

## Deduction of paternity index from DNA mixture

Xiang Hai Liao<sup>a</sup>, Tai Shing Lau<sup>b,\*</sup>, Karenda Fai Ngor Ngan<sup>c</sup>, Jun Wang<sup>a</sup>

<sup>a</sup>Department of Biochemistry, The Chinese University of Hong Kong, N.T., Hong Kong SAR, PR China

<sup>b</sup>Department of Statistics, The Chinese University of Hong Kong, N.T., Hong Kong SAR, PR China

<sup>c</sup>Molecular Technology Innomed Limited, Harcourt House, Room 1206, 39 Gloucester Road, Wanchai, Hong Kong SAR, PR China

Received 19 September 2001; accepted 3 May 2002

### Abstract

Determination of individual genotypes in DNA mixture remains a challenge in forensic science. Using an approach of mixture of distributions, this article provides formula for calculation of paternity index (PI) in cases where only tissue mixture of the mother and alleged father, the genotypes of the mother and child, but not that of the alleged father are available. The formula has been used to solve a real case using mother's vaginal tissue contaminated with semen from alleged father.

© 2002 Elsevier Science Ireland Ltd. All rights reserved.

**Keywords:** Paternity index; DNA mixture; Mixture distribution

### 1. Introduction

In forensic cases, DNA is often extracted from a stain that contains body fluids or tissue from more than one individual. It is usually difficult to separate individual's DNA from each other, mechanically. Given the genotypes of the mixture and the suspect's genotype [1], used likelihood ratio to evaluate different hypothesis that the mixture came from the suspect and another man, and the hypothesis that the mixture came from two other men. Weir et al. [2] assumed that the mixture came from some known number of contributors and known profiles of contributors, and provided the formula for computing the probabilities in both, the numerator and denominator of the usual likelihood ratio. Clayton et al. [3] discussed the difficulties with the analysis of the mixed forensic stains and suggested the interpretation of the mixture profile should be independent of the reference samples. Further, they suggested that similar to [1], all pairwise combinations of allelic pairs are listed and their possibilities are evaluated based on the peak areas in the mixture profile. The problems raised and solutions suggested by [1–3], are not specified to the paternity case.

In this article, we have developed methods to compute the paternity index (PI) for cases where only the mixture of the mother and alleged father, and the genotypes of the mother and child, but not that of the alleged (known) father are available, Table 1.

### 2. Results

Assuming the mixture comes from two persons, alleged father and mother, and there are at the most four alleles observed, we can determine all possible genotypes of the alleged father. We combine all the possible cases of the alleged father corresponding to the same genotypes for mother ( $G_M$ ) and child ( $G_C$ ) and the same observed alleles of the mixture ( $E$ ). Let  $H_p$  denote the hypothesis that the alleged person is the father. Let  $H_d$  be the hypothesis that the alleged father is not the biological father. The PI (likelihood ratio) is thus computed as [4]:

$$\begin{aligned} \text{PI} &= \frac{\Pr(E|H_p)}{\Pr(E|H_d)} = \frac{\sum_A \Pr(G_C|G_M, G_A, H_p)\Pr(G_M, G_A|H_p)}{\sum_A \Pr(G_C|G_M, G_A, H_d)\Pr(G_M, G_A|H_d)} \\ &= \frac{\sum_A \Pr(G_C|G_M, G_A, H_p)\Pr(G_M, G_A)}{\sum_A \Pr(G_C|G_M, G_A, H_d)\Pr(G_M, G_A)} \end{aligned}$$

Notice that the sum is taken over all possible genotypes of the alleged father, and the genotypes of mother and alleged

\* Corresponding author. Tel.: +852-26097927;

fax: +852-26035188.

E-mail address: tslau@sparc2.sta.cuhk.edu.hk (T.S. Lau).

Table 1

Provides the paternity index for cases when the genotype of the alleged father is not known

Phenotype (mixture)	Mother's genotype	Child's genotype	Alleged father's genotype	Paternity index (PI)
P	PP	PP	PP	1/p
PQ	PQ	PQ	PP, PQ, QQ	1/(p+q)
		PP	PP, PQ	(p+q)/p(p+2q)
	PP	PQ, QQ	PQ, QQ	(p+q)/q(2p+q)
PQR	PP	PP	PQ	0.5/p
		PQ	QR	0.5/q
	PQ	PP	PR	0.5/p
		QQ	QR	0.5/q
		PQ	PR, QR	0.5/(p+q)
PQRS	PQ	PR	PR, QR, RR	(p + q + r)/r(2p + 2q + r)
		PR	RS	0.5/r

father are independent of  $H_p$  and  $H_d$ . The probability  $\Pr(G_M, G_A)$  can be computed using Hardy-Weinberg equilibrium and combinatorics.

