

# 中国朝鲜族 DXS7132、DXS6854 和 GATA31E08 基因座遗传多态性

张永吉<sup>1</sup>，徐京男<sup>2</sup>，张美花<sup>1</sup>，韩松瑛<sup>1</sup>

( 1. 延边大学基础医学院法医学教研室 , 延吉 133000 ； 2. 延边州公安局刑事科学技术研究所 , 延吉 133000 )

**摘要：**利用复合 PCR 扩增、变性聚丙烯酰胺凝胶电泳和银染方法对 261 名中国朝鲜族人群 3 个 X-STR 基因座研究发现 , 在 DXS7132 、 DX6854 和 GATA31E08 中分别检出 8 个、 6 个和 10 个等位基因 ; 基因频率分别分布在 0.006 ~ 0.344 、 0.015 ~ 0.450 和 0.005 ~ 0.317 之间 ; 分别检出 19 、 17 和 21 种基因型 ; 女性个体基因型频率分布均符合 Hardy-Weinberg 平衡 (  $P > 0.05$  ) ; 观测杂合度分别为 0.706 、 0.743 、 0.772 ; 多态信息含量分别为 0.710 、 0.730 、 0.670 ; 女性个人识别力分别为 0.898 、 0.911 、 0.857 ; 男性个人识别力分别为 0.766 、 0.802 、 0.734 ; 联体单倍型多样性大于 0.988 。结果表明 3 个基因座均具有较高的杂合度、多态信息量和个人识别力 , 是较理想的遗传标记系统 , 在研究中国朝鲜族起源和与其他民族之间的亲缘关系等方面具有较高的应用价值 , 并且对中国朝鲜族群体的亲子鉴定和个人识别提供宝贵的基础资料。

**关键词：**朝鲜族 ; 短串联重复序列 ; DXS7132 ; DX6854 ; GATA31E08

中图法分类号 : Q987 文献标识码 : A 文章编号 : 1000-3193(2010)01-0095-05

短串联重复序列 ( Short tandem repeat , STR ) 通常由长度为 2 ~ 6 个 bp 的核心序列串联重复而成 , 广泛存在于人类 DNA 编码区及非编码区 , 具有高度遗传多态性<sup>[1]</sup> , 并遵循孟德尔遗传规律。对中国朝鲜族常染色体和 Y 染色体 STR 的研究已经很多 , 但对 X 染色体 STR 的多态性资料比较少。然而 X 染色体 STR 遗传多态性对法医学个人识别、亲权鉴定和民族遗传学研究也起很重要的作用 , 如父 - 女关系确定、缺少双亲的姐妹之间关系确定和单倍型的研究等<sup>[2]</sup>。为了观察中国朝鲜族群体遗传多态性 , 我们选择了吉林延边地区的朝鲜族个体。延边朝鲜族自治州是朝鲜族群体从朝鲜半岛迁入到中国之后所形成的主要聚集地。我们应用复合扩增技术对这一群体 DXS7132 、 DXS6854 和 GATA31E08 基因座多态性进行了调查 , 获得其遗传多态性资料 , 为民族遗传学研究、法医学个人识别及亲子鉴定应用奠定基础。

## 1 对象与方法

### 1.1 研究对象

采集中国吉林延边地区朝鲜族 261 名 ( 女性 101 名 , 男性 160 名 ) 随机个体血样 ( 知情同意 ) , 用 Chelex-100 方法提取 DNA<sup>[3]</sup> , -20°C 冰冻保存。

收稿日期 : 2008-04-11 ; 定稿日期 : 2008-10-06

基金项目 : 国家自然科学基金 ( 30660201 )

作者简介 : 张永吉 ( 1971- ) , 男 , 延边大学基础医学院副教授 , 医学博士 , 主要从事法医物证学研究。 E-mail : zhangyj@ybu.edu.cn

