

·基础研究·

北方汉族31个X染色体短串联重复片段突变及遗传多态性研究

徐妍, 刘岩, 郝世诚, 张金佩, 刘革新, 陶玉婷, 王显众, 袁丽
(中国政法大学证据科学教育部重点实验室, 北京 100088)

摘要:【目的】调查31个X-STR基因座在中国北方汉族的遗传多态性、连锁不平衡情况和突变率。【方法】采集209个中国北方汉族健康志愿者家系样本,用Microreader™ 19X和AGCU X19 STR荧光检测试剂盒检测,统计基因座的突变率。选取家系中父亲或母亲各207个样本,组成男性和女性无关个体组并进行遗传多态性研究和连锁不平衡检验。【结果】共检出344个等位基因,基因频率分布在0.001 6-0.810 0之间。其中多态性信息含量最丰富的基因座是DXS10135(PIC=0.912 4)。个体识别能力在男、女群体中分别为0.330 5-0.918 1、0.534 0-0.987 6。累积个人识别概率在男、女群体中分别为 $1-4.566 4 \times 10^{-19}$ 、 $1-1.578 4 \times 10^{-31}$ 。平均排除概率在三联体和二联体中分别为0.312 7-0.912 4、0.193 5-0.844 6。累积平均排除概率在三联体和二联体中分别为 $1-2.281 9 \times 10^{-17}$ 、 $1-2.509 0 \times 10^{-12}$ 。经分析,有1对X-STR(DXS10103-DXS10101)存在明显连锁不平衡现象。在414次减数分裂中观察到19个X-STR基因座有突变,平均突变率为0.002 3,突变率最高的基因座是DXS10135,为0.012 1。【结论】获得的31个X-STR遗传学数据为法医学鉴定提供了基础数据。DXS10103-DXS10101存在连锁不平衡,在应用时需按照单倍型频率分析。应用高突变率基因座时需保持谨慎,不要因为个别基因座不符合遗传规律而排除具有某种亲缘关系。

关键词: X染色体;短串联重复片段;北方汉族;遗传多态性;突变

中图分类号: D919.4 **文献标志码:** A **文章编号:** 1672-3554(2022)04-0511-11

DOI: 10.13471/j.cnki.j.sun.yat-sen.univ(med.sci).2022.0401

Mutation and Genetic Polymorphism of 31 X-STR Loci in Northern Han Nationality

XU Yan, LIU Yan, HAO Shi-cheng, ZHANG Jin-pei, LIU Ge-xin,
TAO Yu-ting, WANG Xian-zhong, YUAN Li

(Key Laboratory of Evidence Science of Ministry of Education, China University of Political Science and Law,
Beijing 100088, China)

Correspondence to: YUAN Li; E-mail: yuanliwcy@126.com

Abstract: 【Objective】 To investigate the genetic polymorphism, linkage disequilibrium and mutation rate of 31 X-STR loci in the Northern Han Chinese. 【Methods】 A total of 209 family samples of healthy volunteers in the Northern Han Chinese were selected. All samples were detected by AGCU X19 and Microreader™ 19X fluorescence detection kits and the mutation rate of loci was counted. The mothers ($n=207$) and the fathers ($n=207$) in the families were selected, respectively, to form the group of unrelated male and female individuals for genetic polymorphism studies and linkage disequilibrium test. 【Results】 A total of 344 alleles were detected with gene frequencies ranging from 0.001 6 to 0.810 0. The locus with the most abundant polymorphism information content (PIC) was DXS10135 (PIC=0.912 4). Discrimination power

收稿日期:2022-03-03

基金项目:教育部人文社会科学研究规划基金(19YJA820050);中国政法大学“双一流”建设项目(1000-10321021)

作者简介:徐妍,硕士生,研究方向:法医学物证学, E-mail: 15732189176@163.com;袁丽,通信作者,教授,主任法医师,研究方向:法医学物证学, E-mail: yuanliwcy@126.com

(DP) was 0.330 5–0.918 1 and 0.534 0–0.987 6 in male and female groups respectively. The cumulative discrimination power was $1-4.566\ 4\times 10^{-19}$ and $1-1.578\ 4\times 10^{-31}$ in male and female groups respectively. Mean exclusion chance (MEC) was 0.312 7–0.912 4 and 0.193 5–0.844 6 in trios and duos. The cumulative mean exclusion chance was $1-2.281\ 9\times 10^{-17}$ and $1-2.509\ 0\times 10^{-12}$ in trios and duos respectively. After analysis, one pair of loci (DXS10103–DXS10101) had obvious linkage disequilibrium. Mutations at 19 X–STR loci were observed in 414 meiotic divisions, with an average mutation rate of 0.002 3. DXS10135 had the highest mutation rate of 0.012 1.【Conclusion】The obtained 31 X–STR genetic data provided basic data for forensic identification. DXS10103–DXS10101 shows linkage disequilibrium, so haplotype frequency could be counted in application. Care should be taken when using loci with high mutation, and don't rule out some kind of kinship only because individual loci do not conform to the laws of inheritance.

