Int J Legal Med (2002) 116:79–86 

ORIGINAL ARTICLE

Wing K. Fung · Yue-Qing Hu

# The statistical evaluation of DNA mixtures with contributors from different ethnic groups

Received: 19 February 2001 / Accepted: 19 July 2001

**Abstract** The effect of a structured population on the evaluation of forensic mixed stains has been considered by the authors and others. However, in countries with multiple racial or ethnic groups, it is not uncommon that contributors to a DNA mixture are of different ethnic groups. A famous example is the OJ Simpson case in which the suspect was an African-American, the victims were Caucasian Americans and the true perpetrator(s) could be from any ethnic group(s). In this paper six common mixture cases are considered and the formulae for likelihood ratios are derived. These formulae can help forensic DNA scientists acquire a better understanding of the problem. The effect of different ethnic groups is illustrated using a case in Hong Kong.

**Keywords** DNA mixtures · Ethnic groups · Forensic science · Likelihood ratio · Population structure

## Introduction

DNA profiling has proved to be a powerful tool for forensic human identification. Consider the following simple situation: a crime has been committed, the perpetrator has left a blood stain at the scene of the crime and a suspect has been identified. Suppose that the stain in the crime scene is typed with alleles $A_iA_j$ at a particular locus and the suspect has the same alleles. It is common to calculate the probability of a random match of DNA alleles at the crime scene and those of the suspect under the hypothesis of innocence. Under the assumption of Hardy-Weinberg (HW) equilibrium, it can be evaluated as:

$$\begin{array}{l} 2p_ip_j \text{ for heterozygous alleles } A_iA_j, i \neq j \\ \quad p_i^2 \text{ for homozygous alleles } A_iA_i, \end{array} \tag{1.1}$$

where $p_i$ and $p_j$ are the proportions or frequencies for alleles $A_i$ and $A_j$, respectively.

In forensic DNA analyses it is not uncommon to find DNA samples containing materials from more than one person. For example, the sample in a rape case may contain materials from the victim, her consensual sexual partner(s) and/or the perpetrator(s). The mixed stain problem is complex as commented by the second National Research Council (NRC-II 1996) report on the evaluation of forensic evidence and some forensic scientists find it difficult to assess mixed stains. The mixed stain problem was discussed by Evett et al. (1991), Weir et al. (1997) and Fukshansky and Bär (1998) and the HW equilibrium was assumed. Researchers have investigated the validity of the assumption in various ethnic or racial groups (see for example Devlin and Risch 1993 and Fung 1996). The HW law, however, is seldom exactly correct (NRC-II 1996) because dependence can have developed among alleles in the same ethnic group during the human evolutionary process. It is sensible to take this dependence structure into account in the evaluation of random match probability. A relatively simple formula for assessing the dependence was provided by Balding and Nichols (1994), which was used in recommendation 4.2 of the NRC-II (1996) report for handling the single sample problem. Recently, Curran et al. (1999) and Fung and Hu (2000a) tackled the mixed stain problem with such a dependence on the population structure.

Based on genetic and statistical justifications, Balding and Nichols (1994) showed that in general, if $y$ copies of allele $A_i$ have been observed among $n$ alleles, then the probability that the next allele to be observed will be $A_i$ is:

$$P(A_i|yA_i \text{ among } n \text{ alleles}) = \frac{y\theta + (1-\theta)p_i}{1+(n-1)\theta} \tag{1.2}$$

W. K. Fung (✉)
Department of Statistics and Actuarial Science,
The University of Hong Kong, Pokfulam Road,
Hong Kong, China
e-mail: wingfung@hku.hk

Y.-Q. Hu
Department of Applied Mathematics, Southeast University,
Nanjing 210018, China

where $\theta$ is the coancestry coefficient, which is essentially the same as Wright's (1951) $F_{st}$. The latter, which is well known to population geneticists, has a genetic interpretation that measures the interpopulation variation in allele frequencies. Equation (1.2) can also be established if a state of evolutionary equilibrium has been established (Curran et al. 1999; Wright 1951). The model of Balding and Nichols (1994) is quite general and their method is endorsed by the NRC-II. It forms the basis of our derivation in this paper for taking dependence into account for mixed stain problems.

Previous reports on mixtures all regarded or assumed that contributors to the mixed stain came from one single ethnic group. In practice, however, it is not uncommon that contributors to a mixed stain belong to different ethnic or racial groups in countries with multiple ethnic groups such as the USA, UK or Singapore. A famous example is the OJ Simpson case in which the defendant was an African-American and the two victims were Caucasians and mixed stains were found in this case. The perpetrator(s) could be African-American(s), Caucasian(s), or from any other group. Extensive studies from various databases indicate that there are substantial differences in frequencies among the major ethnic groups (NRC-II 1996) and ignorance of the ethnic group information of contributors to the mixed stain may not be allowed in courtrooms and could also be misleading in assessing the weight of evidence of the mixed stains. Recently, Fukshansky and Bär (1999) and Fung and Hu (2001) studied such problems, but the HW law was assumed. Triggs et al. (2000) offered a coherent method to estimate likelihood ratios for DNA match probabilities from mixed racial populations. Buckleton et al. (1998) dealt with the situation when multiple hypotheses are postulated for mixtures.

The purpose of this paper is to evaluate the mixed stain problem with contributors of different structured ethnic groups. The likelihood ratios (*LR*s) for six common cases as given in Curran et al. (1999) are presented. These formulae can help forensic DNA scientists acquire a better understanding of the problem. A general formula for *LR* is being sought and it will be reported elsewhere.

## Likelihood ratio

Suppose a crime was committed and a mixed stain was collected from the crime scene. Some crime-related persons, for example the victim and the suspect(s), were typed for their DNA characteristics. Usually, a pair of propositions, the prosecution proposition ($H_p$) and the defence proposition ($H_d$) are given to explain who the contributors to the mixed stain were. We can have propositions like, for example:

-$H_p$ The contributors were the victim and the suspect
-$H_d$ The contributors were the victim and one unknown person.

The likelihood ratio is often used to assign the weight of the genetic evidence in this circumstance (Weir et al. 1997; Fung and Hu 2000a, b):

$$LR = \frac{P(\text{Evidence}|H_p)}{P(\text{Evidence}|H_d)}, \tag{2.1}$$

where the evidence is all the information carried by the typed persons and that found in the mixed stain. If the likelihood ratio is $L$, it means that the evidence is $L$ times more likely to have arisen under explanation $H_p$ than under explanation $H_d$. Under either proposition, $H_p$ or $H_d$, every typed person would be declared to have contributed to the mixed stain or not.

