

Expressmarker22系统检测广东汉族人群中的标准梯度外等位基因

马晓燕, 李少英, 何文智, 王燕超, 王晓蔓, 冼嘉嘉, 黎青

(广州医科大学附属第三医院 法医物证司法鉴定所//广东省普通高校生殖与遗传重点实验室, 广东 广州 510150)

摘要:【目的】分析 Expressmarker22 (EX22) 系统检测广东汉族人群中的梯度标准物范围之外(OL)等位基因频率及序列组成, 完善频率数据库。【方法】应用 EX22 系统对 3495 例广东汉族无关个体进行 STR 分型, 统计等位基因分型 OL 的等位基因频率并筛选稀有等位基因。通过比对 STRBase 数据库和相关文献, 对暂未报道的 OL 样本进行单基因座扩增及克隆测序。【结果】共发现分布于 10 个基因座中的 33 种 OL 等位基因, 频率为 0.3%~4.6%, 其中稀有等位基因 25 种, 11 个未见报道。测序结果显示以核心序列的不完整重复多见。【结论】OL 等位基因数据有助于提高个体识别率和非父排除率, 有效补充法医 DNA 多态性数据, 完善数据库建设。

关键词: Expressmarker22; 短串联重复序列; 梯度标准物范围之外的等位基因; 稀有等位基因; 测序
中图分类号: R394 **文献标志码:** A **文章编号:** 1672-3554(2018)02-0227-06

Off-ladder Alleles Detected by Expressmarker22 Kit in Guangdong Han Population

MA Xiao-yan, LI Shao-ying, HE Wen-zhi, WANG Yan-chao, WANG Xiao-man, XIAN Jia-jia, LI Qing
(Forensic Identification Institute//Key Laboratory of Reproduction and Genetics of Guangdong Higher Education Institutes, The Third Affiliated Hospital of Guangzhou Medical University, Guangzhou 510150, China)

Corresponding to: LI Qing, E-mail: 81292522@163.com

Abstract: 【Objective】 To study the rare alleles frequencies and sequences of Expressmarker 22 kit in Guangdong Han Population. 【Methods】 3495 Samples from unrelated individuals in Guangdong Han Population were screened by using AGCU Expressmarker 22 kit (EX22) and ABI 3100 Genetic Analyzer. Then analyzed the frequencies of the off-ladder (OL) alleles and sequenced the rare alleles obtained based on comparison with the STRBase database and literature. 【Results】 33 off-ladder alleles with 25 rare alleles were found in 10 STR loci, and allele frequencies ranged from 0.0003~0.0046. Sequencing of the 11 unreported rare OL alleles showed that most of them have incomplete repeats. 【Conclusion】 Off-ladder alleles especially the rare alleles are helpful to improve the power of discrimination and the power of exclusion and to provide samples which will be added its allele into ladders These OL-alleles will supplement forensic DNA database.

Key words: EX22; short tandem repeat; off-ladder allele; rare allele; sequence

[J SUN Yat-sen Univ (Med Sci), 2018, 39(2): 227-232]

短串联重复序列(short tandem repeat, STR)是目前法医 DNA 实验室运用最广泛的遗传标记。随着商业化试剂盒的广泛应用, STR 分型数据在各实验室间得以共享, 但由于 STR 基因座核心序

列短, 易出现变异, 故在不同地区和种族人群中等位基因分布有一定的差异。不同实验室相继发现常用 STR 试剂盒的某些基因座检出等位基因梯度标准物范围之外的等位基因(off-ladder allele),

收稿日期: 2017-12-20

基金项目: 广东省科技计划项目(2013B022000023)

作者简介: 马晓燕, 硕士, 主检法医师/技师, 专业方向: 法医遗传学, E-mail: yalena_m@foxmail.com; 黎青, 通信作者, 博士, 主任法医师/主任技师, 硕士生导师, 研究方向: 遗传学/妇产科学, E-mail: 81292522@163.com

OL)^[1-11]。这些罕见等位基因在人群中频率极低,罕见等位基因的出现,尤其是落入邻近基因座的OL等位基因造成的假三带型给检测结果判读带来一定困难。目前已有的一些关于PowerPlex16(PP16, Promega公司)体系或单个基因座OL等位基因的报道^[1-3,5],但国内广泛应用的国产试剂盒AGCU Expressmarker 22(EX22,无锡中德美联生物技术有限公司)的OL等位基因的研究数据尚缺乏。本文就EX22试剂盒在广东汉族人群中的OL等位基因进行系统分析,以期对广东汉族群体遗传多态性数据进行有效的补充和完善,为该试剂盒在实际应用中出现的复杂情况的判读提供参考。