Key words: X chromosome; short tandem repeats; Northern Han; genetic polymorphisms; mutation

[J SUN Yat-sen Univ (Med Sci), 2022, 43(4): 511–521]

女性含有两条X染色体,除连锁基因外,在减数分裂时可以发生同源重组,随机遗传给子代一条X染色体,与常染色体类似。男性体内只有一条X染色体,且除拟常染色区外,X染色体与Y染色体缺乏同源重组^[1],因此减数分裂时男性个体的X–STR以一种类似于Y–STR的单倍型方式传递给其女儿。这种独特的遗传方式使X–STR在含有女性的亲缘关系鉴定案件,如全同胞姐妹、同父异母半同胞姐妹、祖母和孙女、父女、母子或母女、三联体鉴定中^[2–3]能够发挥独特的应用价值,并且可以利用X–STR的基因频率或单倍型频率计算亲缘关系的似然率。当然,当联合应用多个X–STR检验时需关注基因座之间是否存在连锁不平衡,以及突变等问题^[4]。在一些家系重建案件中,检验更多的X–STR来获得更多的遗传信息^[5],以推导出重建人员更多确切的基因分型,往往能获得非常好的效果。既往报道的X–STR基因座数量有限,而在复杂亲缘关系鉴定时往往需要检测更多的X–STR,这些基础群体数据缺乏,亟需有群体遗传信息来支撑案件的分析。基于以上背景,本文对北方汉族群体和家系样本检测31个X–STR基因座,开展了X–STR遗传多态性、连锁不平衡检验和突变情况的研究。

1 材料与方法

1.1 样本

在获取知情同意条件下采集209个中国北方汉族家系样本,具体为:205个母亲–女儿–父亲三联体,2个母亲–女儿二联体,2个父亲–女儿二联

体。所有家系样本已经常染色体STR检验,亲权指数均超过10 000,支持有亲子关系。家系样本用于观察31个X–STR基因座在414次减数分裂过程中的突变率。选择所有家系中母亲/父亲样本,分别组成207个女性/男性无关个体群体,总共无关个体为414人。本研究所设计样本实验均经中国政法大学证据科学研究院伦理委员会批准[批准号No. 2020003],所有研究对象均签署了知情同意书。

1.2 X–STR分型检测

用Chelex–100法提取基因组DNA。分别按照AGCU X19 STR试剂盒和Microreader™ 19X Direct ID System试剂盒说明书扩增各样本DNA。扩增产物经Applied Biosystems 3130遗传分析仪检测、GeneMapper® ID–X v1.4软件分析,获得31个X–STR基因座的基因型,分别为DXS101、DXS981、DXS6789、DXS6795、DXS6800、DXS6803、DXS6807、DXS6809、DXS6810、DXS7132、DXS7133、DXS7423、DXS7424、DXS8378、DXS9902、DXS9907、DXS10074、DXS10075、DXS10079、DXS10101、DXS10103、DXS10134、DXS10135、DXS10148、DXS10159、DXS10162、DXS10164、GATA165B12、GATA172D05、GATA31E08和HPRTB。

1.3 统计学分析

对于209个家系的X–STR分型结果,依据遗传规律观察是否存在突变的基因座,并计算发生突变的基因座突变率。仅针对207例女性X–STR基因座结果,使用Arlequin v3.5软件进行31个X–STR遗传标记的Hardy–Weinberg平衡检验及连锁不平衡检验。使用Arlequin v3.5软件对男、女群体等位基因频率分布进行显著性差异分析。使用

StatsX v2.0软件统计各基因座的等位基因频率、观察杂合度(heterozygosity observed, Hobs)、期望杂合度(expected heterozygosity, Hexp)、基因差异度(gene diversity, GD)、女性个人识别率(discrimination power of female, DP_f)、男性个人识别率(discrimination power of male, DP_m)、多态性信息量(polymorphism information content, PIC)、平均父权排除概率(mean exclusion chance, MEC):父女或母子二联体X染色体标记的平均父权排除概率(MEC_Desmarais_duo)、涉及女儿的三联体X染色体标记的平均父权排除概率(MEC_Desmarais)。此外,将本研究的北方汉族群体分别与德国、日本、广东汉族、四川藏族群体^[6-9]进行显著性差异分析。

2 结果

2.1 31个X-STR基因座的突变情况

根据遗传规律观察209个家系的X-STR分型,经分析在19个基因座(DXS981、DXS10148、DXS7132、HPRTB、DXS9907、DXS10162、DXS7424、DXS10159、DXS10075、GATA172D05、DXS10135、DXS9902、DXS6810、DXS10164、DXS10101、DXS10079、DXS10103、DXS6789、GATA165B12)上发现29个新发突变。总家系平均突变率为0.0023。在这19个X-STR基因座上突变率最高的基因座是DXS10135,为0.0121;其次是HPRTB,为0.0072;DXS7132、DXS10159、DXS981、DXS10162基因座突变率均为0.0048(表1)。