For the simplicity of exposition, let $M$ denote the genetic profile of the mixed stain, which is simply a listing of the distinct alleles found in the mixed stain. We also let $K$ denote the collection of genotypes of the known typed persons. Thus the evidence can be represented as $(M,K)$, and the likelihood ratio can be expressed as

$$\begin{aligned} LR &= \frac{P(M,K|H_p)}{P(M,K|H_d)} \\ &= \frac{P(K|H_p)}{P(K|H_d)}\frac{P(M|K,H_p)}{P(M|K,H_d)} \end{aligned}$$

using the third law of probability. The terms $P(K|H_p)$ and $P(K|H_d)$ denote the probabilities of having the alleles for the typed persons. They are equal because whether $H_p$ or $H_d$ is true or not does not provide any information about the uncertainty of $K$. Hence the likelihood ratio can be simplified to:

$$LR = \frac{P(M|K,H_p)}{P(M|K,H_d)}. \tag{2.2}$$

So the evaluation of *LR* is reduced to the evaluation of the probability $P(M|K,H)$ for some proposition $H$.

## Likelihood ratios for common cases

Let $A_1,A_2,\ldots,A_j,\ldots$ denote the alleles and let $p_{ij}$ be the allele proportion or frequency for allele $A_j$ in ethnic group $i$, $i \in G = \{a,b,\ldots\}$, $j = 1,2,\ldots$, with a coancestry coefficient $\theta_i$. We assume independence of alleles among ethnic groups. Also let $K_a,K_b,\ldots$ be the genotypes of the known contributors from groups $a,b,\ldots$, respectively and let $X_a,X_b,\ldots$ be the genotypes of the unknown contributors from groups $a,b,\ldots$, respectively. In the following, $V$ is the abbreviation for the victim's genotype and $S$ for the suspect's genotype. We are going to derive the likelihood ratios for six common mixture cases as listed in Curran et al. (1999).

### Four-allele mixture, heterozygous suspect and heterozygous victim

Suppose a mixed stain $M$ was recovered and the victim and one suspect were typed. Consider the case that $M =$

$\{A_1,A_2,A_3,A_4\}$ and the genotypes of the suspect and the victim are $S = A_1A_2$ and $V = A_3A_4$, respectively. Therefore $K = (A_1A_2,A_3A_4)$. The two alternative propositions are:

- $H_p$ The contributors of $M$ were the victim and the suspect
- $H_d$ The contributors of $M$ were the victim and one unknown person.

It can be seen that there could be three persons involved in this case, the victim, the suspect and an unknown person. The maximum number of ethnic groups to which they may belong is 3. In theory, the total number of all possible combinations about which person belongs to which ethnic group is $3^3 = 27$. Without loss of generality, we assume that the unknown contributor belongs to ethnic group $a$, which is termed the unknown-based viewpoint. From this viewpoint, we discuss the following situations:

- $C_1$ $V$, $S$ and $X$ come from group $a$
- $C_2$ $V$, $X$ come from group $a$ and $S$ comes from group $b$
- $C_3$ $S$, $X$ come from group $a$ and $V$ comes from group $b$
- $C_4$ $X$ comes from group $a$ but $V$ and $S$ are not from group $a$,

where $X$ is the genotype of the unknown contributor.

Under proposition $H_p$, it is obvious that $P(M|K,H_p) = 1$. Under $H_d$, from the known contributor $V = A_3A_4$, we know that the genotype of the unknown contributor for explaining the mixture $M = \{A_1,A_2,A_3,A_4\}$ must be $A_1A_2$, i.e., $X_a = A_1A_2$. So $P(M|K,H_d) = P(X_a = A_1A_2|K_a,H_d)$, or simply $P(X_a = A_1A_2|K_a)$. Thus, the $LR$ is just the reciprocal of $P(X_a = A_1A_2|K_a)$ for various $K_a$ in $C_1 - C_4$.

If $C_1$ holds, $K_a = (A_1A_2,A_3A_4)$, so:

$$P(X_a = A_1A_2|K_a) = P(X_a = A_1A_2|K_a = (A_1A_2, A_3A_4)).$$

To evaluate the probability using equation (1.2), we must pass from the situation of genotypes to that of alleles. In this case, we have:

$$\begin{aligned} &P(X_a = A_1A_2|K_a = (A_1A_2, A_3A_4)) \\ &= 2P(A_1, A_2|A_1, A_2, A_3, A_4) \\ &= 2\left[\frac{\theta_a + (1-\theta_a)p_{a1}}{1+3\theta_a}\right]\left[\frac{\theta_a + (1-\theta_a)p_{a2}}{1+4\theta_a}\right] \\ &= 2\frac{[\theta_a + (1-\theta_a)p_{a1}][\theta_a + (1-\theta_a)p_{a2}]}{(1+3\theta_a)(1+4\theta_a)}, \end{aligned}$$

where the factor 2 appears because of the two possible ways of inheritance.

When $C_2$ holds, $K_a = A_3A_4$, so:

$$\begin{aligned} P(X_a = A_1A_2|K_a) &= P(X_a = A_1A_2|K_a = A_3A_4) \\ &= 2\frac{(1-\theta_a)p_{a1}(1-\theta_a)p_{a2}}{(1+\theta_a)(1+2\theta_a)} \end{aligned}$$

If $C_3$ holds, $K_a = A_1A_2$, so:

$$\begin{aligned} P(X_a = A_1A_2|K_a) &= P(X_a = A_1A_2|K_a = A_1A_2) \\ &= 2\frac{[\theta_a + (1-\theta_a)p_{a1}][\theta_a + (1-\theta_a)p_{a2}]}{(1+\theta_a)(1+2\theta_a)}. \end{aligned}$$

When situation $C_4$ holds, we have $K_a = \phi$, then:

$$\begin{aligned} P(X_a = A_1A_2|K_a) &= P(X_a = A_1A_2) \\ &= 2\left[\frac{(1-\theta_a)p_{a1}}{1+(0-1)\theta_a}\right]\left[\frac{(1-\theta_a)p_{a2}}{1+(1-1)\theta_a}\right] \\ &= 2(1-\theta_a)p_{a1}p_{a2}. \end{aligned}$$

The $LR$s of these cases are presented in Table 1. It should be noted that there are just four distinct formulae by which the ethnicity of the unknown person is fixed, but each formula represents an equivalence class of possible formulae.

## Three-allele mixture, homozygous victim and heterozygous suspect

Suppose that $M = \{A_1,A_2,A_3\}$, $S = A_1A_2$ and $V = A_3A_3$. The two alternative propositions are:

- $H_p$ The contributors were the victim and the suspect
- $H_d$ The contributors were the victim and one unknown person.

