S338*, *D2S2382*, *D2S305*, *D2S151*, *D2S2368*, *D2S391*, *D11S912*, *D11S4090*, *D11S4147*, *D11S4190*, *D11S4149*, *D11S4126*, and *D11S4094* was 11, 11, 11, 10, 8, 8, 9, 12, 7, 11, 8, 10, 5, 5, and 6. The distribution of allele frequencies of the 15 STR was consistent with Hardy-Weinberg equilibrium ($P > 0.05$). Heterozygosity (H) value was 0.4216 ~ 0.8517, the average power of discrimination (DP) was 0.6568 ~ 0.9598, polymorphism information content (PIC) was 0.4078 ~ 0.8366, and probability of paternity exclusion (EPP) was 0.3135 ~ 0.8537. **Conclusion** The 15 STR loci have relatively high genetic

收稿日期 (Date of reception) 2007-10-11

作者简介 (Biography) 史晓薇 (1971-), 女, 陕西绥德人, 硕士, 主治医师, 主要从事儿童心理行为和骨关节分子机制的研究。

通讯作者 (Corresponding author) 郭雄, E-mail: guox@mail.xjtu.edu.cn

基金项目 (Foundation items) 国家自然科学基金 (30731252, 30630058). This work was supported by National Natural Science Foundation of P. R. China (30731252, 30630058).

polymorphism in Shaanxi Han population, which provides the genetic structure of Chinese Han groups, and is also useful in anthropology and forensic science.

Key words: short tandem repeat; genetic polymorphism; genescan; Hardy-Weinberg equilibrium

[*J Cent South Univ (Med Sci)*, 2008, 33(7):0587-05]

短串联重复序列(short tandem repeat, STR)是以2~6 bp为核心单位的串联重复序列,广泛分布于人类基因组中。具有高度多态性和遗传稳定性,所包含的遗传信息是以往任何遗传性标记所无法比拟的,被广泛应用于遗传图的构建、疾病相关基因的定位、克隆、人类学以及法医学等领域,是目前应用最广泛、研究较深入的遗传标志之一^[1-2]。

由于基因座的等位基因频率、杂合度等遗传参数因群体而异,目前STR的群体遗传数据大多数来自欧美等地的白种人群的研究,因此有必要对中国各民族的STR多态性分布进行调查,建立中华民族STR基因的数据库。2号染色体与高血压、冠状动脉硬化性心脏病等心血管疾病和多种骨关节疾病有关;多种恶性肿瘤的疾病相关基因已定位于11号染色体上,(<http://www.ncbi.nlm.nih.gov>)。因此,本调查选择了2号染色体上8个STR(*D2S151*, *D2S2382*, *D2S391*, *D2S2368*, *D2S335*, *D2S396*, *D2S338*, *D2S305*)和11号染色体上的7个STR(*D11S912*, *D11S4090*, *D11S4147*, *D11S4190*, *D11S4149*, *D11S4126*, *D11S4094*)基因座,对其遗传数据进行分析,旨在为中国人疾病连锁分析和间接基因诊断提供科学依据。

1 材料与方 法

1.1 样本 在陕西麟游县、永寿县的和咸阳地区选择175例相互无血缘关系的健康个体作为本次研究对象,年龄为(25.10±1.23)岁,在知情自愿的原则下,各抽取静脉血2 mL,EDTA抗凝备用。

1.2 方 法

1.2.1 DNA提取 采用美国Promega公司生产的外周血基因组DNA提取试剂盒提取基因组DNA,操作按试剂盒说明(<http://www.promega.com>),经定量后稀释为50 mg/L备用。

1.2.2 多态基因座及引物 所选的15个STR基因座来源于美国人类连锁合作中心(CHLC)研制的人类基因组扫描引物组。15个基

因座的扩增引物及试剂均购自美国ABI公司(<http://www.appliedbiosystems.com>)。

1.2.3 PCR扩增 反应体系为15:1的引物混合液,9 μL PCR混合液(包括PCR缓冲液,2.5 mmol/dNTP混合液,Tap聚合酶5 U/μL,25 mmol/L的MgCl₂),1.2 μL的DNA和3.8 μL的双蒸水。反应条件为:95℃预变性12 min,然后进行94℃变性15 s,55℃15 s,72℃退火30 s,循环10个周期;89℃退火15 s,55℃延伸15 s,72℃延伸30 s,循环10个周期;最后72℃延伸10 min,4℃冰箱避光保存待用。PCR扩增在PE9600下完成。