表1 X-STR基因座的新发突变情况
Table 1 Denovo mutations at the X-STR loci

Locus	Number of mutations	Paternal mutation	Maternal mutation	Unknown mutation direction	Mutations with increased number of steps	Mutations with reduced number of steps	mutation rates	Binominal 95% CI
DXS7132	2	2				2	0.004 8	(6.00×10^{-4} , 1.73×10^{-2})
DXS10164	1		1			1	0.002 4	(1.00×10^{-4} , 1.34×10^{-2})
HPRTB	3	3			1	1	0.007 2	(1.50×10^{-3} , 2.10×10^{-2})
DXS10159	2	2				2	0.004 8	(6.00×10^{-4} , 1.73×10^{-2})
DXS9907	1	1			1		0.002 4	(1.00×10^{-4} , 1.34×10^{-2})
DXS981	2	1		1	1	1	0.004 8	(6.00×10^{-4} , 1.73×10^{-2})
DXS10162	2	2				2	0.004 8	(6.00×10^{-4} , 1.73×10^{-2})
DXS10079	1	1				1	0.002 4	(1.00×10^{-4} , 1.34×10^{-2})
DXS7424	1		1			1	0.002 4	(1.00×10^{-4} , 1.34×10^{-2})
DXS6789	1		1			1	0.002 4	(1.00×10^{-4} , 1.34×10^{-2})
DXS10075	1			1			0.002 4	(1.00×10^{-4} , 1.34×10^{-2})
GATA172D05	1			1	1		0.002 4	(1.00×10^{-4} , 1.34×10^{-2})
DXS10135	5	5			2	3	0.012 1	(3.90×10^{-3} , 2.80×10^{-2})
DXS9902	1	1				1	0.002 4	(1.00×10^{-4} , 1.34×10^{-2})
DXS6810	1	1			1		0.002 4	(1.00×10^{-4} , 1.34×10^{-2})
DXS10101	1	1			1		0.002 4	(1.00×10^{-4} , 1.34×10^{-2})
DXS10148	1	1			1		0.002 4	(1.00×10^{-4} , 1.34×10^{-2})
DXS10103	1	1				1	0.002 4	(1.00×10^{-4} , 1.34×10^{-2})
GATA165B12	1			1		1	0.002 4	(1.00×10^{-4} , 1.34×10^{-2})

mutation rate = number of mutation cases/number of meiosis (M), M=414.

2.2 31个X-STR基因座的遗传多态性

207例女性样本的31个X-STR基因座基因型分布P值均大于0.05,均符合Hardy-Weinberg平衡。414例无关个体的31个X-STR基因座共检测出了344个不同的等位基因。对男、女群体等位基因分布频率进行显著性差异分析后的P值均大于0.05,表明31个X-STR等位基因频率分布在男、女

两个群体中不存在显著差异,同时女性群体符合Hardy-Weinberg平衡,故将女性、男性样本合并计算每个基因座的等位基因频率。表2列出了合并后的基因座等位基因频率。等位基因的频率分布在0.001 6~0.810 0之间,其中DXS6800的等位基因16频率最高,为0.810 0。

表2 31个X-STR基因座在中国北方汉族群体中的等位基因频率

Table 2 Allele frequencies of 31 X-STR loci in Northern Han of Chinese

(n=414)

Loci		Loci		Loci		Loci	
DXS101		DXS981		DXS6789		DXS6795	
A	F	A	F	A	F	A	F
18	0.001 6	10	0.001 6	14	0.003 2	9	0.051 5
19	0.003 2	11	0.001 6	15	0.148 1	10	0.159 4
21	0.008 1	12	0.030 6	16	0.256 0	11	0.278 6
22	0.037 0	12.3	0.074 1	17	0.046 7	12	0.041 9
23	0.115 9	13	0.138 5	18	0.004 8	13	0.433 2
24	0.301 1	13.3	0.177 1	19	0.030 6	14	0.030 6
25	0.212 6	14	0.309 2	20	0.236 7	15	0.004 8
26	0.198 1	14.2	0.001 6	21	0.185 2	DXS6809	
27	0.083 7	14.3	0.067 6	22	0.074 1	A	F
28	0.019 3	15	0.140 1	23	0.011 3	29	0.009 7
29	0.008 1	15.3	0.017 7	24	0.001 6	30	0.017 7
30	0.006 4	16	0.037 0	25	0.001 6	31	0.159 4
31	0.003 2	17	0.003 2	DXS6807		32	0.190 0
32	0.001 6	DXS6803		A	F	33	0.286 6
DXS6800		A	F	10.3	0.003 2	34	0.202 9
A	F	10	0.003 2	11	0.394 5	35	0.090 2
14	0.001 6	10.3	0.001 6	12	0.016 1	36	0.033 8
15	0.001 6	11	0.138 5	13	0.032 2	37	0.006 4
16	0.810 0	11.3	0.133 7	13.3	0.004 8	38	0.003 2
18	0.011 3	12	0.106 3	14	0.336 6	DXS7423	
19	0.101 4	12.3	0.473 4	14.3	0.001 6	A	F
20	0.001 6	13	0.027 4	15	0.183 6	13	0.003 2
21	0.019 3	13.3	0.107 9	15.3	0.001 6	14	0.277 0
22	0.051 5	14.3	0.008 1	16	0.025 8	15	0.655 4
23	0.001 6	DXS7132		DXS7133		16	0.061 2
DXS6810		A	F	A	F	17	0.003 2
A	F	11	0.011 3	6	0.008 1	DXS9907	
15	0.001 6	12	0.082 1	8	0.004 8	A	F
16	0.011 3	13	0.222 2	9	0.750 4	9	0.006 4
17	0.157 8	14	0.349 4	10	0.190 0	10	0.001 6