It is clear that, under proposition $H_p$, $P(M|K,H_p) = 1$. Under $H_d$, since the mixture $M = \{A_1,A_2,A_3\}$ and the known contributor $V = A_3A_3$, $X_a = A_1A_2$. Similar to the derivations for a 4-allele mixture with a heterozygous suspect and victim, we have $P(M|K,H_d) = P(X_a = A_1A_2|K_a)$ and can obtain the likelihood ratios listed in Table 2. It is noted that the $LR$s in Tables 1 and 2 are the same.

**Table 1** Likelihood ratio for $M = \{A_1,A_2,A_3,A_4\}$, $S = A_1A_2$ and $V = A_3A_4$ with $H_p$ the contributors were the victim and the suspect, and $H_d$, the contributors were the victim and one unknown person

| Ethnicity | | | Likelihood ratio |
|---|---|---|---|
| $X$ | $V$ | $S$ | |
| $a$ | $a$ | $a$ | $(1+3\theta_a)(1+4\theta_a)/\{2[\theta_a+(1-\theta_a)p_{a1}][\theta_a+(1-\theta_a)p_{a2}]\}$ |
| $a$ | $a$ | $b$ | $(1+\theta_a)(1+2\theta_a)/[2(1-\theta_a)^2p_{a1}p_{a2}]$ |
| $a$ | $b$ | $a$ | $(1+\theta_a)(1+2\theta_a)/\{2[\theta_a+(1-\theta_a)p_{a1}][\theta_a+(1-\theta_a)p_{a2}]\}$ |
| $a$ | $\bar{a}$ | $\bar{a}$ | $1/[2(1-\theta_a)p_{a1}p_{a2}]$ |

$\bar{a}$ means not group a

**Table 2** Likelihood ratio for $M = \{A_1,A_2,A_3\}$, $S = A_1A_2$ and $V = A_3A_3$ with $H_p$ the contributors were the victim and the suspect, and $H_d$, the contributors were the victim and one unknown person

| Ethnicity | | | Likelihood ratio |
|---|---|---|---|
| $X$ | $V$ | $S$ | |
| $a$ | $a$ | $a$ | $(1+3\theta_a)(1+4\theta_a)/\{2[\theta_a+(1-\theta_a)p_{a1}][\theta_a+(1-\theta_a)p_{a2}]\}$ |
| $a$ | $a$ | $b$ | $(1+\theta_a)(1+2\theta_a)/[2(1-\theta_a)^2p_{a1}p_{a2}]$ |
| $a$ | $b$ | $a$ | $(1+\theta_a)(1+2\theta_a)/\{2[\theta_a+(1-\theta_a)p_{a1}][\theta_a+(1-\theta_a)p_{a2}]\}$ |
| $a$ | $\bar{a}$ | $\bar{a}$ | $1/[2(1-\theta_a)p_{a1}p_{a2}]$ |

Three-allele mixture, heterozygous victim, and homozygous suspect

Consider that the DNA profiles of the mixed stain, suspect and victim are $M = \{A_1,A_2,A_3\}$, $S = A_1A_1$, and $V = A_2A_3$, respectively. The propositions are

-$H_p$ The contributors were the victim and the suspect
-$H_d$ The contributors were the victim and one unknown person.

From the unknown-based viewpoint, we have the same situations $C_1,\ldots,C_4$ as in the case of a 4-allele mixture with a heterozygous victim and suspect.

Under proposition $H_p$, we have $P(M|K,H_p) = 1$. Under $H_d$, since the mixture $M = \{A_1,A_2,A_3\}$ and the known contributor $V = A_2A_3$, $X_a$ must be one of the following forms: $A_1A_1$, $A_1A_2$, or $A_1A_3$. Thus, $P(M|K,H_d) = P(X_a = A_1A_1|K_a) + P(X_a = A_1A_2|K_a) + P(X_a = A_1A_3|K_a)$.

If $C_1$ holds, $K_a = (A_1A_1,A_2A_3)$, so:

$$\begin{aligned}
&P(M|K,H_d)\\
&= P(X_a = A_1A_1|K_a = (A_1A_1,A_2A_3))\\
&\quad + P(X_a = A_1A_2|K_a = (A_1A_1,A_2A_3))\\
&\quad + P(X_a = A_1A_3|K_a = (A_1A_1,A_2A_3))\\
&= \frac{[2\theta_a + (1-\theta_a)p_{a1}][3\theta_a + (1-\theta_a)p_{a1}]}{(1+3\theta_a)(1+4\theta_a)}\\
&\quad + 2\frac{[2\theta_a + (1-\theta_a)p_{a1}][\theta_a + (1-\theta_a)p_{a2}]}{(1+3\theta_a)(1+4\theta_a)}\\
&\quad + 2\frac{[2\theta_a + (1-\theta_a)p_{a1}][\theta_a + (1-\theta_a)p_{a3}]}{(1+3\theta_a)(1+4\theta_a)}\\
&= \frac{[2\theta_a + (1-\theta_a)p_{a1}][7\theta_a + (1-\theta_a)(p_{a1} + 2p_{a2} + 2p_{a3})]}{(1+3\theta_a)(1+4\theta_a)}.
\end{aligned}$$

The $LR$ is the reciprocal of $P(M|K,H_d)$. The $LR$s when $C_2$, $C_3$, $C_4$ hold can also be obtained in a similar way (Table 3).

Four-allele mixture, heterozygous suspect and one unknown

Consider the situation that the mixed stain did not originate from the victim. Instead, the contributors of the mixed stain were two perpetrators. Suppose one suspect was arrested and we have $M = \{A_1,A_2,A_3,A_4\}$, $S = A_1A_2$. The propositions are:

-$H_p$ The contributors were the suspect and one unknown person
-$H_d$ The contributors were two unknown persons.

This is a more complicated case and the two unknown persons with genotypes $X_1$ and $X_2$, respectively, may come from different ethnic groups. From the unknown-based viewpoint, we discuss the following situations:

-$C_1$ $X_1$, $X_2$ and $S$ come from group $a$
-$C_2$ $X_1$ and $X_2$ come from group $a$, $S$ comes from group $b$
-$C_3$ $X_1$ comes from group $a$, $X_2$ and $S$ come from group $b$
-$C_4$ $X_1$ comes from group $a$, $X_2$ comes from group $b$, and $S$ comes from group $c$
-$C_5$ $X_1$ and $S$ come from group $a$, $X_2$ comes from group $b$.

