Int J Leg Med (1994) 107:127–131

ORIGINAL ARTICLE

Ranajit Chakraborty · Li Jin · Yixi Zhong

Paternity evaluation in cases lacking a mother and nondetectable alleles

Received: 28 February 1994 / Received in revised form: 16 August 1994

Abstract In parentage testing the formulae for computing paternity index and exclusion probability generally ignores the presence of nondetectable alleles at the loci tested. In contrast, it is now known that even when paternity testing is done with hypervariable DNA markers, nondetectable alleles should not be ignored. This work presents simple formulae needed with this consideration, to analyze paternity evaluation from DNA markers in cases where the mother of the disputed child is unavailable for testing. It is shown that even a modest frequency of nondetectable alleles (e.g., 2–5% per locus) may have a substantial impact on the paternity index when the child and/or the alleged father exhibits a single-banded DNA profile at a locus. Use of such formulae can generate a high probability of exclusion and a high paternity index when multiple independently segregating hypervariable DNA markers are used.

Key words Paternity testing · Deficiency cases · DNA

Zusammenfassung Bei der Vaterschaftsbestimmung ignorieren im allgemeinen die Formeln für die Berechnung des Paternitätsindex und der Ausschlußchance die Anwesenheit nicht-nachweisbarer Allele an den untersuchten Loci. Im Gegenteil, es ist jetzt bekannt, daß selbst, wenn Vaterschaftsuntersuchungen mit hypervariablen DNA-Markern durchgeführt werden, nicht nachweisbare Allele nicht ignoriert werden sollten. Diese Untersuchung präsentiert einfach Formeln, welche unter dieser Bedingung benötigt werden, um die Vaterschaftsbestimmung mit DNA-Markern in solchen Fällen durchzuführen, in denen die Mutter des zu untersuchenden Kindes nicht verfügbar ist. Es wird gezeigt, daß sogar eine mäßige Frequenz nicht detektierbarer Allele (z.B. 2–5% per Locus) einen substantiellen Einfluß auf den Paternitätsindex haben kann, wenn das Kind und/oder der Putativvater ein Einzelbandenmuster an einem Locus haben. Die Benutzung solcher Formeln kann eine hohe Wahrscheinlichkeit des Ausschlusses und einen hohen Paternitätsindex generieren, wenn zahlreiche, unabhängig voneinander segregierende hypervariable DNA-Marker benutzt werden.

Schlüsselwörter Vaterschaftsbestimmung
Defizienfälle · DNA

R. Chakraborty (✉) · Y. Zhong
Center for Demographic and Population Genetics
Graduate School of Biomedical Sciences
University of Texas at Houston Health Science Center
P.O. Box 20334, Houston, Texas 77225, USA

L. Jin
Genetics Department
Stanford University Medical School
Stanford, California, USA

Introduction

Out of approximately 140,000 cases of genetic testing of parentage analysis in the US per year, it is estimated that in nearly 5% of them the mother is unavailable for testing [1]. In such cases, paternity determination involves genetic typing based on the child (C) and the alleged father (AF) alone. Although such cases require some care in the analysis, both in terms of the determination of exclusion probability (P_E) and the evaluation of paternity index (PI), their treatment is rather simple particularly in the context of DNA typing. This is so, because paternity analysis in motherless cases is a special case of more general 'deficiency' cases [2–4], or the component probabilities for such analyses can be numerically obtained from some general pedigree analysis algorithms [5, 6]. Evaluations of exclusion probability and paternity index in motherless cases have been also specifically discussed [7, 8].

While these theories apply to all types of Mendelian genetic markers (serological, biochemical, and DNA), it should be noted that since a motherless case (or any general deficiency case) offers a lower probability of exclusion (and consequently, a lower paternity index) than any complete analysis (with the full trio analyzed), certain considerations of laboratory typing of markers are even more important in the analysis of such deficiency cases.

