NCJRS
National Criminal Justice Reference Service

Office of Justice Programs
BJA BJS NIJ OJJDP OVC SMART

The author(s) shown below used Federal funding provided by the U.S. Department of Justice to prepare the following resource:

Document Title:	**Interpreting DNA Mixtures**
Author(s):	**B S Weir; C M Triggs; L Starling; L I Stowell; K A J Walsh; J Buckleton**
Document Number:	**172960**
Date Received:	**1996**
Award Number:	**95-IJ-CX-0007**

This resource has not been published by the U.S. Department of Justice. This resource is being made publicly available through the Office of Justice Programs' National Criminal Justice Reference Service.

Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

If you have issues viewing or accessing this file contact NCJRS.gov.

172960

Interpreting DNA mixtures.

B.S. Weir
Program in Statistical Genetics, Department of Statistics,
North Carolina State University, Raleigh, NC 27695-8203, USA

C.M. Triggs
Department of Mathematics, University of Auckland,
Private Bag, Auckland, New Zealand

L. Starling, L.I. Stowell, K.A.J. Walsh, J. Buckleton
ESR:Forensic, Mt Albert Science Centre,
Private Bag 92-021, Auckland, New Zealand

February 26, 1996

PROPERTY OF
National Criminal Justice Reference Service (NCJRS)
Box 6000
Rockville, MD 20849-6000

Introduction

The use of genetic markers has proved invaluable in identifying perpetrators, or exonerating falsely accused suspects, in a large variety of crimes. However, the interpretation of genetic profiles from biological samples requires care when the samples contain material from more than one person, and this is especially common in rape cases. The sample may contain material from the victim or consensual sexual partners as well as from the perpetrator, or there may be multiple perpetrators. Evidence of this type was discussed by Evett et al. (1991), and more recently by Aitken (1995). The present discussion is intended to expand on these treatments by generalizing the calculations to arbitrary numbers of contributors, by allowing for "null" components in the profile and by allowing for dependencies among the profile frequencies of the contributors. Details are given for several specific situations.

To keep the treatment as simple as possible, only discrete genetic systems will be treated. These include binned RFLP markers, conventional blood group markers, and the current STR markers. It will be assumed that a profile has been determined for an evidentiary sample, that it indicates more than one contributor to the sample, and that it may contain the profiles of one or more known people. The task is to assign a numerical weight to this evidence, and this in turn requires that databases are available from which to estimate frequencies of components of the profile.

Likelihood Ratios

The standard framework for assigning weight to evidence will be used. A likelihood ratio for comparing two alternative explanations, C and $\bar{C}$, for the evidentiary profile is formed as the ratio of the probabilities of the profile under the alternatives. The likelihood ratio is written as L and the profile as E, so

$$L = \frac{\Pr(E|C)}{\Pr(E|\bar{C})}$$

If a jury is being asked to make a choice between explanations C and $\bar{C}$, it can be told "The profile E is L times more likely to have arisen under explanation C than under $\bar{C}$." A full Bayesian analysis would use the likelihood ratio to modify prior odds on C to produce posterior odds after the evidence:

$$\frac{\Pr(C|E)}{\Pr(\bar{C}|E)} = L \times \frac{\Pr(C)}{\Pr(\bar{C})}$$

The evidentiary profile in a rape case is likely to have been determined from a vaginal swab. Although preferential extraction (Gill et al., 1985) may minimize the DNA from the victim in the male fraction, it may not eliminate her DNA. Reference samples of both the victim and any consensual partners will be needed for the purposes of identifying any profile components that must have been contributed by the perpetrator. If all evidentiary profile components match the components of the victim, possible consensual partners and a suspect in the case, with no unmatched components, then the prosecution explanation would be that E can be accounted for entirely by these known people. The probability $\Pr(E|C)$ is 1. The defense explanation may be that, although

the victim and consensual partners contributed to the profile, the perpetrator is unknown. The probability $\Pr(E|\bar{C})$ involves the chance of obtaining the perpetrator's profile.