续表

Loci		Loci		Loci		Loci	
18	0.545 9	15	0.256 0	11	0.043 5	11	0.038 6
19	0.273 8	16	0.061 2	12	0.003 2	12	0.396 1
20	0.009 7	17	0.014 5	DXS9902		13	0.471 8
DXS7424		18	0.003 2	A	F	14	0.074 1
A	F	DXS8378		7	0.003 2	15	0.011 3
11	0.004 8	A	F	9	0.020 9	DXS10101	
12	0.003 2	9	0.025 8	9.3	0.001 6	A	F
13	0.058 0	10	0.545 9	10	0.447 7	26	0.004 8
14	0.143 3	11	0.285 0	11	0.339 8	26.2	0.001 6
15	0.320 5	12	0.133 7	12	0.177 1	27	0.008 1
16	0.354 3	13	0.009 7	12.1	0.001 6	27.2	0.001 6
17	0.103 1	DXS10075		13	0.008 1	27.3	0.001 6
18	0.012 9	A	F	DXS10079		28	0.017 7
DXS10074		13	0.001 6	A	F	28.2	0.029 0
A	F	14	0.001 6	15	0.006 4	29	0.038 6
7	0.001 6	15	0.009 7	16	0.012 9	29.2	0.041 9
12	0.001 6	16	0.220 6	17	0.054 8	30	0.090 2
13	0.003 2	16.2	0.020 9	18	0.093 4	30.2	0.069 2
14	0.019 3	17	0.413 8	18.2	0.001 6	31	0.214 2
15	0.058 0	17.2	0.041 9	19	0.228 7	31.1	0.001 6
16	0.188 4	18	0.235 1	20	0.289 9	31.2	0.091 8
17	0.339 8	18.2	0.006 4	21	0.190 0	32	0.148 1
18	0.256 0	19	0.040 3	21.2	0.001 6	32.2	0.064 4
18.2	0.001 6	20	0.008 1	22	0.078 9	32.3	0.001 6
18.3	0.003 2	DXS10134		23	0.033 8	33	0.096 6
19	0.103 1	A	F	24	0.004 8	33.2	0.038 6
20	0.022 5	25	0.001 6	25	0.001 6	34	0.033 8
21	0.001 6	27.3	0.001 6	26	0.001 6	34.2	0.001 6
DXS10103		30	0.003 2	DXS10135		35	0.003 2
A	F	30.2	0.001 6	A	F	DXS10148	
14	0.001 6	31	0.003 2	15	0.003 2	A	F
15	0.014 5	32	0.030 6	16	0.001 6	18	0.120 8
16	0.307 6	33	0.030 6	17	0.009 7	19	0.030 6
17	0.141 7	33.2	0.001 6	18	0.041 9	20	0.024 2
18	0.186 8	34	0.083 7	19	0.096 6	20.1	0.003 2
18.2	0.001 6	35	0.198 1	20	0.104 7	21	0.004 8
19	0.264 1	35.2	0.001 6	21	0.148 1	21.1	0.008 1
20	0.070 9	35.3	0.008 1	22	0.112 7	22.1	0.066 0
21	0.009 7	36	0.188 4	23	0.083 7	23	0.001 6
22	0.001 6	36.3	0.001 6	24	0.075 7	23.1	0.067 6

续表

Loci		Loci		Loci		Loci	
DXS10159		37	0.191 6	25	0.054 8	24.1	0.101 4
A	F	37.3	0.041 9	26	0.051 5	24.2	0.001 6
21	0.003 2	38	0.119 2	26.1	0.001 6	25.1	0.127 2
22	0.011 3	38.3	0.009 7	27	0.048 3	26	0.003 2
23	0.051 5	39	0.040 3	28	0.040 3	26.1	0.146 5
24	0.301 1	39.3	0.003 2	29	0.029 0	27.1	0.112 7
25	0.254 4	40	0.006 4	29.2	0.001 6	27.2	0.004 8
26	0.202 9	40.3	0.003 2	30	0.033 8	28.1	0.072 5
27	0.117 6	41	0.001 6	31	0.014 5	28.2	0.003 2
28	0.046 7	41.3	0.004 8	32	0.017 7	29.1	0.064 4
29	0.009 7	43.3	0.004 8	32.2	0.001 6	29.2	0.001 6
30	0.001 6	DXS10162		33	0.012 9	30.1	0.020 9
GATA172D05		A	F	34	0.004 8	30.2	0.001 6
A	F	13	0.003 2	35	0.001 6	31.1	0.008 1
5	0.003 2	15	0.008 1	36	0.003 2	32.2	0.001 6
6	0.041 9	16	0.041 9	38	0.004 8	33.1	0.001 6
7	0.006 4	17	0.191 6	DXS10164		GATA165B12	
8	0.135 3	17.2	0.006 4	A	F	A	F
9	0.112 7	18	0.304 3	8	0.037 0	9	0.233 5
10	0.418 7	18.2	0.001 6	9	0.017 7	10	0.563 6
11	0.233 5	18.3	0.001 6	10	0.565 2	11	0.185 2
12	0.048 3	19	0.278 6	11	0.235 1	12	0.016 1
GATA31E08		20	0.115 9	12	0.101 4	13	0.001 6
A	F	21	0.037 0	13	0.022 5	HPRTB	
5	0.001 6	22	0.008 1	14	0.012 9	A	F
6	0.001 6	23	0.001 6	15	0.008 1	10	0.001 6
7	0.070 9					11	0.049 9
8	0.022 5					12	0.320 5
9	0.175 5					13	0.394 5
10	0.238 3					14	0.183 6
11	0.376 8					15	0.045 1
12	0.101 4					16	0.004 8
13	0.008 1						
14	0.003 2						

A = allele, F = frequency.