1 材料与amp;方法

1.1 材料及DNA提取

收集本实验室无关个体样本共3 495例,所有个体均来源于广东地区汉族人群。样本为血FTA卡(英国Whatman公司)或毛发、指甲、羊水、绒毛等组织。血FTA卡样本直接扩增;头发、指甲、羊

水、绒毛等样本按照常规Chelex-100法(美国Bio-Rad公司)提取DNA。

1.2 扩增及OL样本确认

扩增及分型按照EX22荧光检测试剂盒(无锡中德美联生物技术有限公司)说明书进行,扩增产物在ABI3100型遗传分析仪上电泳,最后由GeneMapper v3.2软件自动分析,得到被测样品各基因座的分型数据。对仪器自动判读为“OL”等位基因的样本重新用该试剂盒或其他同类试剂盒进行复查,若结果和初检一致,则确定为OL等位基因样本。

1.3 单基因检测样本的筛选

将1.2中确定的OL等位基因与STRBase数据库(<http://www.cstl.nist.gov/strbase>)和相关文献报道的OL进行比对后,筛选暂未或鲜少有报道的(出现例数不超过2例)OL等位基因样本,进行单基因座扩增、检测和克隆测序。

1.4 单基因座扩增及克隆测序

1.4.1 单基因座扩增引物 引物序列见表1,引物由大连宝生物公司合成。

表1 单基因座检测的大片段克隆引物
Table 1 Primers for single locus cloning sequencing

Locus	Primer-F(5'-3')	Primer-R(5'-3')	Amplicons range/bp
Penta E	TTAAAGTTGACGTCTCATTGCATTG	TAGCCTCAAACATCAGTTTCCTCT	574-694
D8S1179	GACCACACGGCCAAGTAGAAG	TGGAGAAACTGAAACCCGTGTC	483-535
D2S441	AAGAGATTCTTAAGACCCACGG	AGAGACAGGGTCTCCCTATGTTG	473-509
D6S1043	GCCCTTCTTCTCCTGCTCTTGAACATA	CTCAATCAGACCCTCATCTCCTCA	611-675
D18S51	ACTAACAAATAGGCCAAGCGTGA	TAAAAATAACAAACCCGACTACC	509-593
D19S433	GCAATGTATAGGTCGTTTC	CGGCTCTGATTAATAATGCTACA	872-904

1.4.2 扩增体系及参数 使用AGCU EX22试剂盒的反应组分,终体系为25 μ L,终浓度为:1 \times Reaction Mix,4 U热启动C-Taq酶,0.4 μ mol/L引物对,基因组DNA约为30~50 ng;PCR反应条件为:95 $^{\circ}$ C 11 min;94 $^{\circ}$ C 1 min,59 $^{\circ}$ C 1 min,72 $^{\circ}$ C 1 min,30个循环;60 $^{\circ}$ C 1 h;扩增产物经1%琼脂糖凝胶电泳以确认目标片段。

1.4.3 PCR产物Ta克隆测序 使用大连宝生物工程有限公司生产的pMD18-T连接体系对PCR产物进行连接,并转入大肠杆菌DH5- α 感受态细胞中铺平板,在平板中挑取单菌落进行标记。对培养的菌液进行PCR,终体系为10 μ L,终浓度

为:1 \times Reaction Mix,1.5 U热启动C-Taq酶,0.4 μ mol/L引物对,菌液DNA 1 μ L;PCR反应条件为:95 $^{\circ}$ C 11 min;94 $^{\circ}$ C 30 s,59 $^{\circ}$ C 30 s,72 $^{\circ}$ C 30 s,30个循环;60 $^{\circ}$ C 1 h;PCR产物经聚丙烯酰胺凝胶电泳验证所挑取的克隆是否是所需的等位基因克隆。如为目的等位基因克隆,则对菌液中片段进行测序。此部分工作由无锡中德美联生物技术有限公司协助完成。