## 1.2 方法

1.2.1 遗传标记 DDX7132( GDB 685605 ), DDX6854( GDB 386910 )和 GATA31E08( GDB : 684783 )基因座引物序列参照数据库( [www.gdb.org](http://www.gdb.org) )。引物由韩国 Genotech 公司合成。

1.2.2 PCR 反应 PCR 反应体系为 20 ,含 1 × 反应缓冲液 ,15 mmol/L MgCl<sub>2</sub> ,3 个基因座的引物各 0.5 μmol/L ,0.25 mmol/L dNTP ,DNA 10—30ng。采用 ABI 2400 型 PCR 仪 ,循环参数为 95℃ 3min→94℃ 30sec ,61℃ 30sec ,72℃ 1min 共 30 循环→72℃ 10min→保存 4℃。

1.2.3 变性凝胶电泳 采用 5% 聚丙烯酰胺、7mol/L 尿素凝胶垂直电泳( SA-32 , Biometra )。预热 20 分钟 ,上样后恒功率 51W ,电泳 2h。停泳后 ,硝酸固定 ,AgNO<sub>3</sub> 染色 ,Na<sub>2</sub>CO<sub>3</sub> 显色 ,乙酸停显<sup>[4]</sup>。

1.2.4 等位基因标记物的制备 利用琼脂糖凝胶回收试剂盒( QIAquick Gel Extraction Kit )回收不同长度纯合子 PCR 产物 ,并以测序结果按重复单位重复次数命名等位基因。

1.2.5 数据处理 通过 Power Stats Software 软件进行数据统计分析 ,得出各基因座的等位基因及基因型频率、各基因座的观察杂合度( heterozygosity observed , Ho )、个体识别力( power of discrimination , PD )、多态性信息含量( polymorphism information content , PIC ) ,非父排除率( power of exclusion , PE )按文献<sup>[5]</sup>对群体数据作 Hardy-Weinberg 平衡吻合度检验( P-value ) ,男性个体识别力或单倍型多样性( Gene or haplotype diversity )按文献<sup>[6]</sup> ,用 Arlequin 2.000 统计软件计算等位基因及群体之间的差异性<sup>[7]</sup>。

## 2 结果与讨论

### 2.1 结果

吉林延边地区朝鲜族 3 个 X 染色体 STR 基因座基因频率分布及群体遗传学指标见表 1。160 名男性 3 个 X 染色体 STR 单倍型频率和多样性见表 2。

### 2.2 讨论

为了利用基因多态性的法医学个人识别、亲子鉴定和民族遗传学研究 ,对多态性较高的基因位点提前做好群体调查工作 ,了解等位基因数目及频率分布是一项奠基工作 ,特别是对于建立各个民族的基因数据库尤为重要。近年来 ,由于 X 染色体的特殊遗传方式 ,X-STR 逐渐受到了法医物证研究者的重视<sup>[2]</sup> ,而对中国朝鲜族群体 X-STR 的研究尚未报道。因此 ,作者选择延边朝鲜族自治州的朝鲜族个体为研究对象 ,利用复合扩增的方法 ,对 DDX7132 、 DDX6854 和 GATA31E08 基因座多态性进行了研究。结果发现 ,DDX7132 、 DDX6854 和 GATA31E08 基因座的基因型分布均符合 Hardy-Weinberg 平衡定律 ;男、女之间的等位基因频率没有差异(各 P 值为 0.79980 ± 0.0266 ,0.87735 ± 0.0110 ,0.81915 ± 0.0150 ) 3 个基因座累计个人识别力高于 0.998 ;杂合性高于 0.706( 该值大于 0.7 有意义 )<sup>[8]</sup> ;多态信息含量高于 0.670( 该值大于 0.5 有高度可提供信息性 )<sup>[9]</sup> 。说明本研究中的 3 个基因座具有遗传稳定性和较高的识别能力 ,适合该群体的亲权鉴定、法医学个人识别和遗传学研究。在男性基因组中只有一个 X 染色体 ,所以 X-STR 以单倍型的形式存在 ,并只遗传给女儿<sup>[10]</sup> 。所以实际应用中应使用单倍型的分布频率来分析检测结果。本研究中以 160 名男性个体作为研究对象 ,调查了 3 个 X-STR 所组成的单倍型种类和频率。结果发现 ,共 91 种单倍型 ,其中 61 种单倍型仅出现一次( 67% ) ,单倍型多样性大于 0.988 ,显示较高