Under proposition $H_p$, since the known contributor $S = A_1A_2$, $X_a = A_3A_4$. Thus $P(M|K,H_p) = P(X_a = A_3A_4|K_a)$. Under $H_d$, since the mixture $M = \{A_1,A_2,A_3,A_4\}$ and there is no known contributor, the genotypes of the two unknown persons can have six possible combinations: $(A_1A_2,A_3A_4)$, $(A_1A_3,A_2A_4)$, $(A_1A_4,A_2A_3)$, $(A_2A_3,A_1A_4)$, $(A_2A_4,A_1A_3)$ and $(A_3A_4,A_1A_2)$. Thus we have:

$$\begin{aligned}
P(M|K,H_d) = {}& P(X_a = (A_1A_2,A_3A_4)|K_a)\\
&+ P(X_a = (A_1A_3,A_2A_4)|K_a)\\
&+ P(X_a = (A_1A_4,A_2A_3)|K_a)\\
&+ P(X_a = (A_2A_3,A_1A_4)|K_a)\\
&+ P(X_a = (A_2A_4,A_1A_3)|K_a)\\
&+ P(X_a = (A_3A_4,A_1A_2)|K_a),
\end{aligned}$$

if the two unknown persons come from the same ethnic group $a$, or:

$$\begin{aligned}
P(M|K,H_d) = {}& P(X_a = A_1A_2|K_a)P(X_b = A_3A_4|K_b)\\
&+ P(X_a = A_1A_3|K_a)P(X_b = A_2A_4|K_b)\\
&+ P(X_a = A_1A_4|K_a)P(X_b = A_2A_3|K_b)\\
&+ P(X_a = A_2A_3|K_a)P(X_b = A_1A_4|K_b)\\
&+ P(X_a = A_2A_4|K_a)P(X_b = A_1A_3|K_b)\\
&+ P(X_a = A_3A_4|K_a)P(X_b = A_1A_2|K_b),
\end{aligned}$$

if the two unknown persons come from different ethnic groups $a$ and $b$, since the genotypes between ethnic groups are statistically independent.

If $C_1$ holds, $K_a = A_1A_2$, so:

$$\begin{aligned}
P(M|K,H_p) &= P(X_a = A_3A_4|K_a = A_1A_2)\\
&= 2\frac{(1-\theta_a)p_{a3}(1-\theta_a)p_{a4}}{(1+\theta_a)(1+2\theta_a)},
\end{aligned}$$

**Table 3** Likelihood ratio for $M = \{A_1,A_2,A_3\}$, $S = A_1A_1$ and $V = A_2A_3$ with $H_p$ the contributors were the victim and the suspect, and $H_d$, the contributors were the victim and one unknown person

| Ethnicity | | | Likelihood ratio |
|---|---|---|---|
| $X$ | $V$ | $S$ | |
| $a$ | $a$ | $a$ | $(1 + 3\theta_a)(1 + 4\theta_a)/\{[2\theta_a + (1 - \theta_a)p_{a1}][7\theta_a + (1 - \theta_a)(p_{a1} + 2p_{a2} + 2p_{a3})]\}$ |
| $a$ | $a$ | $b$ | $(1 + \theta_a)(1 + 2\theta_a)/\{(1 - \theta_a)p_{a1})[5\theta_a + (1 - \theta_a)(p_{a1} + 2p_{a2} + 2p_{a3})]\}$ |
| $a$ | $b$ | $a$ | $(1 + \theta_a)(1 + 2\theta_a)/\{[2\theta_a + (1 - \theta_a)p_{a1}][3\theta_a + (1 - \theta_a)(p_{a1} + 2p_{a2} + 2p_{a3})]\}$ |
| $a$ | $\bar{a}$ | $\bar{a}$ | $1/\{p_{a1}[\theta_a + (1 - \theta_a)(p_{a1} + 2p_{a2} + 2p_{a3})]\}$ |

and

$$P(M|K,H_p)=24\frac{[\theta_a+(1-\theta_a)p_{a1}][\theta_a+(1-\theta_a)p_{a2}]\,(1-\theta_a)p_{a3}(1-\theta_a)p_{a4}}{(1+\theta_a)(1+2\theta_a)(1+3\theta_a)(1+4\theta_a)}.$$

Thus, the likelihood ratio obtained is:

$$LR=\frac{(1+3\theta_a)(1+4\theta_a)}{12[\theta_a+(1-\theta_a)p_{a1}][\theta_a+(1-\theta_a)p_{a2}]}.$$

The $LR$, when $C_2$ holds, can be derived in a similar way; see Table 4.

If $C_3$ holds, $K_a=\phi$, $K_b=A_1A_2$, thus:

$$P(M|K,H_p)=P(X_a=A_3A_4)=2(1-\theta_a)p_{a3}p_{a4}.$$

The denominator of the $LR$ is more complicated. It is evaluated as:

$$\begin{aligned}P(M|K,H_d)=\ &P(X_a=A_1A_2)P(X_b=A_3A_4|K_b=A_1A_2)\\&+P(X_a=A_1A_3)P(X_b=A_2A_4|K_b=A_1A_2)\\&+P(X_a=A_1A_4)P(X_b=A_2A_3|K_b=A_1A_2)\\&+P(X_a=A_2A_3)P(X_b=A_1A_4|K_b=A_1A_2)\\&+P(X_a=A_2A_4)P(X_b=A_1A_3|K_b=A_1A_2)\\&+P(X_a=A_3A_4)P(X_b=A_1A_2|K_b=A_1A_2)\end{aligned}$$

which can be expressed as $4(1-\theta_a)/[(1+\theta_b)(1+2\theta_b)]$ multiplied by the sum of the six terms:

$$\begin{aligned}&[p_{a1}p_{a2}(1-\theta_b)p_{b3}(1-\theta_b)p_{b4}]\\&+[p_{a1}p_{a3}[\theta_b+(1-\theta_b)p_{b2}](1-\theta_b)p_{b4}]\\&+[p_{a1}p_{a4}[\theta_b+(1-\theta_b)p_{b2}](1-\theta_b)p_{b3}]\\&+[p_{a2}p_{a3}[\theta_b+(1-\theta_b)p_{b1}](1-\theta_b)p_{b4}]\\&+[p_{a2}p_{a4}[\theta_b+(1-\theta_b)p_{b1}](1-\theta_b)p_{b3}]\\&+[p_{a3}p_{a4}[\theta_b+(1-\theta_b)p_{b1}][\theta_b+(1-\theta_b)p_{b2}]].\end{aligned}$$

The $LR$s when $C_4$ and $C_5$ hold can be obtained similarly (Table 4).

Three-allele mixture, heterozygous suspect, and one unknown

Suppose the mixed stain profile is $M=\{A_1,A_2,A_3\}$, the suspect profile is $S=A_1A_2$ and the propositions are:

- $H_p$ The contributors were the suspect and one unknown person
- $H_d$ The contributors were two unknown persons.