2 结果

3 495例样本在21个常染色体STR基因座中

共发现33种OL等位基因,分布于PentaE、PentaD、D2S441、D7S820、D2S1338、D6S1043、D21S11、D18S51、D8S1179、D19S433共10个基因座,各基因座OL等位基因的种类为1~6种,其中在PentaE基因座发现的OL等位基因种类及数量最多。基因频率为0.000 3~0.004 6(表2),部分分型图谱见图1;有25种属于稀有等位基因(基因频率<0.1%),11个OL暂未或鲜少有报道,分别为D18S51-“17.3”、“28”;D8S1179-“12.2”;PentaE-“25.4”;D6S1043-

“12.3”、“17.3”、“18.2”、“22.3”;D2S441-“10.1”、“12.3”;D19S433-“4”。根据片段大小和电泳漂移率计算出的OL等位基因值显示33种OL中仅有6种为核心序列的完整重复,而其余17种OL等位基因类型均为核心序列的不完整重复。对11个暂未或鲜少有报道的OL进行测序的结果显示,序列核心重复次数与根据片段大小和漂移率计算出的等位基因大小一致(表3)。

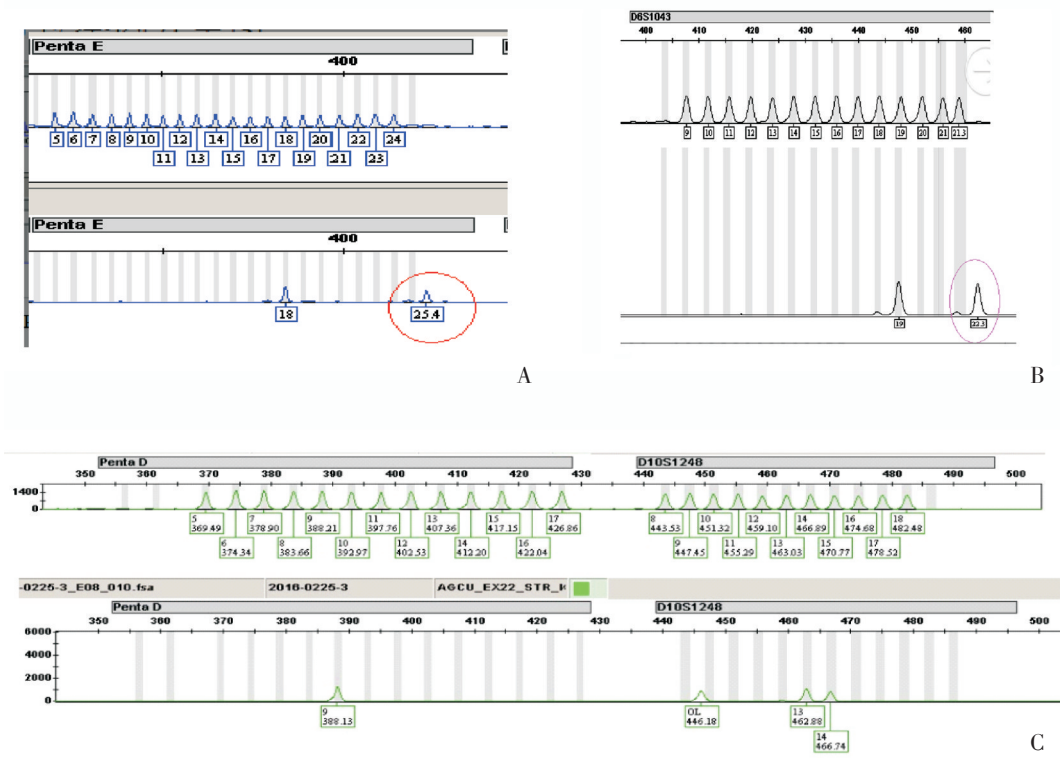
表2 各基因座OL等位基因的种类及频率
Table 2 The type and frequency of OL alleles

(n=3 495)