中国北方汉族群体31个X-STR基因座的法医学参数见表3。31个X-STR基因座的PIC在0.312 7-0.912 4之间, DP_M 在0.330 5-0.918 1之间, DP_F 在0.534 0-0.987 6之间, MEC_Desmarais在0.312 7-0.912 4, MEC_Desmarais_duo在0.193 5-0.844 6之间。根据

ChrX-STR.org 2.0数据库(<http://www.chrx-str.org/>)中的公式,31个X-STR的女性群体累积个人识别概率为 $1-1.578 4 \times 10^{-31}$,男性群体累积个人识别概率为 $1-4.566 4 \times 10^{-19}$ 。累积的平均父权排除概率,三联体为 $1-2.281 9 \times 10^{-17}$,二联体为 $1-2.509 0 \times 10^{-12}$ 。

表3 31个X-STR基因座在中国北方汉族群体中的法医学参数
Table 3 Forensic parameters of 31 X-STR loci in Northern Han of Chinese

Locus	N-Al- leles	H _{obs}	H _{exp}	GD	PIC	DP _M	DP _F	MEC_Des- marais	MEC_Des- marais_duo
DXS8378	5	0.623 2	0.595 3	0.603 1	0.539 6	0.602 1	0.779 1	0.539 5	0.392 5
DXS7423	5	0.478 3	0.495 9	0.490 7	0.420 2	0.490 0	0.670 1	0.420 2	0.282 9
DXS10148	25	0.903 4	0.905 5	0.905 7	0.896 4	0.904 3	0.983 0	0.896 4	0.819 3
DXS10159	10	0.787 4	0.780 8	0.785 8	0.752 4	0.784 5	0.921 5	0.752 4	0.623 1
DXS10134	25	0.845 4	0.868 3	0.862 5	0.846 2	0.861 1	0.965 8	0.846 2	0.746 2
DXS7424	8	0.753 6	0.739 3	0.738 2	0.694 7	0.737 0	0.888 5	0.694 7	0.555 7
DXS10164	8	0.589 4	0.602 9	0.613 6	0.567 7	0.612 6	0.805 0	0.567 7	0.419 8
DXS10162	13	0.797 1	0.781 6	0.777 6	0.742 4	0.776 3	0.916 1	0.742 4	0.611 2
DXS7132	8	0.763 3	0.756 7	0.753 4	0.712 5	0.752 2	0.898 9	0.712 5	0.575 5
DXS10079	14	0.840 6	0.810 2	0.809 5	0.782 7	0.808 2	0.937 7	0.782 7	0.661 1
DXS6789	12	0.806 8	0.810 8	0.814 7	0.787 7	0.813 4	0.939 5	0.787 7	0.667 0
DXS101	14	0.811 6	0.809 7	0.803 8	0.775 6	0.802 5	0.934 0	0.775 6	0.651 9
DXS10103	10	0.768 1	0.770 7	0.776 6	0.740 3	0.775 3	0.914 5	0.740 3	0.607 9
DXS10101	22	0.869 6	0.891 9	0.891 6	0.881 0	0.890 2	0.978 8	0.881 0	0.796 8
HPRTB	7	0.719 8	0.701 3	0.704 5	0.651 3	0.703 4	0.860 0	0.651 3	0.507 5
DXS6809	10	0.806 8	0.805 8	0.806 7	0.778 1	0.805 4	0.934 8	0.778 1	0.654 5
DXS10075	11	0.719 8	0.716 7	0.722 0	0.677 6	0.720 8	0.878 9	0.677 6	0.536 4
DXS10074	13	0.821 3	0.763 9	0.769 8	0.734 1	0.768 6	0.911 9	0.734 1	0.601 2
DXS10135	26	0.913 0	0.922 0	0.919 6	0.912 4	0.918 1	0.987 6	0.912 4	0.844 6
DXS6795	7	0.705 3	0.704 0	0.705 1	0.658 2	0.703 9	0.866 6	0.658 2	0.515 1
DXS6803	9	0.724 6	0.715 3	0.716 2	0.685 0	0.715 1	0.888 8	0.685 0	0.543 7
DXS6807	10	0.719 8	0.704 7	0.696 5	0.640 7	0.695 4	0.852 9	0.640 7	0.496 5
DXS9907	7	0.642 5	0.628 8	0.614 3	0.538 1	0.613 4	0.775 2	0.538 1	0.394 3
GATA172D05	8	0.724 6	0.745 3	0.736 2	0.699 0	0.735 0	0.893 8	0.699 0	0.560 3
DXS9902	8	0.623 2	0.650 9	0.653 3	0.585 7	0.652 2	0.812 6	0.585 8	0.439 0
DXS7133	6	0.439 6	0.420 2	0.399 5	0.355 8	0.398 8	0.595 5	0.355 8	0.227 1
DXS6810	6	0.565 2	0.596 9	0.602 9	0.538 5	0.601 9	0.778 1	0.538 4	0.390 9
GATA31E08	10	0.821 3	0.761 4	0.755 7	0.718 7	0.754 5	0.904 0	0.718 7	0.582 9
DXS6800	9	0.333 3	0.335 2	0.331 0	0.312 7	0.330 5	0.534 0	0.312 7	0.193 5
DXS981	13	0.792 3	0.826 1	0.822 8	0.800 6	0.821 5	0.947 2	0.800 6	0.684 4
GATA165B12	5	0.579 7	0.580 6	0.594 2	0.532 9	0.593 3	0.774 2	0.532 9	0.384 7