This case is the same as that for a 4-allele mixture with a heterozygote suspect and one unknown, except now the mixed stain only shows three alleles. If $H_p$ is true, we have:

$$\begin{aligned}P(M|K,H_p)=\ &P(X_a=A_1A_3|K_a)+P(X_a=A_2A_3|K_a)\\&+P(X_a=A_3A_3|K_a).\end{aligned}$$

If $H_d$ is true, the genotypes of the two unknown persons would have 12 possible combinations: $(A_1A_1,A_2A_3)$, $(A_2A_3,A_1A_1)$, $(A_1A_2,A_1A_3)$, $(A_1A_3,A_1A_2)$, $(A_2A_2,A_1A_3)$, $(A_1A_3,A_2A_2)$, $(A_2A_1,A_2A_3)$, $(A_2A_3,A_2A_1)$, $(A_3A_3,A_1A_2)$, $(A_1A_2,A_3A_3)$, $(A_3A_1,A_3A_2)$, and $(A_3A_2,A_3A_1)$ and $P(M|K,H_p)$ is the sum of the conditional probabilities for these combinations. For brevity, the details of the derivation are omitted; see Table 5 for expressions of $LR$s.

Four-allele mixture, two heterozygous suspects

If two suspects were arrested and the DNA profiles are $M=\{A_1,A_2,A_3,A_4\}$, $S_1=A_1A_2$ for suspect 1, and $S_2=A_3A_4$ for suspect 2, the two propositions are:

- $H_p$ The contributors were the two suspects
- $H_d$ The contributors were two unknown persons.

From the unknown-based viewpoint, we discuss the following situations:

- $C_1$ $X_1$, $X_2$, $S_1$ and $S_2$ come from group $a$
- $C_2$ $X_1$, $X_2$ and $S_1$ come from group $a$, $S_2$ comes from group $b$
- $C_3$ $X_1$ and $X_2$ come from group $a$, $S_1$ and $S_2$ not from group $a$

**Table 4** Likelihood ratio for $M=\{A_1,A_2,A_3,A_4\}$, $S=A_1A_2$ with $H_p$ the contributors were the suspect and one unknown person, and $H_d$, the contributors were two unknown persons

| Ethnicity | | | Likelihood ratio |
|---|---|---|---|
| $X_1$ | $X_2$ | $S$ | |
| $a$ | $a$ | $a$ | $(1+3\theta_a)(1+4\theta_a)/\{12[\theta_a+(1-\theta_a)p_{a1}][\theta_a+(1-\theta_a)p_{a2}]\}$ |
| $a$ | $a$ | $b$ | $(1+\theta_a)(1+2\theta_a)/[12p_{a1}p_{a2}(1-\theta_a)^2]$ |
| $a$ | $b$ | $a$ | $(1+\theta_b)(1+2\theta_b)p_{a3}p_{a4}/[2\{p_{a1}p_{a2}(1-\theta_b)^2p_{a3}p_{a4}$ $+p_{a1}p_{a3}[\theta_b+(1-\theta_b)p_{b2}](1-\theta_b)p_{b4}+p_{a1}p_{a4}[\theta_b+(1-\theta_b)p_{b2}](1-\theta_b)p_{b3}$ $+p_{a2}p_{a3}[\theta_b+(1-\theta_b)p_{b1}](1-\theta_b)p_{b4}+p_{a2}p_{a4}[\theta_b+(1-\theta_b)p_{b1}](1-\theta_b)p_{b3}$ $+p_{a3}p_{a4}[\theta_b+(1-\theta_b)p_{b1}][\theta_b+(1-\theta_b)p_{b2}]\}]$ |
| $a$ | $b$ | $c$ | $p_{a3}p_{a4}/[2(1-\theta_b)(p_{a1}p_{a2}p_{b3}p_{b4}+p_{a1}p_{a3}p_{b2}p_{b4}+p_{a1}p_{a4}p_{b2}p_{b4}$ $+p_{a2}p_{a3}p_{b1}p_{b4}+p_{a2}p_{a4}p_{b1}p_{b3}+p_{a3}p_{a4}p_{b1}p_{b2})]$ |
| $a$ | $b$ | $a$ | $(1-\theta_a)^2p_{a3}p_{a4}/[2(1-\theta_b)([\theta_a+(1-\theta_a)p_{a1}][\theta_a+(1-\theta_a)p_{a2}]p_{a3}p_{a4}$ $+[\theta_a+(1-\theta_a)p_{a1}](1-\theta_a)p_{a3}p_{b2}p_{b4}+[\theta_a+(1-\theta_a)p_{a1}](1-\theta_a)p_{a4}p_{b2}p_{b3}$ $+[\theta_a+(1-\theta_a)p_{a2}](1-\theta_a)p_{a3}p_{b1}p_{b4}+[\theta_a+(1-\theta_a)p_{a2}](1-\theta_a)p_{a4}p_{b1}p_{b3}$ $+(1-\theta_a)^2p_{a3}p_{a4}p_{b1}p_{b2})]$ |

**Table 5** Likelihood ratio for $M = \{A_1,A_2,A_3\}$, $S = A_1A_2$ with $H_p$ the contributors were the suspect and one unknown person, and $H_d$, the contributors were two unknown persons