Locus	OL-allele			Locus	OL-allele			Locus	OL-allele		
	Type	n	Frequency		Type	n	Frequency		Type	n	Frequency
PentaE	14.2	1	0.000 3	PentaD	11.2	1	0.000 3	D2S441	10.1	6	0.001 7
	18.4	6	0.001 7		12.1	1	0.000 3		11.1	2	0.000 6
	19.4	9	0.002 6		19	2	0.000 6		12.3	2	0.000 6
	20.4	1	0.000 3		20	1	0.000 3		14.1	2	0.000 6
	25.4	1	0.000 3		20.4	2	0.000 6		15.1	1	0.000 3
	27	1	0.000 3		21	2	0.000 6				
D7S820	9.1	16	0.004 6	D6S1043	12.3	5	0.001 4	D8S1179	12.2	1	0.000 3
	9.2	1	0.000 3		17.3	2	0.000 6				
	9.3	1	0.000 3		18.2	2	0.000 6				
	10.1	4	0.001 1		22.3	4	0.001 1		D2S1338	24.1	1
D21S11	30.3	15	0.004 3	D18S51	17.3	1	0.000 3	D19S433	4	1	0.000 3
	33.1	1	0.000 3		28	1	0.000 3		12.1	1	0.000 3

表3 11种OL等位基因序列分析
Table 3 The sequence analysis of 11 rare alleles

Locus	OL-allele	Core Sequence	OL-allele sequence type
D18S51	17.3	(AGAA)11 AAA(AGAA)6	II
	28	(AGAA)28	I
D8S1179	12.2	(TCTA)1(TCTG)1(TCTA)1TA(TCTA)9	II
PentaE	25.4	AAGA(AAAGA)25	II
D6S1043	12.3	(AGAT)7 GAT(AGAT)5	II
	17.3	(AGAT)8 GAT(AGAT)3(ACAT)1(AGAT)5	II
	18.2	(AGAT)12AT(ACAT)1(AGAT)5	II
	22.3	(AGAT)14GAT(AGAT)2(ACAT)1(AGAT)5	II
D2S441	10.1	A(TCTA)10	II
	12.3	(TCTA)4TCA(TCTA)8	II
D19S433	4	AAGG AAAG AAGG TAGG (AAGG)2	I



A: Locus Penta E OL 25.4, B: Locus D6S1043 OL 22.3, C: Locus Penta D OL 21

图1 部分OL等位基因的分型图谱

Fig.1 The profiling of OL alleles

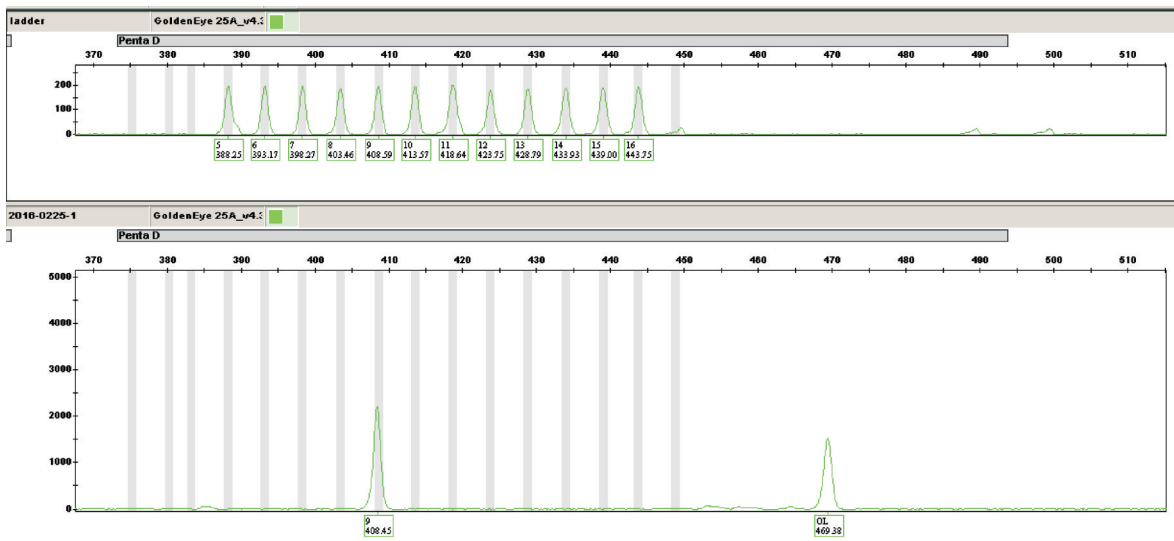


图2 Penta D基因座OL等位基因(图1C)的验证

Fig.2 The certification profiling of OL alleles in Penta D locus

3 讨论

3.1 OL等位基因类型

等位基因分型标准物(Allelic ladder)是指某

一STR基因座在人群中常见的等位基因的混合物,是STR基因座分型的标准。检测时,将检测样品和Ladder同时进行电泳,若检测样品的等位基因在ladder范围内,分型软件会自动识别为对应的等位基因。若等位基因在范围外或者在ladder