For example, silent alleles are known to exist for many biochemical markers [9]. In the restriction fragment length polymorphism (RFLP) analysis of hypervariable DNA markers, sometimes alleles which produce aberrantly small or large size DNA fragments cannot be scored [10–12], and in polymerase chain reaction (PCR)-based protocols, alterations at the priming sequences may induce nondetectibility of certain alleles [13, 14]. In the reverse dot blot technique of DNA profiling as well (such as the case with typing at the HLA-DQα locus), the frequency of nondetectable alleles may not be negligible for some populations [15].

Even though a general pedigree analysis algorithm [7, 8] can readily consider the consequences of such phenomena, in the specific evaluations of parentage analysis of deficiency cases [2–4, 7, 8, 16] this was not done. The purpose of this paper is to examine the effect of the presence of nondetectable alleles on the exclusion probability (P_E) and paternity index (PI) computations in cases where the mother is unavailable for testing. With numerical computations, we illustrate that even with a moderate frequency of nondetectable alleles, the exclusion probability as well as the paternity index may be substantially affected for a specific set of observed DNA profiles for the AF-C pair.

Theory

Exclusion probability for an observed DNA profile for a child:

Most DNA typing involves autosomal codominant multiallelic loci, at each of which the DNA profile of a child may exhibit either a single-band pattern, or a 2-band pattern. When nondetectable alleles exist, not all single-band profiles would represent homozygosity of the individuals. Let the $k+1$ segregating alleles at the locus be represented by A_1, A_2, ..., A_k, and A_0, in which the DNA fragment sizes produced by alleles A_1, A_2, ..., A_k are all detectable, but the DNA fragment resulting from the allele A_0 remains undetected. Let p_1, p_2, ..., p_k, and r be the frequencies of these alleles in a population, so that

$$\sum_{i=1}^{k} p_i + r = 1. \tag{1}$$

When the child's DNA profile at such a locus exhibits a two-banded pattern (say, A_iA_j), any alleged father (AF) who has either of the 2 alleles (A_i and/or A_j) in his DNA profile would not be excluded. Therefore, under the assumption of Hardy-Weinberg equilibrium (HWE), the probability of exclusion, when the child's genotype is A_iA_j, is

$$P_E(A_iA_j) = [1 - (p_i + p_j)]^2, \tag{2}$$

which depends only on the frequencies of the 2 alleles present in the child's genotype.

In contrast, when the child's DNA profile at the locus exhibits a single-band pattern (say, A_i–), under the same assumption of HWE, the probability of exclusion becomes

$$P_E(A_i-) = [1 - (p_i + r)]^2, \tag{3}$$

which requires the knowledge of the frequency (r) of nondetectable alleles (A_0) in addition to the frequency (p_i) of the allele present in the child's DNA profile.

Two points are worth noting at this point. First, in the presence of nondetectable alleles, the inference of paternal exclusion when the child's (C) DNA pattern is single banded is somewhat different from the traditional codominant systems. An alleged father showing any single-banded DNA profile cannot be excluded even when the single-banded profile of a child (C) does not exhibit a match of the detectable fragment size found in the DNA profile of the alleged father (AF). This is so because AF and C might share the same nondetectable allele. Second, equation (2) apparently indicates that the exclusion probability for a two-banded DNA pattern of a child may not be affected in the presence of nondetectable alleles (since r does not explicitly appear in this formula), but it is not really so. When in DNA databases nondetectable alleles are neglected, the gene count estimates of allele frequencies are overestimates of their true values (because, in such situations, the estimated allele frequencies are truly the estimates of $p_i/(1-r)$, instead of simple p_i, see [17]), so that the use of estimated allele frequencies p_i and p_j in equation (2) without any adjustment for the presence of nondetectable alleles will give only a lower bound for the exclusion probability, P_E.

Table 1 shows some numerical illustrations of the effect of the presence of nondetectable alleles on exclusion probabilities in motherless cases where the child's profile is single-banded. It is clear that the presence of a nondetectable allele ($r > 0$) lowers the exclusion probability, and the effect becomes more severe as the frequency of the detectable obligatory allele (A_i) increases. Since in general at the hypervariable loci used in RFLP databases the frequency of nondetectable alleles is less than 10% [11, 12, 17, 18], the computations in Table 1 indicate that the exclusion probability could be reduced by 20–25% per locus, when the child's DNA profile is single-banded. For most loci, r rarely exceeds 5% [18], and hence the effect may be somewhat less severe (e.g., 13% for $p_i = 0.25$, and $r = 0.05$).