表 1 3 个 X 染色体 STR 基因座的等位基因分布频率和群体统计数据

Tab. 1 Allele frequencies and statistical values for 3 X-STR loci of Korean ethnic group in China

等位基因 Allele	DXS7132		DXS6854		GATA31E08	
	男 male	女 female	男 male	女 female	男 male	女 female
7	—	0.005	—	—	—	—
8	0.013	0.005	—	—	—	—
9	0.019	0.010	0.050	0.069	—	—
10	0.094	0.064	0.050	0.015	0.231	0.218
11	0.213	0.198	0.219	0.223	0.431	0.450
12	0.344	0.347	0.256	0.272	0.106	0.149
13	0.244	0.292	0.269	0.317	0.088	0.084
14	0.075	0.059	0.106	0.079	0.100	0.084
15	0.006	0.025	0.031	0.015	0.044	0.015
16	0.006	0.005	0.006	—	—	—
杂合度( H )		0.706		0.743		0.772
多态信息量( PIC )		0.710		0.730		0.670
个人识别力( PD )	0.766	0.898	0.802	0.911	0.734	0.857
PE	0.468		0.497		0.549	
P 值		0.060		0.963		0.795

H : heterozygosities , PIC : polymorphic information contents , PD : power of discrimination , PE : power of exclusion , p : Hardy-Weinberg equilibrium

表 2 中国延边地区朝鲜族 DDXS7132、DXS6854 和 GATA31E08 基因座单倍型分布频率

Tab. 2 Haplotype and diversity for DDXS7132、DXS6854 and GATA31E08 STR loci

of Korean ethnic group in China

类型 N	单倍型 Haplotype	频率 Frequency	类型 N	单倍型 Haplotype	频率 Frequency	类型 N	单倍型 Haplotype	频率 Frequency
1	9/10/10	0.025	32	12/8/13	0.006	63	13/12/10	0.006
2	9/12/12	0.006	33	12/9/11	0.013	64	13/12/15	0.013
3	9/14/15	0.006	34	12/10/12	0.006	65	13/12/13	0.006
4	10/8/11	0.025	35	12/11/11	0.025	66	13/13/11	0.019
5	10/9/11	0.025	36	12/11/12	0.006	67	13/13/10	0.006
6	10/10/12	0.031	37	12/11/14	0.006	68	13/13/13	0.006
7	10/11/10	0.006	38	12/11/10	0.006	69	13/13/14	0.006
8	10/11/11	0.006	39	12/11/13	0.019	70	13/13/12	0.019
9	10/12/11	0.025	40	12/12/11	0.006	71	13/14/12	0.006
10	10/12/14	0.019	41	12/12/12	0.006	72	13/14/13	0.006
11	10/13/10	0.006	42	12/12/10	0.006	73	13/14/10	0.006
12	10/13/13	0.019	43	12/12/15	0.006	74	13/14/11	0.006
13	10/14/11	0.025	44	12/13/10	0.013	75	13/14/14	0.006
14	10/14/15	0.013	45	12/13/11	0.006	76	13/15/11	0.006
15	10/15/11	0.031	46	12/13/15	0.013	77	13/15/10	0.006
16	11/9/14	0.006	47	12/13/12	0.006	78	13/16/11	0.006
17	11/9/11	0.013	48	12/13/14	0.006	79	14/9/10	0.006
18	11/10/11	0.013	49	12/14/12	0.013	80	14/11/10	0.006
19	11/11/11	0.006	50	12/14/11	0.006	81	14/11/14	0.006
20	11/11/13	0.006	51	12/15/11	0.006	82	14/11/11	0.006
21	11/11/14	0.006	52	13/9/10	0.013	83	14/12/11	0.006