两个等位基因之间,则识别为off-ladder(OL)。

OL等位基因主要有两种类型^[3]:(I)核心序列的完整重复:如D18S51的“28”,等位基因核心序列为(AGAA)28;这一类在图谱上通常是位于两个邻近基因座ladder之间或某个基因座ladder之外。本次调查发现的多数OL属于核心序列的不完整重复。根据陆惠玲^[3]、Anne Deucher^[12]等的报道,等位基因侧翼序列的碱基缺失和插入也会引起STR长度的改变,这类比较少见,本调查目前未观察到侧翼变异导致的OL。(II)核心序列的不完整重复:这一类等位基因在核心重复序列中出现个别碱基的缺失或插入,如D18S51的“17.3”,等位基因序列为(AGAA)11 AAA (AGAA)6,其核心重复序列“AGAA”重复了17次,在中间第12个重复处缺失了一个“G”;在图谱上多出现在ladder的两个等位基因之间。第II类:重复单位不完整重复,位于两个相邻等位基因之间;

3.2 OL等位基因对图谱判读的影响

图谱中出现OL时,首先要排除杂峰的影响,如stutter峰、DNA过浓导致的拔起峰峰、电泳系统杂质导致的干扰峰以及电泳环境不当导致的漂移等。在排除了各种情况之后,仍然有一种特殊的OL会影响图谱的判断,即因OL造成的假三带型。如图1C中,D10S1248基因座为三带型,其中有一个OL峰。但观察相邻的PentaD基因座,发现该基因座为较矮的单峰,不符合纯合子单峰高于杂合子双峰的规律,此时应考虑该OL属于PentaD基因座的可能性,需更换其他PentaD基因座毗邻关系的检测体系进行验证。本例通过片段

大小和漂移率计算出该OL为PentaD的“21”,同时我们采用Goldeneye25A试剂盒进行验证,结果一致(图2)。类似等位基因落入相邻基因座区域造成误判的现象其他文献也报道过^[13-14]。我们目前观察到类似PentaD基因座OL等位基因造成D10S1248基因座假三带型的案例共3例,这也提示我们,在应用EX22试剂盒时,遇到PentaD和D10S1248基因座上的OL等位基因或三带型的判读需谨慎。

3.3 OL等位基因发现的意义

在本次调查中发现,在21个STR基因座中,PentaE和PentaD基因座所含的OL等位基因数目较多,均为6种;其次为D2S441、D7S820和D6S1043基因座,分别为5、4和4种;这一结果与以往的报道^[1-5]既相互印证又进行补充,除PowerPlex系统所含PentaE、PentaD和D7S820基因座以外,我们还发现D2S441和D6S1043基因座在广东人群中存在较多的OL等位基因。其中多数为稀有等位基因(频率<0.1%),这些OL等位基因将有助于提高个体识别率和非父排除率。

在33个OL中,有8个OL等位基因的频率超过了稀有等位基因的范畴。其中D7S820基因座的OL等位基因9.1在广东汉族人群中的基因频率为0.0046,D21S11基因座的等位基因30.3在广东汉族人群中的基因频率为0.0043,远大于稀有等位基因的评定标准。提示我们,在ladder及识别软件中添加这些OL等位基因尤其是频率较高的OL等位基因,将有助于提高检测效能,方便群体数据在实验室间的共享^[15]。

参考文献

- [1] Mizuno N, Sekiguchi K, Sato H, et al. Variant alleles on the PentaE locus in the PowerPlex™ 16 kit [J]. *J Forens Sci*, 2003, 48(2): 358-361.
- [2] 曾艳红,孙宏钰,童大跃,等. Powerplex™ 16体系在中国人群中罕见等位基因及其类型[J]. *中国法医学杂志*, 2004, 19(2): 78-79.
Zeng YH, Sun HY, Tong DY, et al. The frequencies of rare alleles in Chienes population using PowerPlex™ 16 System [J]. *Chin J Forens Med*, 2004, 19(2): 78-79.
- [3] 陆惠玲,台运春,刘超,等. Powerplex™ 16体系OL等位基因序列分析及命名探讨[J]. *法医学杂志*, 2006, 22(3): 186-189.
Lu HL, Tai YC, Liu C, et al. Sequences of Off-ladder Alleles of PowerPlex™ 16 Kit in Chinese Han Population [J]. *J Forens Med*, 2006, 22(3): 186-189.
- [4] René LM, Lara B, Kamelija MT, et al. Variant alleles, triallelic patterns, and point mutations observed in nuclear short tandem repeat typing of populations in Bosnia and Serbia [J]. *Croat Med J*, 2007, 48(4): 494-502.