2.3 31个X-STR基因座的连锁不平衡分析结果

在31个X-STR基因座配对组成的465个配对

基因座中,一共30对基因座连锁不平衡检验 $P < 0.05$,见表4。经Bonferroni校正($0.05/465=0.000$)

表4 连锁不平衡检验中 P 值小于0.05的基因座对
Table 4 Locus pairs with P value less than 0.05 in linkage disequilibrium test

Pairs of loci	P	Pairs of loci	P
DXS8378-DXS9907	0.003 2	DXS7132-DXS6809	0.022 1
DXS8378-GATA165B12	0.002 2	DXS7132-DXS6803	0.046 4
DXS7423-DXS10162	0.003 4	DXS7132-GATA172D05	0.008 5
DXS7423-DXS10103	0.040 8	DXS101-HPRTB	0.009 3
DXS7423-DXS981	0.035 1	DXS101-DXS6800	0.004 6
DXS10148-DXS9907	0.035 9	DXS10103-DXS10101	0.000 0
DXS10148-GATA172D05	0.006 8	DXS10103-HPRTB	0.018 0
DXS10159-GATA172D05	0.002 7	HPRTB-GATA172D05	0.041 0
DXS7424-DXS7132	0.049 2	DXS10075-DXS10074	0.043 8
DXS7424-DXS7133	0.031 8	DXS10074-DXS10135	0.042 1
DXS7424-DXS981	0.046 9	DXS10135-DXS6795	0.045 0
DXS10164-DXS101	0.035 1	DXS10135-DXS7133	0.039 0
DXS7132-DXS10103	0.002 7	DXS6807-DXS6810	0.038 2
DXS7132-DXS10101	0.043 2	DXS6807-DXS981	0.027 8
DXS9907-DXS6800	0.048 2	DXS9902-GATA165B12	0.008 3

The black bold is after Bonferroni correction, $P < 0.05/465 = 0.000\ 107\ 526\ 9$.

107 526 9), 仍有 1 对基因座 (DXS10103-DXS10101) 具有统计学意义, 存在连锁不平衡。DXS10103-DXS10101 单倍型共有 60 种, 单倍型频率在 0.004 8-0.087 0 之间(表 5), 其中单倍型 16-31 的频率最高, 为 0.087 0。

2.4 北方汉族群体与其他四个群体之间 X-STR 基因座等位基因频率分布的比较

本研究的北方汉族群体分别与德国、日本、广东汉族、四川藏族群体比较 P 值见表 6。北方汉族群体与德国群体在可比较的 13 个基因座上均存在明显差异, 与日本群体在可提供数据比较的 14 个基因座中有 8 个基因座存在差异, 与广东汉族群体在 11 个基因座中的 3 个基因座上存在差异、与四川藏族群体在可比较的 19 个基因座中的 6 个基因座上存在差异。

3 讨论

经家系分析, 31 个 X-STR 在 414 次减数分裂中发生了 29 个突变, 且均为一步突变, 未观察到有一个家系出现 2 个突变。除 4 个突变来源不明外, 有 22 个突变来源于父亲, 有 3 个突变来源于母亲, 父

源性突变与母源性突变的观察值比例约为 7.3:1, 表明了父源性突变相比较于母源性突变来说更容易发生。这种突变机率的差异可能来源于生殖细胞分化形成卵子和精子过程中细胞分裂次数的差异^[10]。对于女性来说, 分裂形成卵子的卵原细胞在胚胎发育结束时几乎已经确定, 而对于男性来说, 精原细胞能不断地进行有丝分裂, 增加细胞数量并分化为精母细胞。因此, 相比于卵子, 精子的形成经历了更多的细胞分裂, 这有可能是父源性突变率更高的原因之一。除突变来源外, 突变率的大小还可能与遗传标记重复序列的结构特征、种群、样本量大小等因素有关^[11-12]。本次研究的北方汉族人群中, 31 个 X-STR 基因座平均突变率为 0.002 3, 与常染色体 STR 和 Y-STR 的平均突变率相近^[13-16]。在北方汉族群体, DXS7132、HPRTB、DXS10159、DXS981、DXS10162 和 DXS10135 基因座的突变率明显高于平均突变率。对 31 个 X-STR 突变情况的研究, 提示我们在检测 X-STR 时, 特别是 DXS7132、HPRTB、DXS10159、DXS981、DXS10162 和 DXS10135 这几个突变率较高的基因座, 需要保持谨慎, 不要因为个别基因座不符合遗传规律而排除具有某种亲缘关系, 尤其是矛盾等位

