

用 STR 进行亲权鉴定时结果的判定及其理论根据

朱运良, 欧雪玲, 蔡贵庆, 伍新尧

(中山大学中山医学院法医学系, 广东 广州 510080)

摘要: **【目的】**探讨亲权鉴定中应该应用的 STR 基因座的数量及出现矛盾基因座时下结论的标准。**【方法】**根据基因频率和遗传规律,用计算机模拟 1 万例亲权鉴定,统计父权指数。根据平均非父排除率和突变率,用二项分布公式计算真父和随机男人出现不同数目矛盾基因座的概率和似然比。**【结果】**在 1 万例模拟亲权鉴定中,当检测 9 个和 13 个 STR 基因座时,分别有 2% 和 0.06% 的 PI 小于 999。检测 17 个 STR 基因座,真父和随机男人出现 1 个矛盾基因座的似然比为 2.3×10^3 。检测 24 个 STR 基因座,真父和随机男人出现 2 个、3 个和 4 个矛盾基因座的似然比分别为 1 087、1.408 和 0.002。**【结论】**亲权鉴定中应该检测 13 个以上的 STR 基因座。出现 1 个和 2 个矛盾基因座时,应该分别增加检测到 17 和 24 个基因座,才能下肯定的结论。检测到 3 个矛盾基因座时,应该用其它方法作进一步的分析。出现 4 个矛盾基因座时,可排除父权关系。

关键词: 亲权鉴定;短串联重复;父权关系指数

中图分类号: D919.2

文献标识码: A

文章编号: 1672-3554(2004)06-0534-04

How to Draw a Conclusion in Parentage Testing Using Short Tandem Repeats as Genetic Markers

ZHU Yun-liang, OU Xue-ling, CAI Gui-qing, WU Xin-yao

(Department of Forensic Medicine, SUN Yat-sen University, Guangzhou 510080, China)

Abstract: **【Objective】**To discuss how many STR loci should be typed in routine parentage testing and how to draw a conclusion if there are inconsistent loci. **【Methods】**Based on allele frequency and the law of inheritance, the paternity indexes of ten thousand cases of computer simulating parentage testing were calculated and summarized. The probabilities of inconsistency under paternity and non-paternity were calculated respectively according to binomial distribution based on the likelihood ratio of true father to random man. **【Results】**If 9 and 13 loci were typed, the percentages of PI less than 999 were 2% and 0.06%, respectively. The likelihood ratio of true father to random man producing one inconsistent locus was 2.3×10^3 if 17 STR loci were used. If there were 2, 3, and 4 inconsistent loci, the likelihood ratios were 1087, 1.408, and 0.002, respectively, after 24 STR loci were typed. **【Conclusion】**It is recommended to type at least 13 loci of STR in routine parentage testing. Seventeen or 24 typed loci are necessary to draw the conclusion when 1 or 2 inconsistent loci are appeared respectively. Further test should be done in order to make a conclusion if 3 inconsistent loci are observed. Four inconsistent loci are necessary to render an opinion of non-paternity.

Key words: paternity testing; short tandem repeats; paternity index

[J SUN Yat-sen Univ (Med Sci), 2004, 25(6): 534 - 537]

STR 的应用极大地提高了亲权鉴定能力。据估计,人类基因组中存在着 20 万个左右的 STR 基因

座^[1],可以应用于亲权鉴定的 STR 基因座也被不断地报道。STR 分型已经成为亲权鉴定的主要手

收稿日期: 2004-01-05

基金项目: 广东省自然科学基金资助项目(79-128);中山大学“211 工程”重点学科资助项目(429008)