- [5] 刘秋玲, 梁艳芳, 吕德坚, 等. Powerplex®16™ System 在中国人群中OL等位基因的研究[J]. 中山大学学报(医学科学版), 2009, 30(3S): 29-32.
Liu QL, Liang YF, Lv DJ, et al. Development of OL Allele of Powerplex®16™ System in China [J]. J Sun Yat-Sen Uni (Med Sci), 2009, 30(3S): 29-32.
- [6] 赵英, 徐卫华, 符生苗. 序列分析鉴定等位基因D18S1345-10[J]. 中华医学遗传学杂志, 2010, 27(6): 704-706.
Zhao Y, Xu WH, Fu SM. Identification of the new allele D18S1364-10 by sequence analysis [J]. Chin J Med Genetics, 2010, 27(6): 704-706.
- [7] 章雅清, 陈维红, 陆惠玲. 广东汉族人群罕见等位基因FGA-13的研究[J]. 中华医学遗传学杂志, 2013, 30(6): 742-744.
Zhang YQ, Chen WH, Lu HL. Identification of a rare allele FGA-13 in Guangdong Han population [J]. Chin J Med Genetics, 2013, 30(6): 742-744.
- [8] 欧阳曙明, 杨旭, 申琴, 等. 应用Identifiler™系统对湖南人群中OL等位基因的观察与分析[J]. 生命科学研究, 2013, 17(3): 193-195.
Ouyang SM, Yang X, Shen Q, et al. Analysis of OL Allele of Identifiler™ System in Hunan Population [J]. Life Sci Res, 2013, 17(3): 193-195.
- [9] 余建华, 高静, 向超杰, 等. 云南汉族人群15个常染色体STR基因座的遗传多态性与OL等位基因的研究[J]. 昆明医科大学学报, 2014, 35(7): 27-31.
Yu JH, Gao J, Xiang CJ, et al. Genetic variability and off-ladder allele of 15 autosomal STR loci in Yunnan Han Population of China [J]. J Kunming Med Univ, 2014, 35(7): 27-31.
- [10] 黄洪武, 徐振亮, 朱诣琦. D13S317基因座“off-ladder”等位基因分析[J]. 沈阳医学院学报, 2013, 15(4): 205-207.
Huang HW, Xu ZL, Zhu YQ. Analysis of off-ladder allele found on the D13S317 locus [J]. J Shenyang Med College, 2013, 15(4): 205-207.
- [11] 陆慧洁, 陈玲, 邱平明. D21S11稀有等位基因落入相邻基因座1例[J]. 法医学杂志, 2016, 32(3): 239-240.
Lu HJ, Chen L, Qiu PM. 1 cases with rare alleles of D21S11 locus fall into adjacent loci [J]. J Forensic Med, 2016, 32(3): 239-240.
- [12] Anne D, Tsoyu C, Iris S. Rare sequence variation in the genome flanking a short tandem repeat locus can lead to a question of “nonmaternity” [J]. J Mol Diagn, 2010, 12(3): 384-389.
- [13] Chen W, Cheng J, Ou X, et al. Identification of the sequence variations of 15 autosomal STR loci in a Chinese population [J]. Ann Hum Biol, 2014, 41(6): 524-530.
- [14] Chen W, Cheng J, Tong D, et al. Identification of a rare off-ladder allele of the D13S325 locus during paternity testing [J]. Leg Med (Tokyo), 2014, 16(1): 48-51.
- [15] Tvedebrink T, Bright JA, Buckleton JS, et al. The effect of wild card designations and rare alleles in forensic DNA database searches [J]. Forens Sci Int Genet, 2015, 16(8): 98-104.

(编辑 孙慧兰)