Table 1 Paternity Exclusion Probabilities (P_E) in motherless cases for a single-banded profile of a child and nondetectable alleles (r)

p_i	P_E for different values of r				
	0.0	0.01	0.02	0.05	0.10
0.01	0.9801	0.9604	0.9409	0.8836	0.7921
0.05	0.9025	0.8836	0.8649	0.8100	0.7225
0.10	0.8100	0.7921	0.7744	0.7225	0.6400
0.25	0.5625	0.5476	0.5329	0.4900	0.4425

Table 2 Paternity Index for all alleged father (AF)-child (C) pairs of DNA profiles when the alleged father is not excluded

Case	DNA profile	X	Y	*PI*
(a) Both two-banded, one shared band	C: A_iA_j AF: A_iA_l	$p_ip_jp_l$	$4p_i^2\,p_jp_l$	$\frac{1}{4p_i}$
(b) Both two-banded, two shared band	C: A_iA_j AF: A_iA_j	$p_ip_j\,(p_i+p_j)$	$(2p_ip_j)^2$	$\frac{1}{4p_i}+\frac{1}{4p_j}$
(c) AF two-banded, C single-banded, one shared band	C: A_i- AF: A_iA_j	$p_ip_j\,(p_i+r)$	$2p_i^2p_j(p_i+2r)$	$\frac{p_i+r}{2p_i(p_i+2r)}$
(d) AF single-banded, C two-banded, one shared band	C: A_iA_j AF: A_i-	$p_ip_j\,(p_i+r)$	$2p_i^2p_j(p_i+2r)$	$\frac{p_i+r}{2p_i(p_i+2r)}$
(e) Both single-banded, one shared band	C: A_i- AF: A_i-	$p_i[p_ir+(p_i+r)^2]$	$p_i^2(p_i+2r)^2$	$\frac{(p_i+r)^2+p_ir}{p_i(p_i+2r)^2}$
(f) Both single-banded, no shared band	C: A_i- AF: A_j-	p_ip_jr	$4p_ip_jr^2$	$\frac{1}{4r}$

Table 3 Effect of nondetectable alleles (r) on Paternity Index (*PI*)

p_i	*PI* for different values of r				
	0.0	0.01	0.02	0.05	0.10
Case: C = A_i-; AF = A_iA_j					
0.01	50.0	33.3	30.0	27.3	26.2
0.05	10.0	8.6	7.8	6.7	6.0
0.10	5.0	4.6	4.3	3.8	3.3
0.25	2.0	1.9	1.8	1.7	1.6
Case: C = A_i-; AF = A_i-					
0.01	100.0	55.6	44.0	33.9	29.7
0.05	20.0	16.7	14.6	11.1	8.8
0.10	10.0	9.1	8.4	6.9	5.6
0.25	4.0	3.8	3.7	3.3	2.9

Paternity Index for an observed pair of DNA profiles for AF and C:

When the alleged father (AF) cannot be excluded, there are 6 possible scenarios of DNA profiles that can be observed for the AF-C pair, 5 of which are discussed by Brenner [8]. Table 2 shows these possibilities, for each of which the paternity index (*PI*) is defined as the ratio of 2 likelihood functions, X and Y, which represent the probabilities of the observed pair of DNA profiles for the C-AF pair, under the 2 alternative hypotheses; namely, AF is the true father of C (giving the likelihood X), and AF is an unrelated random man, and not the biological father of C (giving the likelihood Y).