作者简介: 朱运良(1963-),男,河南洛阳人,博士生,讲师,伍新尧,通讯作者。E-mail: xyaow@gzsums.edu.cn

段^[2]。我国多数实验室也采用这一方法。但是,亲权鉴定时要检测多少个 STR 基因座才能做出判断,这一问题还没有统一的标准。父权概率达到 0.9973 的父权认定标准^[3]是否仍然适用,也没有从理论上进行探讨和论证。STR 基因座的等位基因数多,多态性高,突变率较高^[2, 4],在亲权鉴定中,遇到突变引起的矛盾基因座也较多。究竟出现多少个矛盾基因座可以下排除亲权关系的结论?这个问题也还没有统一的认识。出现矛盾基因座而其数目又达不到排除的标准时,要考虑矛盾基因座是由于突变所产生。此时要考虑突变因素,计算父权指数 (paternity index, PI) 和父权概率,如果父权概率能达到认定标准,就认定有亲生关系。但是,如何计算矛盾基因座的 PI 值还没有一个公认的方法。据美国血库联合会 (American association of blood banks, AABB)^[2]对 46 个亲权鉴定实验室的统计,突变基因座的 PI 计算方法有以下几种:①直接用突变率。②用突变率作为传递概率 (即用突变率作为 $PI = X/Y$ 中的 X)。③用一个常数 (0.002 或者 0.003)。④用突变率除以平均非父排除概率 (64% 的实验室用这一方法)。另外还有 Ayres^[5]介绍的方法。为了找到一个较为合理、科学的亲权鉴定判断标准,我们结合有关数据和我室多年的实践经验,参考国外相关做法,运用概率学理论,对亲权鉴定时需要检测的 STR 基因座的特性和数量,肯定父权关系时父权指数 (或父权概率) 的标准,以及出现矛盾基因座时检验以及下结论的策略等问题作一探讨。

1 材料和方法

1.1 基因座

9 个基因座包括 D3S1358、vWA、FGA、D8S1179、D21S11、D18S51、D5S818、D13S317、D7S820。13 个基因座指的是 13 个 CODIS (combined DNA index system) 基因座 (除去以上 9 个基因座外,加上 D16S539、TH01、TPOX 和 CSF1PO)。15 个基因座则指 13 个 CODIS 加上 PentaE 和 PentaD。17 个基因

座和 24 个基因座分别再加上基因座 D8S384、D11S554 和 D8S384、D11S554、D12S391、PLA、D13S631、FES/FPS、D19S253、D20S85 和 D10S89。

1.2 所用基因座的群体数据和参数

各个基因座的等位基因频率和非父排除概率均采用我实验室的数据。平均非父排除概率用 15 个基因座 (13 个 CODIS 基因座加上 PentaE 和 PentaD) 的非父排除概率的平均值 (0.5714)。CODIS 13 个基因座的突变率来源于 AABB 2002 年年报^[2]。平均突变率用这 13 个位点的突变率的平均值 (0.0017) 计算均假设母子关系确定,需鉴定孩子的亲生父亲。计算父权概率时假设父权前概率为 0.5。

1.3 计算出现矛盾基因座的概率

设共检查了 n 个基因座,其中在 m 个基因座上产生了矛盾。 m 服从二项分布。根据二项分布公式 $P(m) = C_n^m (1-\pi)^{n-m} \pi^m$ (P 为概率, C_n^m 为从 n 个元素中抽取 m 个元素的组合) 分别计算当 $n = 9, 13, 15, 17, 24$ 时,存在亲生关系的情况下和不存在亲生关系的情况下产生 0, 1, 2, 3 和 4 个矛盾基因座的概率。 π 分别等于 0.5714 (不存在亲生关系时) 和 0.0017 (存在亲生关系时)。

1.4 计算和统计

计算机模拟鉴定用 Microsoft Excel 进行。计算和统计用 Microsoft Excel 和 SPSS 进行。

2 结果

2.1 计算机模拟检验

计算机模拟检验及 PI 的计算按照如下方法进行:每个基因座中,以基因频率作为抽样概率,分别抽出母亲的 2 个等位基因,组成母亲基因型。随机抽出这 2 个等位基因中的 1 个,作为孩子的 1 个基因,再以基因频率作为抽样概率抽出孩子的另 1 个基因,把它作为生父基因,和另外抽出的 1 个基因 (同样以基因频率作为抽样概率) 组成父亲基因型。计算各基因座 PI 和累积 PI。统计 1 万次计算机模拟鉴定的最小 PI、最大 PI 和 PI 中位数,结果如表 1。