| Ethnicity | | | Likelihood ratio |
|---|---|---|---|
| $X_1$ | $X_2$ | $S$ | |
| $a$ | $a$ | $a$ | $(1 + 3\theta_a)(1 + 4\theta_a)[5\theta_a + (1 - \theta_a)(2p_{a1} + 2p_{a2} + p_{a3})]$<br>$/\{12[\theta_a + (1 - \theta_a)p_{a1}][\theta_a + (1 - \theta_a)p_{a2}]$<br>$\times[5\theta_a + (1 - \theta_a)(p_{a1} + p_{a2} + p_{a3})]\}$ |
| $a$ | $a$ | $b$ | $(1 + \theta_a)(1 + 2\theta_a)[\theta_a + (1 - \theta_a)(2p_{a1} + 2p_{a2} + p_{a3})]$<br>$/\{12(1 - \theta_a)^2p_{a1}p_{a2}[3\theta_a + (1 - \theta_a)(p_{a1} + p_{a2} + p_{a3})]\}$ |
| a | $b$ | $b$ | $p_{a3}(1 + \theta_b)(1 + 2\theta_b)[\theta_a + (1 - \theta_a)(2p_{a1} + 2p_{a2} + p_{a3})]$<br>$/[2([\theta_b + (1 - \theta_b)p_{b1}][2\theta_b + (1 - \theta_b)p_{b1}]p_{a2}(1 - \theta_a)p_{a3}$<br>$+ [\theta_b + (1 - \theta_b)p_{b2}](1 - \theta_b)p_{b3}p_{a1}[\theta_a + (1 - \theta_a)p_{a1}]$<br>$+ 2[\theta_b + (1 - \theta_b)p_{b1}](1 - \theta_b)p_{b3}p_{a1}(1 - \theta_a)p_{a2}$<br>$+ 2[\theta_b + (1 - \theta_b)p_{b1}][\theta_b + (1 - \theta_b)p_{b2}]p_{a1}(1 - \theta_a)p_{a3}$<br>$+ [\theta_b + (1 - \theta_b)p_{b2}][2\theta_b + (1 - \theta_b)p_{b2}]p_{a1}(1 - \theta_a)p_{a3}$<br>$+ [\theta_b + (1 - \theta_b)p_{b1}](1 - \theta_b)p_{b3}p_{a2}[\theta_a + (1 - \theta_a)p_{a2}]$<br>$+ 2[\theta_b + (1 - \theta_b)p_{b2}](1 - \theta_b)p_{b3}p_{a2}(1 - \theta_a)p_{a1}$<br>$+ 2[\theta_b + (1 - \theta_b)p_{b2}][\theta_b + (1 - \theta_b)p_{b1}]p_{a2}(1 - \theta_a)p_{a3}$<br>$+ [\theta_b + (1 - \theta_b)p_{b3}](1 - \theta_b)p_{b3}p_{a1}(1 - \theta_a)p_{a2}$<br>$+ [\theta_b + (1-\theta_b)p_{b1}][\theta_b + (1 - \theta_b)p_{b2}]p_{a3}[\theta_a + (1 - \theta_a)p_{a3}]$<br>$+ 2[\theta_b + (1 - \theta_b)p_{b2}](1 - \theta_b)p_{b3}p_{a3}(1 - \theta_a)p_{a1}$<br>$+ 2[\theta_b + (1 - \theta_b)p_{b1}](1 - \theta_b)p_{b3}p_{a3}(1 - \theta_a)p_{a2})]$ |
| $a$ | $b$ | $c$ | $p_{a3}[\theta_a + (1 - \theta_a)(2p_{a1} + 2p_{a2} + p_{a3})]$<br>$/[2(p_{a1}[\theta_a + (1 - \theta_a)p_{a1}]p_{b2}(1 - \theta_b)p_{b3}$<br>$+ p_{a2}(1 - \theta_a)p_{a3}p_{b1}[\theta_b + (1 - \theta_b)p_{b1}]$<br>$+ 2p_{a1}(1 - \theta_a)p_{a2}p_{b1}(1 - \theta_b)p_{b3}$<br>$+ 2p_{a1}(1 - \theta_a)p_{a3}p_{b1}(1 - \theta_b)p_{b2}$<br>$+ p_{a2}[\theta_a + (1 - \theta_a)p_{a2}]p_{b1}(1 - \theta_b)p_{b3}$<br>$+ p_{a1}(1-\theta_a)p_{a3}p_{b2}[\theta_b + (1 - \theta_b)p_{b2}]$<br>$+ 2p_{a2}(1 - \theta_a)p_{a1}p_{b2}(1 - \theta_b)p_{b3}$<br>$+ 2p_{a2}(1 - \theta_a)p_{a3}p_{b2}(1 - \theta_b)p_{b1}$<br>$+ p_{a3}[\theta_a + (1 - \theta_a)p_{a3}]p_{b1}(1 - \theta_b)p_{b2}$<br>$+ p_{a1}(1 - \theta_a)p_{a2}p_{b3}[\theta_b + (1 - \theta_b)p_{b3}]$<br>$+ 2p_{a3}(1 - \theta_a)p_{a1}p_{b3}(1 - \theta_b)p_{b2}$<br>$+ 2p_{a3}(1 - \theta_a)p_{a2}p_{b3}(1 - \theta_b)p_{b1})]$ |
| $a$ | $b$ | $a$ | $p_{a3}(1 - \theta_a)[5\theta_a + (1 - \theta_a)(2p_{a1} + 2p_{a2} + p_{a3})]$<br>$/[2([\theta_a + (1 - \theta_a)p_{a1}][2\theta_a + (1 - \theta_a)p_{a1}]p_{b2}(1 - \theta_b)p_{b3}$<br>$+ [\theta_a + (1 - \theta_a)p_{a2}](1 - \theta_a)p_{a3}p_{a1}[\theta_b + (1 - \theta_b)p_{b1}]$<br>$+ 2[\theta_a + (1 - \theta_a)p_{a1}](1 - \theta_a)p_{a3}p_{b1}(1 - \theta_b)p_{b2}$<br>$+ 2[\theta_a + (1 - \theta_a)p_{a1}][\theta_a + (1 - \theta_a)p_{a2}]p_{b1}(1 - \theta_b)p_{b3}$<br>$+ [\theta_a + (1 - \theta_a)p_{a2}][2\theta_a + (1 - \theta_a)p_{a2}]p_{b1}(1 - \theta_b)p_{b3}$<br>$+ [\theta_a + (1 - \theta_a)p_{a1}](1 - \theta_a)p_{a3}p_{b2}(\theta_b + (1 - \theta_b)p_{b2})$<br>$+ 2[\theta_a + (1 - \theta_a)p_{a2}](1 - \theta_a)p_{a3}p_{b2}(1 - \theta_b)p_{b1}$<br>$+ 2[\theta_a + (1 - \theta_a)p_{a2}][\theta_a + (1 - \theta_a)p_{a1}]p_{b2}(1 - \theta_b)p_{b3}$<br>$+ [\theta_a + (1 - \theta_a)p_{a3}](1-\theta_a)p_{a3}p_{b1}(1 - \theta_b)p_{b2}$<br>$+ [\theta_a + (1 - \theta_a)p_{a1}][\theta_a + (1 - \theta_a)p_{a2}]p_{b3}[\theta_b + (1 - \theta_b)p_{b3}]$<br>$+ 2[\theta_a + (1 - \theta_a)p_{a2}](1 - \theta_a)p_{a3}p_{b3}(1 - \theta_b)p_{b1}$<br>$+ 2[\theta_a + (1 - \theta_a)p_{a1}](1 - \theta_a)p_{a3}p_{b3}(1 - \theta_b)p_{b2})]$ |

-$C_4$ $X_1$ and $S_1$ come from group $a$, $X_2$ and $S_2$ come from group $b$
-$C_5$ $X_1$ from group $a$, $X_2$ from group $b$ and $S_1$ and $S_2$ not from groups $a$ or $b$
-$C_6$ $X_1$, $S_1$ and $S_2$ come from group $a$, $X_2$ comes from group $b$
-$C_7$ $X_1$ and $S_1$ from group a, $X_2$ from group $b$ and $S_2$ from group $c$.