表5 DXS10103-DXS10101单倍型频率

Table 5 Haplotype frequency of DXS10103-DXS10101

Haplotype	Number	frequency	Haplotype	Number	frequency
15-31	1	0.004 8	19-28	2	0.009 7
15-32	1	0.004 8	19-28.2	4	0.019 3
16-29	3	0.014 5	19-29	3	0.014 5
16-30	7	0.033 8	19-29.2	5	0.024 2
16-30.2	1	0.004 8	19-30	3	0.014 5
16-31	18	0.087 0	19-30.2	6	0.029 0
16-32	15	0.072 5	19-31	8	0.038 6
16-33	10	0.048 3	19-31.2	12	0.058 0
16-34	4	0.019 3	19-32	3	0.014 5
17-28.2	3	0.014 5	19-32.2	6	0.029 0
17-29.2	1	0.004 8	19-33	1	0.004 8
17-30	8	0.038 6	19-33.2	2	0.009 7
17-30.2	1	0.004 8	20-27	1	0.004 8
17-31	9	0.043 5	20-28	1	0.004 8
17-32	4	0.019 3	20-28.2	1	0.004 8
17-32.2	1	0.004 8	20-29	1	0.004 8
17-32.3	1	0.004 8	20-29.2	1	0.004 8
17-33	5	0.024 2	20-30	1	0.004 8
17-34	1	0.004 8	20-30.2	1	0.004 8
18-28.2	3	0.014 5	20-31	2	0.009 7
18-29.2	2	0.009 7	20-31.2	1	0.004 8
18-30	3	0.014 5	20-32	2	0.009 7
18-30.2	3	0.014 5	20-32.2	1	0.004 8
18-31	4	0.019 3	20-33	1	0.004 8
18-31.1	1	0.004 8	20-33.2	1	0.004 8
18-31.2	3	0.014 5	20-34	2	0.009 7
18-32	6	0.029 0	20-34.2	1	0.004 8
18-32.2	7	0.033 8	21-30.2	1	0.004 8
18-33	3	0.014 5	21-31	2	0.009 7
18-33.2	1	0.004 8	22-29	1	0.004 8

基因来自于父源。

在31个X-STR基因座中,共28个基因座表现出了高度的多态性^[17],即PIC>0.5,多态性较差的基因座有DXS7423、DXS7133、DXS6800。31个X-STR的女性群体累积个人识别概率为 $1-1.578\ 4\times 10^{-31}$,男性群体累积个人识别概率为 $1-4.566\ 4\times 10^{-19}$,可以作为常染色体STR的有力补充来满足个人识别的需求。31个X-STR三联体累积的平均父

权排除概率为 $1-2.281\ 9\times 10^{-17}$,二联体累积的平均父权排除概率为 $1-2.509\ 0\times 10^{-12}$,可以有效地辅助常染色体STR进行复杂亲缘关系的鉴定。

本次研究的北方汉族女性群体31个X-STR基因座的基因型结果经连锁不平衡检验,在Bonferro-ni校正($P<0.05/465=0.000\ 107\ 526\ 9$)后,仍有1对基因座(DXS10103-DXS10101)具有显著性差异,存在连锁不平衡现象。因此,在案件涉及北方汉族的,

表6 北方汉族群体与其他四个群体之间X-STR基因座等位基因频率分布的差异检验结果
 Table 6 Difference test results of allele frequency distribution of X-STR loci between northern Han population and other four populations

Locus	Germany	Japanese	Guangdong Han	Sichuan Tibetan
DXS101	-	0.884 3	-	0.903 9
DXS981	-	0.000 0	-	-
DXS6789	0.000 0	0.014 5	-	0.271 3
DXS6800	-	0.046 4	-	-
DXS6803	-	0.000 1	-	-
DXS6807	-	0.522 8	-	-
DXS6809	0.002 1	0.663 7	-	0.419 7
DXS7132	0.000 0	0.030 9	0.746 3	0.996 2
DXS7133	-	0.017 1	-	-
DXS7423	0.000 0	0.342 2	0.043 3	0.869 9
DXS7424	-	0.000 0	-	0.731 7
DXS8378	0.000 0	-	0.469 4	0.000 0
DXS10074	0.000 0	-	0.332 4	0.066 3
DXS10075	-	-	-	0.006 7
DXS10134	0.000 0	-	0.538 3	0.379 1
DXS10079	0.000 0	-	0.943 4	0.036 9
DXS10101	0.000 0	-	0.095 3	0.251 8
DXS10103	0.000 0	-	0.982 6	0.362 1
DXS10135	0.000 0	-	0.364 9	0.768 6
DXS10148	0.000 0	-	0.000 5	0.024 5
DXS10159	-	-	-	0.194 6
GATA31E08	-	0.222 9	-	-
DXS10162	-	-	-	0.333 1
DXS10164	-	-	-	0.022 6
GATA165B12	-	-	-	-
HPRTB	0.003 8	0.068 0	0.034 7	0.000 0
GATA172D05	-	0.000 1	-	-

Significance level $\alpha = 0.05$. The black bold indicates significant difference.