2.2 有矛盾基因座时亲权关系的判定

表 1 1 万例计算机模拟检验所得到的 PI 值的描述统计

Table 1 The descriptive statistics of PI of ten thousand cases of computer simulation

Typed loci (n)	Minimum PI	Maximum PI	Median of PI	Percentage of PI < 999 (%)
9 loci	160	2.0×10^{10}	3.6×10^{-4}	2.0
13 loci	322	6.5×10^{13}	1.1×10^{-6}	0.06
15 loci	1 931	3.2×10^{15}	2.6×10^{-7}	0
17 loci	5 534	4.7×10^{16}	1.6×10^{-8}	0

亲权鉴定案出现矛盾基因座,可能是孩子和争议男子之间不存在父子关系;也可能有父子关系,矛盾基因座是由于突变所产生。根据平均非父排除率和平均突变率(分别是 15 个和 13 个基因座平均值,见材料与方法),用二项式定理分别计算这两种情况下产生 0 到 4 个矛盾基因座的概率(计算方法见材料与方法)。结果如表 2。

表 2 在非父和父亲情况下出现不同数目矛盾基因座的概率

Table 2 The probability of inconsistencies under non-paternity and paternity

Typed loci (<i>n</i>)	Inconsistent loci (<i>n</i>)	P_{nt}	P_t	LR
9 loci	0	4.9×10^{-4}	9.8×10^{-2}	2.0×10^3
	1	5.9×10^{-3}	1.5×10^{-2}	2.6
	2	3.1×10^{-2}	1.0×10^{-4}	3×10^{-3}
	3	9.7×10^{-2}	4.3×10^{-7}	4.4×10^{-6}
	4	1.9×10^{-1}	1.1×10^{-9}	5.7×10^{-9}
13 loci	0	1.7×10^{-5}	9.8×10^{-1}	5.9×10^4
	1	2.9×10^{-4}	2.2×10^{-2}	7.7×10^1
	2	2.3×10^{-3}	2.0×10^{-4}	9.9×10^{-2}
	3	1.1×10^{-2}	1.4×10^{-6}	1.0×10^{-4}
	4	3.7×10^{-2}	6.2×10^{-9}	1.7×10^{-7}
15 loci	0	3.0×10^{-6}	9.7×10^{-1}	3.2×10^5
	1	6.1×10^{-5}	2.5×10^{-2}	4.4×10^2
	2	5.7×10^{-4}	3.0×10^{-4}	5.4×10^{-1}
	3	3.3×10^{-3}	2.3×10^{-6}	7.0×10^{-4}
	4	1.3×10^{-2}	1.2×10^{-8}	9.0×10^{-7}
17 loci	0	5.6×10^{-7}	9.7×10^{-1}	1.7×10^6
	1	1.2×10^{-5}	2.9×10^{-2}	2.3×10^3
	2	1.3×10^{-4}	4.0×10^{-4}	2.9
	3	9.0×10^{-4}	3.4×10^{-6}	3.8×10^{-3}
	4	4.2×10^{-3}	2.1×10^{-8}	4.9×10^{-6}
24 loci	0	1.5×10^{-9}	9.6×10^{-1}	6.5×10^8
	1	4.7×10^{-8}	4.0×10^{-2}	8.4×10^5
	2	7.3×10^{-7}	8.0×10^{-2}	1.1×10^3
	3	7.1×10^{-6}	1.0×10^{-5}	1.4
	4	5.0×10^{-5}	9.0×10^{-8}	2.0×10^{-3}

P_{nt} : The probability of inconsistency arising from sample error under condition that there is non-paternity; P_t : The probability of inconsistency arising from mutations under condition that there is paternity; LR: Likelihood ratio, P_t/P_{nt}