1.2.4 扩增产物电泳及基因分析 每个样本的PCR扩增产物按照荧光染料FAM(蓝色荧光):VIC(绿色荧光):NED(黄色荧光)(荧光信息见表1)以2:2:4的比例混匀后,取0.5 μL与9 μL的去离子甲酰胺和0.5 μL的GS-500Liz混匀,95℃变性3 min迅速冷却,检查无气泡后放入ABI 3100遗传分析仪进行毛细管电泳,用GeneScan Analysis 3.7和Genotyper3.7软件进行检测分型。

1.3 统计学处理 用Excell 2003完成数据库的建立,用直接计数法计算每个基因座的等位基因频率和基因型频率。采用χ²检验方法进行Hardy-Weinberg平衡检验,采用文献[3-6]的方法计算各位点杂和度(heterozygosity value, H)、个体识别力(average power of discrimination, DP)、非父排除率(probability of paternity exclusion, EPP)^[3-4]、多态信息量(polymorphism information content, PIC)^[5]和随机个体相同表型偶合率(matching probability, PM)^[6]等多态信息含量。

2 结 果

2.1 各位点等位基因频率和基因型频率数据

D2S335, *D2S396*, *D2S338*, *D2S2382*, *D2S305*, *D2S151*, *D2S2368*, *D2S391*, *D11S912*, *D11S4090*, *D11S4147*, *D11S4190*, *D11S4149*, *D11S4126*, *D11S4094* 分别检出11, 11, 11, 10, 8, 8, 9, 12, 7, 11, 8, 10, 5, 5, 6个等位基因,各座位等位基因片段大

小及频率分布见表 1。15 个 STR 基因座不同基因型的观察值和期望值差异无统计学意义,均符合 Hardy-Weinberg 平衡 ($P > 0.05$, 表 2)。

2.2 15 个 STR 基因座在陕西汉族人群中的群体遗传学指标 从表 2 可见,15 个 STR 基因座

中 $D2S151$, $D2S391$, $D2S2368$, $D2S335$, $D2S396$, $D2S338$, $D2S305$, $D11S4090$, $D11S4147$, $D11S4190$, $D11S4094$ 基因座杂和度 H , 个体识别力 DP , 多态信息量 PIC , 非父排除率 EPP 都比较高 ($H > 0.7$, $DP > 0.8$, $PIC > 0.6$, $EPP > 0.5$)。

表 1 陕西汉族人群 15 个 STR 基因座等位基因频率 ($n = 175$)

Tab. 1 Allele frequencies of the 15 STR loci in Han population of Shaanxi ($n = 175$)

等位基因 Allele	$D2S335^3$		$D2S396^3$		$D2S338^3$		$D2S2382^2$		$D11S4190^2$	
	片段大小 (bp)	频率 Frequency								
1	185	0.023	227	0.009	268	0.006	304	0.020	217	0.006
2	189	0.006	231	0.009	270	0.251	306	0.751	227	0.006
3	191	0.174	233	0.034	272	0.077	308	0.014	231	0.242
4	193	0.040	235	0.131	274	0.009	314	0.014	233	0.029
5	195	0.017	237	0.114	276	0.020	316	0.031	237	0.051
6	197	0.034	239	0.260	278	0.049	318	0.014	239	0.340
7	199	0.134	241	0.131	280	0.080	320	0.094	241	0.123
8	201	0.091	243	0.131	282	0.066	326	0.014	243	0.057
9	203	0.231	245	0.109	284	0.166	328	0.031	245	0.023
10	205	0.214	247	0.049	286	0.043	330	0.014	247	0.123
11	207	0.034	249	0.023	288	0.234	-	-	-	-
等位基因 Allele	$D2S151^2$		$D2S2368^3$		$D2S391^3$		$D11S4090^1$		$D11S4147^3$	
	片段大小 (bp)	频率 Frequency								
1	240	0.029	93	0.237	143	0.006	165	0.043	223	0.009
2	242	0.029	95	0.017	145	0.049	167	0.009	225	0.493
3	244	0.051	101	0.023	147	0.143	169	0.026	227	0.100
4	246	0.174	103	0.017	149	0.429	171	0.026	231	0.009
5	248	0.189	105	0.251	151	0.209	173	0.034	233	0.014
6	250	0.311	107	0.277	153	0.063	175	0.083	235	0.006
7	252	0.209	109	0.060	155	0.094	177	0.277	237	0.123
8	254	0.009	111	0.074	159	0.009	179	0.306	239	0.077
9	-	-	113	0.043	181	0.091	241	0.100	-	-
10	-	-	-	-	183	0.054	243	0.049	-	-
11	-	-	-	-	185	0.037	245	0.020	-	-
12	-	-	-	-	187	0.014	-	-	-	-
等位基因 Allele	$D11S912^2$		$D2S305^3$		$D11S4149^1$		$D11S4126^3$		$D11S4094^3$	
	片段大小 (bp)	频率 Frequency								
1	106	0.566	320	0.149	217	0.231	139	0.023	186	0.007
2	108	0.214	322	0.254	219	0.063	141	0.048	188	0.220
3	110	0.080	324	0.023	221	0.620	143	0.620	190	0.229
4	112	0.034	328	0.109	223	0.080	145	0.303	192	0.120
5	114	0.086	330	0.269	233	0.006	147	0.006	194	0.363
6	116	0.014	332	0.117	-	-	-	-	196	0.061
7	118	0.006	334	0.074	-	-	-	-	-	-
8	-	-	338	0.006	-	-	-	-	-	-