Under $H_p$, it is clear that $P(M|K,H_p) = 1$. Under $H_d$, since the mixture $M = \{A_1,A_2,A_3,A_4\}$ and there is no known contributor, the genotypes of the two unknown persons would have six possible combinations (see the 4-allele mixture with a heterozygote suspect and one unknown for a similar situation). The $LR$ is the reciprocal of $P(M|K,H_d)$. Instead of having one person typed as in the 4-allele mixture with a heterozygote suspect and one unknown we now have two persons typed. However, the $LR$s can be derived using a similar approach; see Table 6 for the $LR$s.

## Case example

We consider a case in Hong Kong which was reported in Fung and Hu (2000b). The mixed stain at D3S1358 was found to be $M = \{14, 15, 17, 18\}$. A suspect was identified and his genotype was $S_1 = (14, 17)$. If the mixed stain did not originate from the victim, we may consider the following set of propositions (P1):

-$H_p$ The contributors of the mixture were the suspect and an unknown
-$H_d$ The contributors were two unknowns.

Suppose the arrested suspect was a Caucasian. The possible ethnic groups of the unknowns are taken as Caucasian (CA), Chinese (CH), or Filipino (PH) (the largest ethnic group in Hong Kong besides Chinese). The frequencies for alleles 14, 15, 17 and 18 are

**Table 6** Likelihood ratio for $M = \{A_1, A_2, A_3, A_4\}$, $S_1 = A_1A_2$, and $S_2 = A_3A_4$ with $H_p$ the contributors were the two suspects, and $H_d$ the contributors were two unknown persons

| Ethnicity | | | | Likelihood ratio |
|---|---|---|---|---|
| $X_1$ | $X_2$ | $S_1$ | $S_2$ | |
| $a$ | $a$ | $a$ | $a$ | $(1 + 3\theta_a)(1 + 4\theta_a)(1 + 5\theta_a)(1 + 6\theta_a)/\{24[\theta_a + (1 - \theta_a)p_{a1}]$<br>$\times[\theta_a + (1 - \theta_a)p_{a2}][\theta_a + (1 - \theta_a)p_{a3}][\theta_a + (1 - \theta_a)p_{a4}]\}$ |
| $a$ | $a$ | $a$ | $b$ | $(1 + \theta_a)(1 + 2\theta_a)(1 + 3\theta_a)(1 + 4\theta_a)/\{24(1 - \theta_a)^2 p_{a3}p_{a4}$<br>$\times[\theta_a + (1 - \theta_a)p_{a1}][\theta_a + (1 - \theta_a)p_{a2}]\}$ |
| $a$ | $a$ | $\bar{a}$ | $\breve{a}$ | $(1 + \theta_a)(1 + 2\theta_a)/[24(1 - \theta_a)^3 p_{a1}p_{a2}p_{a3}p_{a4}]$ |
| $a$ | $b$ | $a$ | $b$ | $(1 + \theta_a)(1 + 2\theta_a)(1 + \theta_b)(1 + 2\theta_b)/(4\{[\theta_a + (1 - \theta_a)p_{a1}]$<br>$\times [\theta_a + (1 - \theta_a)p_{a2}][\theta_b + (1 - \theta_b)p_{b3}][\theta_b + (1 - \theta_b)p_{b4}]$<br>$+ [\theta_a + (1 - \theta_a)p_{a1}](1 - \theta_a)p_{a3}(1 - \theta_b)p_{b2}[\theta_b + (1 - \theta_b)p_{b4}]$<br>$+ [\theta_a + (1 - \theta_a)p_{a1}](1 - \theta_a)p_{a4}(1 - \theta_b)p_{b2}[\theta_b + (1 - \theta_b)p_{b3}]$<br>$+ [\theta_a + (1 - \theta_a)p_{a2}](1 - \theta_a)p_{a3}(1 - \theta_b)p_{b1}[\theta_b + (1 - \theta_b)p_{b4}]$<br>$+ [\theta_a + (1 - \theta_a)p_{a2}](1 - \theta_a)p_{a4}(1 - \theta_b)p_{b1}[\theta_b + (1 - \theta_b)p_{b3}]$<br>$+ (1 - \theta_a)^2 p_{a3}p_{a4}(1 - \theta_b)^2 p_{b1}p_{b2}\})$ |
| $a$ | $b$ | $\overline{ab}$ | $\overline{ab}$ | $1/[4(1 - \theta_a)(1 - \theta_b)(p_{a1}p_{a2}p_{b3}p_{b4} + p_{a1}p_{a3}p_{b2}p_{b4}$<br>$+ p_{a1}p_{a4}p_{b2}p_{b3} + p_{a2}p_{a3}p_{b1}p_{b4} + p_{a2}p_{a4}p_{b1}p_{b3} + p_{a3}p_{a4}p_{b1}p_{b2})]$ |
| $a$ | $b$ | $a$ | $a$ | $(1 + 3\theta_a)(1 + 4\theta_a)/[4(1 - \theta_b)$<br>$\times([\theta_a + (1 - \theta_a)p_{a1}][\theta_a + (1 - \theta_a)p_{a2}]p_{b3}p_{b4}$<br>$+ [\theta_a + (1 - \theta_a)p_{a1}][\theta_a + (1 - \theta_a)p_{a3}]p_{b2}p_{b4}$<br>$+ [\theta_a + (1 - \theta_a)p_{a1}][\theta_a + (1 - \theta_a)p_{a4}]p_{b2}p_{b3}$<br>$+ [\theta_a + (1 - \theta_a)p_{a2}][\theta_a + (1 - \theta_a)p_{a3}]p_{b1}p_{b4}$<br>$+ [\theta_a + (1 - \theta_a)p_{a2}][\theta_a + (1 - \theta_a)p_{a4}]p_{b1}p_{b3}$<br>$+ [\theta_a + (1 - \theta_a)p_{a3}][\theta_a + (1 - \theta_a)p_{a4}]p_{b1}p_{b2})]$ |
| $a$ | $b$ | $a$ | $c$ | $(1 + \theta_a)(1 + 2\theta_a)/[4(1 - \theta_b)((1 - \theta_a)^2 p_{a3}p_{a4}p_{b1}p_{b2}$<br>$+ [\theta_a + (1 - \theta_a)p_{a2}](1 - \theta_a)p_{a4}p_{b1}p_{b3}$<br>$+ [\theta_a + (1 - \theta_a)p_{a2}](1 - \theta_a)p_{a3}p_{b1}p_{b4}$<br>$+ [\theta_a + (1 - \theta_a)p_{a1}](1 - \theta_a)p_{a4}p_{b2}p_{b3}$<br>$+ [\theta_a + (1 - \theta_a)p_{a1}](1 - \theta_a)p_{a3}p_{b2}p_{b4}$<br>$+ [\theta_a + (1 - \theta_a)p_{a1}][\theta_a + (1 - \theta_a)p_{a2}]p_{b3}p_{b4})]$ |