3 讨论

传统方法亲权鉴定方法(血型,红细胞酶型,血清蛋白型,同种异型和 HLA 等)的非父排除概率低,可用的基因座不多,多数基因座的分型方法较

烦琐,综合各方面的因素,制定了 99.73% 的父权概率(相当于 $PI = 367$)作为父权认定标准。STR 和其它 DNA 遗传标记在亲权关系鉴定中的应用,极大提高了亲权鉴定能力(表 1),因此应该相应提高亲权认定标准。如果以父权概率 99.9% ($PI = 999$)作为父权认定标准,用 STR 作为遗传标记是很容易达到的。从表 1 可以看出,检测 9 个 STR 基因座,仅有 2% 的案例达不到这个标准,而检测 13 个时只有 0.06% 达不到。从表 2 可以看出,在非父的情况下,由于抽样原因,检测到 0 个矛盾基因座(此时将产生错误结论)的概率在用 9 个基因座时为 4.9×10^{-4} ,在 13 个基因座时为 1.7×10^{-5} 。检测 9 个基因座,在非父的情况下,检测到少于 4 个矛盾基因座(亲权否定的界限,见后)的概率大于 13%,而检测 13 个基因座时这一数值为 1.4%。综合这些因素,借鉴一些国家和地区的做法,为保证鉴定结论不出差错,我们建议,普通的亲权鉴定中应该检测 13 个以上的 STR 基因座,以 99.9% 的父权概率作为父权认定的标准。STR 的突变率较高,根据 AABB 2002 年年报统计,13 个 CODIS 基因座的突变率在 0.02% (TH01 和 TPOX)和 0.34% (GWA)之间(平均 0.17%)。一般认为在出现 1 个、2 个甚至 3 个矛盾基因座时,不能排除有亲权关系^[2];也有人认为出现 2 个矛盾基因座时,可以排除亲权关系^[6]。表 2 中的 LR 可以像 PI 一样用来计算父权概率。从表 2 中可以看出,检测 17 个基因座出现 1 个矛盾基因座,检测 24 个基因座出现 2 个矛盾基因座时,所得到的 LR 均大于 999(父权概率大于 99.9%)。在这两种情况下将矛盾基因座考虑成由突变产生,在考虑突变的情况下计算 PI(方法如前所述),在鉴定书中报告突变基因座,并说明突变基因座的亲权指数计算方法。本室近几年来进行的 5124 例亲权鉴定中,共发现突变 113 例,其中 2 个基因座发生突变的 4 例,没有发现有 3 个矛盾基因座的案例。而据 AABB 统计,在 310 490 例亲权鉴定中,发现了 47 例(0.015%) 2 个基因座发生突变,5 例(0.002%) 3 个基因座发生突变(包括突变率更高的遗传标记,如小卫星 DNA)^[2]。表 2 中可以看出,检验 24 个基因座后,如果发现有 3 个矛盾基因座,LR 为 1.408,因此不能作出排除或者肯定的结论,应该用其它办法来进行进一步鉴定,如对突变形式进行分析,对矛盾基因座进行 DNA 测序,检测 Y 染色体上的 STR 等。根据表 2 中的数据,检测

17 个以下的基因座,出现 3 个矛盾基因座时,根据 LR 应该下排除结论。但是考虑到计算方法的局限性(未考虑突变形式),STR 等位基因的复杂性(如零等位基因, null allele),以及错误结论可能带来的严重后果,建议出现 3 个矛盾基因座时也要继续加做 STR 基因座,特别是矛盾等位基因均相差 1 个重复单位或者不能否定有零等位基因存在的情况下。如果出现 4 个矛盾基因座,则可以作出排除亲权关系的结论。本实验室近年来的 5 124 例亲权鉴定案例中,均做到 15 个 STR 基因座以上,在发现矛盾基因座时,均按照以上原则进行检验和鉴定,没有发现有争议的案例,这说明以上判断原则具有合理性和可靠性。如果出现 2 个矛盾基因座时否定亲权关系,可能会产生错误的结论。