统计分析中DXS10103-DXS10101按照单倍型频率来计算。同时,本次连锁不平衡检验也反映出,虽然在X染色体上报导了4个连锁群,但是实际检验出来的X-STR连锁不平衡并没有理论上的多,即使是位于同一个连锁群内,也有可能不存在连锁不平衡的情况,这在其他研究中也有类似的报道^[18]。

不同群体之间的显著性差异研究可以发现,北

方汉族群体与德国群体存在明显差异,与日本群体、四川藏族群体、广东汉族群体均存在一定程度上的差异。人种差异越大、地域分布越广泛,群体之间的差异也越大。这可能源于更为相近的地理距离或更为接近的种群在遗传上更具有同源性。因此在实践中应用X-STR时选择相对应的群体是至关重要的。

参考文献

- [1] 曾祥培, 李海霞, 孙宏钰. X-STR基因座的法医学应用研究进展[J]. 中国司法鉴定, 2010, 1: 58-62.
Zeng XP, Li HX, Sun HY. The research progress of X chromosome STR loci in forensic genetics [J]. Chin J Forensic Sci, 2010, 1: 58-62.
- [2] 吴小洁, 孙宏钰, 张胤鸣, 等. 姐妹关系鉴定中X-STR分型的应用2例[J]. 中国法医学杂志, 2011, 26(6): 479-480.
Wu XJ, Sun HY, Zhang YM, et al. Two cases of application of X-STR typing in sister relationship identification[J]. Chin J Forensic Med, 2011, 26(6): 479-480.
- [3] 张胤鸣, 刘素娟, 吴小洁, 等. X-STR分型用于祖孙关系鉴定1例[J]. 法医学杂志, 2012, 28(3): 236-237.
Zhang YM, Liu SJ, Wu XJ, et al. A case of X-STR typing used to identify the relationship between grandparents and grandchildren[J]. J Forensic Med, 2012, 28(3): 236-237.
- [4] Liu YX, Zhang WQ, Jia YS, et al. Multistep microsatellite mutation in a case of non-exclusion parentage [J]. Forensic Sci Int Genet, 2015, 16: 205-207.
- [5] Diegoli TM. Forensic typing of short tandem repeat markers on the X and Y chromosomes[J]. Forensic Sci Int Genet, 2015, 18: 140-151.
- [6] Hering S, Klimova A, Edelmann J. German population data for 18 X-STRs: a hexaplex PCR adding two clusters of X-STRs to the Argus X-12 set and expanding the German haplotype databases [J]. Int J Legal Med, 2020, 134(6): 2061-2062.
- [7] Nakamura Y, Minaguchi K. Sixteen X-chromosomal STRs in two octaplex PCRs in Japanese population and development of 15-locus multiplex PCR system [J]. Int J Legal Med, 2010, 124(5): 405-414.
- [8] 任峥, 曾祥培, 陈文静, 等. 十二个X染色体STR基因座荧光标记复合扩增体系的法医学应用[J]. 中山大学学报(医学科学版), 2011, 32(1): 109-115.
Ren Z, Zeng XP, Chen WJ, et al. Forensic application of twelve X chromosome str fluorescence-labeled multiplex amplification system [J]. J Sun Yat-Sen Univ (Med Sci), 2011, 32(1): 109-115.
- [9] He G, Li Y, Zou X, et al. X-chromosomal STR-based genetic structure of Sichuan Tibetan minority ethnicity group and its relationships to various groups [J]. Int J Legal Med, 2018, 132(2): 409-413.
- [10] García MG, Gusm OL, Catanesi CI, et al. Mutation rate of 12 X-STRs from investigator Argus X-12 kit in Argentine population [J]. Forensic Sci Inter Gene Suppl, Series 6 2017: e562-564.
- [11] Diegoli TM, Linacre A, Schanfield MS, et al. Mutation rates of 15 X chromosomal short tandem repeat markers [J]. Int J Legal Med, 2014, 128(4): 579-587.
- [12] 李运丽, 宋诚诚, 郝世诚, 等. 高变异Y染色体短串联重复片段检验体系的建立及对北方汉族群体调查[J]. 中山大学学报(医学科学版), 2021, 42(4): 550-556.
Li YL, Song CC, Hao SC, et al. Establishment of high mutation Y-Chromosome short tandem repeat multiplex amplification system and investigation on Han ethnic group in northern China [J]. J Sun Yat-Sen Univ (Med Sci), 2021, 42(4): 550-556.
- [13] Brinkmann B, Klitschar M, Neuhuber F, et al. Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat [J]. Am J Hum Genet, 1998, 62(6): 1408-1415.
- [14] Weber JL, Wong C. Mutation of human short tandem repeats [J]. Hum Mol Genet, 1993, 2(8): 1123-1128.
- [15] Wu W, Ren W, Hao H, et al. Mutation rates at 42 Y chromosomal short tandem repeats in Chinese Han population in eastern China [J]. Int J Legal Med, 2018, 132(5): 1317-1319.
- [16] 侯一平, 丛斌, 王保捷, 等. 法医物证学[M]. 人民卫生出版社, 2020年, 95.
Hou YP, Cong B, Wang BJ, et al. Forensic genetics [M]. People's Med Publ Press, 2020, 95.
- [17] Chen M, Ren H, Liu ZY, et al. Genetic polymorphisms and mutation rates of 16 X-STRs in a Han Chinese population of Beijing and application examples in second-degree kinship cases [J]. Int J Legal Med, 2020, 134(1): 163-168.
- [18] Szibor R. X-chromosomal markers: past, present and future [J]. Forensic Sci Int Genet, 2007, 1(2): 93-99.