续表 2

类型 N	单倍型 Haplotype	频率 Frequency	类型 N	单倍型 Haplotype	频率 Frequency	类型 N	单倍型 Haplotype	频率 Frequency
22	11/11/12	0.006	53	13/9/14	0.006	84	14/13/13	0.006
23	11/12/11	0.044	54	13/9/11	0.006	85	14/13/14	0.006
24	11/12/10	0.019	55	13/10/11	0.006	86	14/13/12	0.006
25	11/13/11	0.006	56	13/11/13	0.006	87	14/13/11	0.006
26	11/13/13	0.006	57	13/11/11	0.006	88	14/14/11	0.006
27	11/13/15	0.025	58	13/11/10	0.006	89	14/14/15	0.006
28	11/13/12	0.006	59	13/11/12	0.013	90	15/14/11	0.006
29	11/13/10	0.013	60	13/12/11	0.013	91	16/11/14	0.006
30	11/14/13	0.050	61	13/12/14	0.006			
31	11/15/10	0.006	62	13/12/12	0.025			

单倍型多样性( haplotype diversity ) : 0.988

的识别力,完全可以适用于朝鲜族群体的同父姐妹关系等鉴定。

为了观察群体之间的差异性,本研究群体与中国汉族、西班牙、蒙古和韩国群体之间进行了比较,结果发现存在着不同程度的差异。例如:在 DXS7132,香港汉族<sup>[11]</sup>有 7 个等位基因,以 14 等位基因最多见;蒙古族<sup>[12]</sup>有 6 个等位基因,以 14 等位基因最多见;西班牙群体<sup>[13]</sup>有 9 个等位基因,以 14 等位基因最多见;韩国群体<sup>[14]</sup>有 9 个等位基因,以 14 等位基因最多见。对中国朝鲜族的本文调查结果中,共检测到 10 个等位基因,以 12 等位基因最多见。在 DXS6854,广东汉族<sup>[15]</sup>有 7 个等位基因,以 10 等位基因最多见。而对中国朝鲜族的本文调查结果中,共检测到 8 个等位基因,以 13 等位基因最多见。在 GATA31E08,韩国群体<sup>[14]</sup>有 9 个等位基因,以 11 等位基因最多见,对中国朝鲜族的本文调查结果,共检测到 6 个等位基因,以 11 等位基因最多见。统计计算结果均显示显著差异性( $P$  值 = 0.000 ± 0.000),说明本研究中的 3 个基因位点均具有较强的民族、种族及地区差异性。

总之,DXS7132、DXS6854 和 GATA31E08 基因座在中国朝鲜族的分布具有很强的多态性,不仅从群体水平上能阐明中国朝鲜族 X-STR 遗传结构及其变化规律,而且丰富了我国朝鲜族 STR 遗传资料数据库,为中国朝鲜族群体的亲子鉴定、个人识别和民族遗传学的研究提供了宝贵的基础资料。

## 参考文献:

- [1] 戚其威,张红岩,马守中等.青岛地区汉族群体短串联重复序列基因座 D1S549、D3S1754 和 D12S375 的遗传多态性[J].中华医学遗传学杂志,2004,21:184-186.
- [2] 吕德坚,刘秋玲. X 染色体 STR 的法医学应用进展[J].中国法医学杂志,2002,17:252-255.
- [3] 金冬雁,黎孟枫主译.分子克隆实验指南[M].第 2 版.北京:科学出版社,1992:955-956.
- [4] Yongji Z, Kibeam Kim, Changho S, et al. Quadruplex PCR for 8 autosomal STR loci in Korean[J]. Koran J Legal Med, 2001, 25:17-28.
- [5] 江三多,吕宝忠.医学遗传数理统计方法[M].第 1 版.北京:科学出版社,1998.11.
- [6] Nei M. Molecular Evolutionary Genetics[A]. Columbia University Press, New York, 1987.
- [7] Yongji Z, Hongjiang Z, Yan C, et al. Population genetics for Y-Chromosomal STRs haplotypes of Chinese Korean ethnic group in northeastern China[J]. Forensic Sci Int, 2007, 173:197-203.
- [8] 赖江华,张保华,郑海波等.中国纳西族 STR 遗传结构研究[J].遗传学报,2002,29(7):576-580.
- [9] Nakamura Y, Leppert M, Oconnell P, et al. Variable number of tandem repeat( VNTR ): markers for human gene