Example 1. If the phenotype of the mixture is P, then the genotype for both parent and the child must be PP. In this case, the PI is given by:

$$\begin{aligned} PI &= \frac{\sum_A \Pr(G_C|G_M, G_A, H_p)\Pr(G_M, G_A)}{\sum_A \Pr(G_C|G_M, G_A, H_d)\Pr(G_M, G_A)} \\ &= \frac{\sum_A \Pr(G_M, G_A)}{\sum_A p\Pr(G_M, G_A)} = \frac{1}{p} \end{aligned}$$

Example 2. If the phenotype of the mixture is PQ, the genotypes of mother and child is PQ and PP, respectively, the PI of the alleged father of genotype PQ and PP are 0.5/p and 1/p, respectively. The PI is given by:

$$\begin{aligned} PI &= \frac{\sum_A \Pr(G_C|G_M, G_A, H_p)\Pr(G_M, G_A)}{\sum_A \Pr(G_C|G_M, G_A, H_d)\Pr(G_M, G_A)} \\ &= \frac{0.5\Pr(PQ, PQ) + \Pr(PQ, PP)}{p\Pr(PQ, PQ) + p\Pr(PQ, PP)} \\ &= \frac{0.5(2pq \times 2pq) + (2pq \times pp)}{p(2pq \times 2pq) + p(2pq \times pp)} \\ &= \frac{p + q}{p(p + 2q)} \end{aligned}$$

Example 3. If the phenotype of the mixture is PQR and the genotype of mother is PQ, while the genotypes of the child are PR and QR, the possible genotypes of the alleged father are PR, QR and RR. Their corresponding PIs are given by 0.5/r, 0.5/r and 1/r, respectively.

$$\begin{aligned} PI &= \frac{\sum_A \Pr(G_C|G_M, G_A, H_p)\Pr(G_M, G_A)}{\sum_A \Pr(G_C|G_M, G_A, H_d)\Pr(G_M, G_A)} \\ &= \frac{0.5\Pr(PQ, PR) + 0.5\Pr(PQ, QR) + \Pr(PQ, RR)}{r[\Pr(PQ, PR) + \Pr(PQ, QR) + \Pr(PQ, RR)]} \\ &= \frac{0.5(2pq \times 2pr) + 0.5(2pq \times 2qr) + (2pq \times rr)}{r[(2pq \times 2pr) + (2pq \times 2qr) + (2pq \times rr)]} \\ &= \frac{p + q + r}{r(2p + 2q + r)} \end{aligned}$$

### 3. Case

The following is a real case, we have recently solved. The sample is mother's vaginal tissue contaminated with semen from alleged father. For a number of reasons, the alleged father could not be genotyped while the genotypes of the nine loci were determined for the mother and the child using their buccal cells. The PI was computed to relate the alleged father to the child.

Locus	Mixture genotype	Mother's genotype	Child's genotype	Paternity index (PI)
TH01	7,9	9	9	1.14
TPOX	8,11	8,11	8	1.34
CSF1PO	10,11,12	10,11	11,12	1.75
D5S818	9,12,13	12,13	12,13	2.67
D13S317	8,9,10,12	10,12	8,10	1.63
D7S820	10,11,12	10,11	11,12	2.56
D3S1358	15,16,18	16,18	16	1.53
vWA	14,16,17,18	14,17	16,17	3.21
FGA	21,22.2,23	22.2,23	21,22.2	4.78

The combination of PI is therefore 699.22, providing the prior probability is 0.5, the posterior probability of paternity is 99.86%.

### 4. Discussion

In this communication, we have described mathematical methods to calculate PI from the DNA mixture of the mother and alleged father. The formula is also relevant to other cases of similar circumstances; one of its potential applications is to determine, with the available mixture stain, whether the rapist is the father of the child of the victim.

The factor of mutation of STR markers should be considered and can be readily incorporated into PI calculation in for an isolated single-locus mismatch between the DNA mixture and the child [5,6].

Population substructure and the possible involvement of relatives are often the other major concern in human identification using the STR systems. We do not think these two related issues will alter significantly the data from our system, as the use of the large numbers of independent genetic loci likely stabilises the possible fluctuations in the allele frequencies among sub-populations [7] and the effect of relatives, even in the case of siblings, would not change the conclusion, regardless the magnitudes of PI [8]. Given the coancestry coefficient [5], derived the likelihood ratios for relatedness testing. Our results can be easily extended to cover the case.

## References

- [1] I.W. Evett, C. Buffery, G. Willott, A guide to interpreting single locus profiles of DNA mixtures in forensic cases, *J. Forensic Sci. Soc.* 31 (1991) 41–47.
- [2] B.S. Weir, C.M. Triggs, L. Starling, L.I. Stowell, K.A.J. Walsh, J. Buckleton, Interpreting DNA mixtures, *J. Forensic Sci.* 42 (1997) 213–222.
- [3] T.M. Clayton, J.P. Whitaker, R. Sparkes, P. Gill, Analysis and interpretation of mixed forensic stains using DNA STR profiling, *Forensic Sci. Int.* 91 (1998) 55–70.
- [4] I.W. Evett, B.S. Weir, *Interpreting DNA Evidence Statistical Genetics for Forensic Scientists*, Sinauer Associates, Sunderland, MA, USA, 1998.
- [5] K.L. Ayres, Relatedness testing in subdivided populations, *Forensic Sci. Int.* 114 (2000) 107–115.
- [6] T. Egeland, P.F. Mostad, B. Olaisen, A computerized method for calculating the probability of pedigrees from genetic data, *Sci. Justice* 37 (1997) 269–274.
- [7] B.S. Weir, W.G. Hill, Population genetics of DNA profiles, *J. Forensic Sci. Soc.* 33 (1993) 218–225.
- [8] J.F.Y. Brookfield, The effect of relatives on the likelihood ratio associated with DNA profile evidence in criminal cases, *J. Forensic Sci. Soc.* 34 (1994) 193–197.