1: FAM-labeled(蓝色荧光); 2: VIC-labeled(绿色荧光); 3: NED-labeled(黄色荧光)。

表2 15个STR基因座在陕西汉族中的多态性指标
Tab.2 Polymorphism marks of 15 STR loci in Han population of Shaanxi

位点	H	PIC	DP	EPP	PM	P
D2S151	0.7856	0.7612	0.9248	0.7385	0.0703	0.6975
D2S2382	0.4216	0.4078	0.6568	0.3685	0.3483	0.9975
D2S2368	0.7925	0.7625	0.9266	0.7488	0.0731	0.1488
D2S391	0.7367	0.7064	0.8914	0.6568	0.0967	0.9735
D2S335	0.8392	0.8183	0.9511	0.8228	0.0668	0.9305
D2S396	0.8517	0.8366	0.9598	0.8515	0.0359	0.9937
D2S338	0.8327	0.8131	0.9549	0.8537	0.0471	0.9799
D2S305	0.8092	0.7833	0.9304	0.7588	0.0622	0.5703
D11S912	0.6195	0.5788	0.8060	0.5096	0.1855	0.6797
D11S4090	0.8057	0.7827	0.9384	0.7860	0.0607	0.8348
D11S4147	0.7106	0.6870	0.8857	0.6619	0.1071	0.3725
D11S4190	0.7886	0.7588	0.9196	0.7312	0.0766	0.5223
D11S4126	0.5207	0.4473	0.7073	0.3135	0.3031	0.9942
D11S4094	0.7497	0.7097	0.9030	0.6787	0.1027	0.1223
D11S4149	0.5525	0.5018	0.7557	0.4187	0.2509	0.1658

H:杂合度;PIC:多态信息含量;DP:个体识别力;EPP:非父排除率;PM:随机个体相同表型偶合率;P:Hardy-Weinberg平衡检验P值。

3 讨 论

多态位点是进行疾病基因连锁分析,遗传病间接诊断,个体识别等分析的重要工具,筛选高杂合度及高信息含量的STR对于疾病基因连锁分析,构建遗传图有十分重要的价值。随着人类基因组计划的实现,越来越多的疾病基因被定位,如高血压易感基因与D2S335连锁^[7],直肠癌和前列腺癌分别与11q23-qter处D11S912和2q23处D2S151有关^[8-9],D2S305和D2S4094与骨关节病连锁^[10-11]。因此,研究这些基因座在中国人群中的多态性分布,对进一步研究这些基因与疾病的关系有重要意义。

本研究结果显示,陕西汉族人群15个STR位点基因型分布符合Hardy-Weinberg平衡($P > 0.05$),该数据在群体遗传学上具有可靠的科学性。对遗传标记的多态性程度及其应用价值一般可用PIC,DP,H和EPP等指标来衡量,其中PIC直接反映出遗传标记所包含或所能提供的遗传信息,H能客观地反映群体的遗传变异水平;一般来说,当 $PIC > 0.50$ 时,标记具有高度的可提供信息性; $0.50 > PIC > 0.25$,标记能够较合理的提供信息;当 $PIC < 0.25$,标记可提供的

信息性较差^[12]。MP是法医学上作同一认定的依据,表示在随机人群中具有相同基因型的频率。DP和EPP反映该遗传标记在法医学个体识别及亲权鉴定中的能力,一般 $DP > 0.8$, $EPP > 0.5$ 时,属于高度多态性遗传标记,具有较高的应用价值^[13],本研究结果显示,这些遗传标记在陕西汉族中,平均杂合度大于0.7,平均多态信息含量在0.6以上,平均DP大于0.8,平均EPP也在0.5以上,说明实验所选择的这些遗传标记在进行民族群体遗传学研究和相关应用研究时是极有价值的^[2]。此结果不仅丰富了我国民族的STR基因的数据,而且将为遗传制图、疾病连锁分析、个体识别和亲权鉴定等的理论和应用提供基础数据。