- mapping [ J ]. *Science* , 1987 , 235( 27 ) : 1616-1622.
- [ 10 ] Szibor R , Krawczak M , Hering S , et al. Use of X-linked markers for forensic purposes [ J ]. *Int J Legal Med* , 2003 , 117 : 67-74.
- [ 11 ] Tang WM , To KY. Four X-chromosomal STRs and their allele frequencies in a Chinese population [ J ]. *Forensic Science International* , 2006 , 162 : 64-65.
- [ 12 ] Liu QB , Li SB. Patterns of genetic polymorphism at the 10 X-chromosome STR loci in Mongol population [ J ]. *Forensic Science International* , 2006 , 158 : 76-79.
- [ 13 ] Aler M , Diz PS , Gomes I , et al. Genetic data of 10 X-STRs in a Spanish population sample [ J ]. *Forensic Science International* , 2007 , 173 : 193-196.
- [ 14 ] Shin SH , Yu JS , Park SW , et al. Genetic analysis of 18 X-linked short tandem repeat markers in Korean population [ J ]. *Forensic Science International* , 2005 , 147 : 35-41.
- [ 15 ] 刘秋玲 , 吕德坚 , 崔葳. 3 个 X 染色体短串联重复的复合扩增及其多态性 [ J ]. 中华医学遗传学杂志 , 2004 , 21( 3 ) : 233-235.

## Genetic Polymorphisms of DXS7132 , DXS6854 , and GATA31E08 STR Loci Found in a Korean Ethnic Group in China

ZHANG Yong-ji<sup>1</sup> , XU Jing-nan<sup>2</sup> , ZHANG Mei-hua<sup>1</sup> , HAN Song-ying<sup>1</sup>

( 1. Department of Forensic Medicine , Yanbian University College of Medicine , Yanji 133000 ;

2. Department of Forensic Science , Yanbian States Police Station , Yanji 133000 )

**Abstract :** The purpose of this research was to understand the allele structure and genetic polymorphism at DXS7132 , DXS6854 , and GATA31E08 short tandem repeats ( STRs ) loci in a Korean ethnic group found in the Jilin Yanbian area , and to construct a preliminary descriptive database. The allele frequencies of the three STRs loci in 261 unrelated individuals from this ethnic group were analyzed by multiplex polymerase chain reaction ( PCR ) and polyacrylamide gel electrophoresis ( PAGE ). The following 8 , 6 and 10 alleles and 19 , 17 and 21 genotypes were observed at the DXS7132 , DXS6854 and GATA31E08 STR loci , respectively. All of these loci ( in females ) met the Hardy-Weinberg equilibrium (  $P > 0.05$  ). Statistical analysis of the three STR loci calculated the heterozygosities at 0.706 , 0.743 , and 0.772 ; the polymorphic information contents ( PIC ) at 0.710 , 0.730 and 0.67. The female powers of discrimination ( PD ) were 0.898 , 0.911 and 0.857. The male genetic diversity of the male samples were 0.766 , 0.802 and 0.734 , and the combined haplotype diversity was more than 0.988. All of these results show that all three of the loci in this study were found to have high heterozygosity and polymorphic information content , and thus , they could provide useful genetic markers. These results could serve as valuable data to enrich the Korean ethnic group genetic database and play an important role in application and understanding of Chinese population genetics.

**Key words :** Korean ethnic group ; Short tandem repeat ; DXS7132 ; DXS6854 ; GATA31E08