As expected, when either or both of AF and C exhibit a single-banded DNA profile, the paternity index (*PI*) is affected by the presence of nondetectable alleles. For a single-banded C, the paternity index of a heterozygote AF is less severely affected in the presence of nondetectable alleles in comparison to a single-banded profile of AF. This is numerically illustrated in Table 3, which also shows that even a modest frequency (say, 2%) of a nondetectable allele can reduce the PI substantially. The reduction is more severe when the detectable obligatory gene in C is rare in the population. The monotonic decay of *PI* with increasing value of r suggests that in order to obtain a lower-bound estimate of PI one must use the upper confidence limit of the point estimates of p_i and r, which may be obtained from the method discussed in [18].

An actual numerical example

Brenner [8] discussed 3 actual examples of paternity testing in motherless cases with hypervariable DNA probes 3′HVR, YNH24, and TBQ7. In one case (Case B) single-banded profiles were observed for either AF (at the TBQ7 locus), or for C (at 3′HVR and YNH24). In Table 4 we used the fragment sizes, and allele frequencies reported by Brenner [8] to illustrate the effect of the presence of nondetectable alleles with some hypothetical (but realistic) values of r. For this specific case, note that in spite of the fact that the AF's profile at the TBQ7 locus was single-banded, the paternity exclusion probability is unaffected in the presence of nondetectable alleles. The reduction in the combined exclusion probability for the observed 3-locus profiles of the child is not severely affected in the presence of nondetectable alleles at one or more of these loci, since even if r was of the level of 5% per locus, the combined probability of exclusion becomes 96.6%, instead of the 97.9% that would result if the nondetectable alleles at these loci were excluded. The effect of nondetectable alleles, however, is more substantial on the paternity index. For the non-excluded 3-locus profile of the alleged father tested in this case, Brenner [8] reports a paternity index of 50 (it actually should be 51.4) where nondetectable alleles were not considered. If at each locus r

Table 4 An illustration in an actual paternity case lacking the mother

Probe	3′HVR		YNH24		TBQ7		Combined
Fragment sizes (kb):							
for AF	3.68	2.34	3.98	3.70	3.51		
for C		2.34		3.70	3.51	5.65	
Frequencies for child's fragments		0.218		0.136	0.082	0.032	
Exclusion probability:							
$r = 0.0$	0.6115		0.7465		0.7850		0.9788
$r = 0.01$	0.5960		0.7293		0.7850		0.9765
$r = 0.02$	0.5806		0.7123		0.7850		0.9741
$r = 0.05$	0.5358		0.6626		0.7850		0.9663
Paternity index:							
$r = 0.0$	2.29		3.68		6.10		51.4
$r = 0.01$	2.20		3.44		5.50		41.6
$r = 0.02$	2.12		3.26		5.10		35.2
$r = 0.05$	1.93		2.90		4.42		24.8

Source of data: Brenner [8]

was 0.05, this value would be 24.8, less than 50% of the value reported [8]. This suggests that when single-banded profiles are found for the child and/or alleged father at one or more DNA loci, perhaps additional loci should be tested when a more stringent probability level for exclusion as well as paternity (say, 99%) is demanded.

Discussion and conclusions

The theory discussed above shows that appropriate consideration of nondetectable alleles must be taken into account in the evaluation of paternity determinations in cases when the mother of the disputed child is unavailable for genetic testing. The use of codominant DNA markers does not obviate this difficulty, since irrespective of the techniques of DNA typing, nondetectable alleles may not be ignored for any locus in all populations.

Numerical illustrations shown in Tables 1 and 3 suggest that particularly when the detectable alleles present in the child's DNA profile are rare in the population, even a small frequency of nondetectable alleles (say, 1–2%) may substantially reduce the exclusionary power of a locus, and reduce the paternity index for any non-excluded alleged father.

Finally, it should be noted that the general expression [7] for the average exclusion probability in motherless cases can be extended to include the nondetectable alleles, in which the average exclusion probability for a multiallelic locus becomes

$$\bar{P}_E = 1 - 4a_2 + 4a_3 - 3a_4 + 2a_2^2 - r^2(2-r)^2 - r(2a_2 - 2a_3 - 5ra_3), \quad (4)$$

where a_n is the sum of the n-th power of all detectable allele frequencies ($n = 2$, 3, and 4), $a_n = \sum p_i^n$, in which the summation is over all k detectable alleles, and r is the frequency of nondetectable alleles.