致谢:感谢陕西省咸阳市防疫站和麟游县、永寿县防疫站协助采集血样,感谢西安交通大学医学院法医学系张洪波、刘青波博士给予的实验指导。

参考文献:

- [1] 曾朝阳,熊伟,熊芳,等. 9号染色体短臂上7个STR基因座在基因扫描中的信息表现[J]. 遗传,2003,25(5):543-548.
ZENG Zhao-yang, XIONG Wei, XIONG Fang, et al. Infor-

- mation behavior of 7 STR loci on chromosome 9P in genescan [J]. Hereditas (Beijing), 2003, 25(5): 543-548.
- [2] 赖江华,张保华,朱波峰,等. 云南白族遗传多态性研究[J]. 西安交通大学学报(医学版), 2002, 23(3): 242-245.
LAI Jiang-hua, ZHANG Bao-hua, ZHU Bo-feng, et al. STR polymorphism in Bai minority in Yunnan province [J]. Journal of Xi'an Jiaotong University (Med Sci), 2002, 23(3): 242-245.
- [3] YU Bing, ZHANG Hong-bo, LI Sheng-bin. X-chromosome STRs polymorphism of Han ethnic group from Northwest China [J]. Forensic Sci Int, 2005, 153(3): 269-271.
- [4] 王振原,方俊邦,王芳,等. 陕西省汉中地区人群 15 个 STR 基因座的遗传多态性分析[J]. 西安交通大学学报(医学版), 2004, 25(4): 333-335.
WANG Zhen-yuan, FANG Jun-bang, WANG Fang, et al. Polymorphism study on 15 STR loci of Han population from Hanzhong area [J]. Journal of Xi'an Jiaotong University (Medi Sci), 2004, 25(4): 333-335.
- [5] 余纯应,杨庆恩,朱传红,等. 汉族人群 5 个 STR 基因座的多态性调查[J]. 中国法医学杂志, 1997, 12(2): 69-74.
YU Chong-ying, YANG Qing-en, ZHU Chuan-hong, et al. A study of 5 STRs polymorphism loci in the Han population in Wuhan [J]. Chin J Forensic Med, 1997, 12(2): 69-74.
- [6] PING Zhiguang, GUO Xiong, GENG Dong, et al. Allele distribution of D12304, D12583, D121640 and D12S1708 in Chinese Han population [J]. Acad J Xjtu, 2005, 17(2): 164-167.
- [7] Kalmy B, Aldasher A, Khalmatov M, et al. Genome-wide scan for premature hypertension supports linkage to chromosome 2 in a large Kyrgyz family [J]. Hypertension, 2006, 48(5): 908-913.
- [8] Connolly K, Gabra H, Milluater J, et al. Identification of a region of frequent loss of heterozygosity at 11q24 in colorectal cancer [J]. Cancer Res, 1999, 59(12): 2806-2809.
- [9] Xu J, Gillander M, Isaccs S, et al. Genome-wide scan for prostate cancer susceptibility genes in the Johns Hopkins hereditary prostate cancer families [J]. Prostate, 2003, 57(4): 320-325.
- [10] Thomas A, Jayesh D. Genomics of osteoarthritis [J]. Curr Opin Rheumatol, 2003, 15(5): 634-640.
- [11] Thompson S D, Moroldo M B, Guyer L, et al. A genome-wide scan for juvenile rheumatoid arthritis in affected sibpair families provides evidence of linkage [J]. Arthritis Rheum, 2004, 50(9): 2920-2930.
- [12] 陈腾,金天博,辛娜,等. 中国甘肃青海 5 个民族群体 STR 基因座遗传多态性及其应用研究[J]. 中南大学学报(医学版), 2006, 31(6): 877-882.
CHEN Teng, JIN Tian-Bo, XIN Na, et al. Genetic polymorphism and application of 9 STR loci of 5 ethnic groups in Gansu and Qinhai [J]. J Cent South Univ (Med Sci), 2006, 31(6): 877-882.
- [13] Amar A, Brautbar C, Motro U, et al. Genetic variation of three tetrameric tandem repeats in four distict Israel ethnic groups [J]. J Forensic Sci, 1999, 44(5): 983-986.

(本文编辑 傅希文)