$\overline{ab}$ means not groups $a$ or $b$

**Table 7** Likelihood ratios with different ethnic groups of two unknowns for proposition sets ($P_1$ contributors were the suspect and an unknown, versus, contributors were two unknowns; and $P_2$ contributors were two suspects, versus, contributors were two unknowns)

| Ethnicity | | | |
|---|---|---|---|
| $X_1$ | $X_2$ | $P_1$ | $P_2$ |
| CA | CA | 1.95 | 36.6 |
| CH | CH | 12.3 | 341 |
| PH | PH | 13.0 | 285 |
| CA | CH | 3.10 | 62.9 |
| CH | CA | 2.39 | 62.9 |
| CA | PH | 2.72 | 56.0 |
| PH | CA | 2.66 | 56.0 |
| CH | PH | 9.43 | 262 |
| PH | CH | 12.0 | 262 |

given as CA: 0.187, 0.213, 0.223 and 0.127, CH: 0.033, 0.331, 0.239 and 0.056, PH: 0.026, 0.267, 0.286 and 0.088, respectively (Pu et al. 1999; Wong et al. 2001). The coancestry coefficients taken are the same value, 0.03, for all three ethnic groups. Based on the formulae given above, we obtained the *LR*s for various ethnic combinations of the unknowns with genotypes $X_1$ and $X_2$. These results are listed in the column 3 of Table 7. There are great differences among the *LR*s. For example, the *LR* if the two unknowns were Filipinos is more than 6 times greater than if they were Caucasians. The ethnicity of the unknowns has a great effect on the *LR*.

If a second suspect (Caucasian) was identified and his profile was S2 = (15,18), we may be interested in the following set of propositions (P2):

-$H_p$ The contributors were the two suspects
-$H_d$ The contributors were two unknowns.

As in the previous situation, we consider different ethnic groups for $X_1$ and $X_2$. The formulae in the section on 4-allele mixture and two heterozygote suspects are helpful in deriving the *LR* results which are shown in column 4 of Table 7. In this case, the *LR* has the highest value when both unknowns were Chinese, while the *LR* is the smallest if both were Caucasians. They are in a ratio of about 9:1. Again, the ethnicity of contributors has a substantial effect on the *LR* values.

## Concluding remarks

When the contributors to a mixed sample come from more than one structured population, the likelihood ratio is used to assess the weight of the DNA evidence. This paper derives the *LR*s for six common mixture cases. The usefulness of the derived formulae is illustrated using a casework example in Hong Kong. The results show that the ethnicity of contributors can have substantial effect on *LR*s.

This paper considers mixture problems with indistinguishable contributors. Sometimes, the major and minor contributors to the mixed sample can be inferred by considering the peak height of the alleles. If so, the number of possible contributors would be reduced and the problem would be simpler. The method we employed here can be used to derive the corresponding (simpler) *LR*s.

**Acknowledgments** The authors thank an associate editor and two referees for their helpful comments that greatly improved the presentation of the paper. The first author is partially supported by the Hong Kong RGC Competitive Earmarked Research Grant (HKU 7136/97H) and the second author is supported in part by the SEU Science Foundation (9207011101).

## References


Balding DJ, Nichols RA (1994) DNA profile match probability calculation: how to allow for population stratification, relatedness, database selection and single bands. Forensic Sci Int 64: 125–140

Buckleton JS, Evett IW, Weir BS (1998) Setting bounds for the likelihood ratio when multiple hypotheses are postulated. Sci Justice 38:23–26

Curran JM, Triggs CM, Buckleton J, Weir BS (1999) Interpreting DNA mixtures in structured populations. J Forensic Sci 44: 987–995

Devlin B, Risch N (1993) Physical properties of VNTR data, and their impact on a test of allelic independence. Am J Hum Genet 53:324–329

Evett IW, Buffery C, Wilcott G, Stoney D (1991) A guide to interpreting single locus profiles of DNA mixtures in forensic case. J Forensic Sci Soc 31:41–47

Fukshansky N, Bär W (1998) Interpreting forensic DNA evidence on the basis of hypotheses testing. Int J Legal Med 111:62–66

Fukshansky N, Bär W (1999) Biostatistical evaluation of mixed stains with contributors of different ethnic origin. Int J Legal Med 112:383–387

Fung WK (1996) Tests on independence of VNTR alleles in the Chinese population in Hong Kong. In: Lee JC, Johnson WO, Zellner A (eds) Modelling and prediction. Springer, New York, pp 294–304

Fung WK, Hu YQ (2000a) Interpreting forensic DNA mixtures: allowing for uncertainty in population substructure and dependence. J R Stat Soc A 163:241–254

Fung WK, Hu YQ (2000b) Interpreting DNA mixtures based on the NRC-II recommendation 4.1. Forensic Sci Commun 2(4): 1–9 (available at http: //www.fbi.gov/hq/lab/fsc/backissu/oct2000/fung.htm)

Fung WK, Hu YQ (2001) The evaluation of mixed stains from different ethnic origin: general result and common cases. Int J Legal Med (in press)

National Research Council Report (1996) The evaluation of forensic DNA evidence. National Academy Press, Washington DC

Pu CE, Hsieh CM, Chen MY, Wu FC, Sun CF (1999) Genetic variation at nine STR loci in populations from the Philippines and Thailand living in Taiwan. Forensic Sci Int 106:1–6

Triggs CM, Harbison SA, Buckleton J (2000) The calculation of DNA match probabilities in mixed race populations. Sci Justice 40:33–38

Weir BS, Triggs CM, Starling L, Stowell LI, Walsh KAJ, Buckleton J (1997) Interpreting DNA mixtures. J Forensic Sci 42: 213–222

Wong DM, Law MY, Fung WK, Chan KL, Li C, Lun TS, Lai KM, Cheung KY, Chiu CT (2001) Population data for 12 STR loci in Hong Kong Chinese. Int J Legal Med 114:281–284

Wright S (1951) The genetical structure of populations. Ann Eugen 15:323–354