Acknowledgements This work was supported by the US Public Health Service Research Grants 92-IJ-CX-K024 from the National Institute of Justice, and GM 41399 from the National Institutes of Health. We thank 2 anonymous reviewers for their comments and suggestions on an earlier version of the paper. We thank the editor for drawing our attention to some earlier work on this subject. The opinions expressed in this work are those of the authors, and these do not constitute an endorsement of the granting agency.

References

1. American Association of Blood Banks (1991) Annual report. Arlington, Virginia
2. Ihm P, Hummel K (1975) Ein Verfahren zur Ermittlung der Vaterschaftswahrscheinlichkeit aus Blutgruppenbefunden unter beliebiger Einbeziehung von Verwandten. Z Immun Forsch 149:405–416
3. Asano M, Minakata K, Hattori H (1980) General formulas of the estimated likelihood ratio Y/X in the diagnosis of paternity of a deceased putative father. Z Rechtsmed 84:125–133
4. Henke J, Baur MP, Hoffman K, Henke L (1991) Application of highly informative single-locus DNA minisatellite probes to complicated deficient paternity cases. Arztl Lab 37:175–179
5. Elston RC, Stewart J (1971) A general model for the genetic analysis of pedigree data. Hum Hered 21:523–542
6. Lange K, Boehnke M (1983) Extensions to pedigree analysis. V. Optimal calculation of Mendelian likelihoods. Hum Hered 33:291–303
7. Garber RA, Morris JW (1983) General equations for the average power of exclusion for genetic systems of n codominant alleles in one-parent and no-parent cases of disputed parentage. In: Walker RH (ed) Inclusion probabilities in parentage testing. Am Assoc Blood Banks, Arlington, pp 277–280
8. Brenner CH (1993) A note on paternity computation in cases lacking a mother. Tranfusion 33:51–54
9. Martin W (1983) Consideration of silent genes in the statistical evaluation of blood group findings in paternity testing. In: Walker RH (ed) Inclusion probabilities in parentage testing. Am Assoc Blood Banks, Arlington, pp 245–265
10. Budowle B, Giusti AM, Waye JS, Baechtel FS, Fourney RM, Adams DE, Presley LA, Deadman HA, Monson KL (1991) Fixed-bin analysis for statistical evaluation of continuous distributions of allelic data from VNTR loci, for use in forensic comparisons. Am J Hum Genet 48:841–855

11. Devlin B, Risch N (1992) A note on Hardy-Weinberg equilibrium of VNTR data by using the Federal Bureau of Investigation's fixed-bin method. Am J Hum Genet 51:549–553
12. Steinberger EM, Thompson LD, Hartmann JM (1993) On the use of excess homozygosity in subpopulation detection. Am J Hum Genet 52:1275–1277
13. Callen DF, Thompson AD, Shen Y, Phillips HA, Richards RI, Mulley JC, Sutherland GR (1993) Incidence and origin of "null" alleles in the (AC)n microsatellite markers. Am J Hum Genet 52 : 922–927
14. Koorey DJ, Bishop GA, McCaughan GW (1993) Allele non-amplification: a source of confusion in linkage studies employing microsatellite polymorphisms. Hum Mol Genet 2:289–291
15. Imanishi T, Akaza T, Kimura A, Tokunaga K, Gojobori T (1992) Allele and haplotype frequencies for HLA and Complement loci in various ethnic groups. In: Tsuji K, Aizawa M, Sasazuki T (eds) HLA 1991, volume 1. Oxford University Press, Oxford, pp 1065–1074
16. Baur MP, Rittner C (1981) Likelihood ratios in paternity cases: calculation and evaluation. Arztl Lab 27:261–270
17. Chakraborty R, de Andrade M, Daiger SP, Budowle B (1992) Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic applications. Ann Hum Genet 56:45–57
18. Chakraborty R, Zhong Y, Jin L, Budowle B (1993) Nondetectability of restriction fragments and independence of DNA fragment sizes within loci in the RFLP typing of DNA. Am J Hum Genet 55:391–401