现在,一般认为 STR 产生突变的机制是复制时的模板滑动^[7]。突变时产生 1 个重复单位的变化占突变总数的 90% 以上^[2, 4]。在本文中,没有考虑这方面的影响因素。如果在矛盾基因座上待排男子所能提供的基因和孩子的生父基因之间相差超过 1 个重复单位,更倾向于排除他们之间有亲生关系。另外在本文的计算过程中也没有考虑共祖系数的影响,如果遇到怀疑争议父(母)与真父(母)可能有亲缘关系时,需了解案情后再定具体的判断标准。

参考文献:

- [1] Collete D, Sabine F, cecle F, *et al.* A comprehensive genetic map of the human genome based on 5264 microsatellites[J]. *Nature*, 1996, 380(14): 152-4.
- [2] America Association of Blood Banks. Annual report summary for testing in 2002[DB/OL]. http://www.aabb.org/about_the_AABB/stds_and_Accred/ptann-rpt02.pdf. 2003-11-01.
- [3] 伍新尧,郭景元,陈跃龙,等. 亲权纠纷案的鉴定[J]. *中山医科大学学报*, 1987, 8(1): 77-9.
- [4] Brinkman B, klintschar 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): 1408-15.
- [5] Ayres K. Relatedness testing in subdivided populations [J]. *Forensic Sci Int*, 2000, 114(2): 107-15.
- [6] Thomson J A, Pilotti V, Stevens P, *et al.* Validation of short tandem repeat analysis for the investigation of disputed paternity[J]. *Forensic Sci Int*, 1999, 100(1): 1-16.
- [7] Levinson G, Gutman G A. Slipped strand mispriming: a major mechanism for DNA sequence evolution[J]. *Mol Biol Evol*, 1987, 4(3): 203-21.

(编辑 黄小延)

(上接第 533 页 from page 533)

associates with alpha-synuclein and promotes the formation of cytosolic inclusions[J]. *Nat Genet*, 1999, 22(1): 110-4.

- [5] 包新民,舒斯云. 大鼠脑立体定向图谱[M]. 北京: 人民卫生出版社, 1991. 49-58.
- [6] Engeler S, Wanner T, Kleiderlein J J, *et al.* Organization of the human synphilin-1 gene, a candidate for Parkinson's disease[J]. *Mamm Genome*, 2000, 11(9): 763-6.
- [7] Ribeiro C S, Carneiro K, Ross C A, *et al.* Synphilin-1 is developmentally localized to synaptic terminals and its association with synaptic vesicles is modulated by alpha-synuclein[J]. *J Biol Chem*, 2002 277(26): 23927-33.
- [8] Choi J Y, Sung Y M, Park H J, *et al.* Rapid purification and analysis of alpha-synuclein proteins: C-terminal truncation promotes the conversion of alpha-synuclein into a protease-sensitive form in *Escherichia coli*[J]. *Biotechnol Appl Biochem*, 2002, 36(Pt 1): 33-40.
- [9] Lücking C B, Brice A. Alpha-synuclein and Parkinson's

disease[J]. *Cell Mol Life Sci*, 2000, 57(13-14): 1894-908.

- [10] Johnson W G. Late-onset neurodegenerative diseases-the role of protein insolubility[J]. *J Anat*, 2000, 196 (Pt 4): 609-16.
- [11] Conway K A, Lee S J, Rochet J C, *et al.* Acceleration of oligomerization, not fibrillization, is a shared property of both α -synuclein mutations linked to early-onset Parkinson's disease: Implications for pathogenesis and therapy[J]. *Biochemistry*, 2000, (97): 571-6.
- [12] Kawamata H, McLean P J, Sharma N, *et al.* Interaction of alpha-synuclein and synphilin-1: effect of Parkinson's disease-associated mutations[J]. *J Neurochem* 2001, 77(3): 929-34.
- [13] Tamsamani J, Guinot P. Antisense oligonucleotides: a new therapeutic approach[J]. *Biotechnol Appl Biochem*, 1997, 26(3): 65-71.
- [14] Agrawal S. Antisens oligonucleotides: towards clinical trials[J]. *Trends Biotech*, 1996, 14(10): 376-87.

(编辑 刘清海)