As an example, suppose a victim reports having been raped by a single perpetrator and to have had no recent consensual partners. The only genetic locus typed has alleles a, b, c, d in the evidentiary profile E, the victim is known to have alleles a, b and a suspect is known to have alleles c, d. The explanations of interest are C: the victim and suspect contributed to E, and $\bar{C}$: the victim but not the suspect contributed to E. If the profiling system does not allow false negatives, $\Pr(E|C) = 1$. If $\bar{C}$ means that the perpetrator is some unknown person unrelated to the suspect, this person must have genotype cd. If that person belongs to some population in which the frequencies of alleles c, d are p_c, p_d, then $\Pr(E|\bar{C}) = 2p_cp_d$. The likelihood ratio is therefore $1/2p_cp_d$, the reciprocal of the frequency of cd genotypes. If several loci are typed, this quantity can be multiplied over loci for which there are independent allele frequencies.

This result is standard, although a point that bears repeating is that "relevant" population refers to those people who might be considered potential perpetrators. The population is defined by circumstances of the crime, and not by any attributes of the suspect. The circumstances may prescribe a particular ethnic group, in which case frequencies should be used for that group. Otherwise, frequencies from different ethnic groups should be considered and possibly weighted according to population composition (Buckleton et al., 1987). The ethnicity of the suspect is not material – although see the section below on population substructuring.

The circumstances of a crime such as rape may dictate that the profile represents the contributions of exactly two contributors. Generally the profile itself, however, merely points to more than one contributor and several explanations could be considered. Suppose C states that the victim $K_1 : ab$, the suspect $K_2 : cd$ and some unknown third person U_1 were contributors, whereas $\bar{C}$ states that the victim K_1 and two unknown people, U_1 and U_2, were the contributors. As before, U_1, U_2 must account for alleles c, d but now they may also possess alleles a, b. The likelihood ratio is found from summing over all genotypes possible for U_1, U_2. Under C, U_1 may have any of the genotypes $g_1 = aa, ab, bb, ac, bc, cc, ad, bd, cd, dd$ and each gives a probability of 1 for the profile. These ten types have a total frequency of $(p_a + p_b + p_c + p_d)^2$, and any other genotype causes E to have zero probability. Under $\bar{C}$, the pair U_1, U_2 can have any of 37 combinations of genotypes g_1, g_2, each of which makes E certain whereas any other combination makes E impossible. The likelihood ratio is

$$\begin{aligned} L &= \frac{\Pr(E|K_1, K_2, U_1)}{\Pr(E|K_1, U_1, U_2)} \\ &= \frac{\sum_{g_1} \Pr(E|K_1, K_2, U_1 = g_1) \Pr(U_1 = g_1)}{\sum_{g_1} \sum_{g_2} \Pr(E|K_1, U_1 = g_1, U_2 = g_2) \Pr(U_1 = g_1, U_2 = g_2)} \\ &= \frac{(p_a + p_b + p_c + p_d)^2}{2p_cp_d[6(p_a + p_b)^2 + 6(p_a + p_b)(p_c + p_d) + (p_c + p_d)^2 - p_cp_d]} \\ &> \frac{1}{12p_cp_d} \end{aligned}$$

In terms of the notation introduced below, $L = P_1(\phi|abcd)/P_2(cd|abcd)$.

General Method

In general, the evidentiary profile contains a set of alleles $\{e\}$ at some locus. There may be known contributors to the profile, having alleles $\{k\}$ between them, where the set $\{k\}$ is wholly contained within the set $\{e\}$. Under at least one of the alternative explanations, it is necessary to account for alleles $\{u\}$ from unknown people. If there are x unknown contributors to the profile, then what is needed is the probability $P_x(\{u\}|\{e\})$ that these people have alleles $\{u\}$ between them, but do not have any alleles not in $\{e\}$. The x unknown people may have alleles in the set $\{k\}$. When there are no unknown contributors, the symbol P_ϕ will be used, and this probability is one. Some examples will clarify the meaning of this probability.

Profiles with one allele

If the profile has only alleles of type a, then all contributors must be aa homozygotes. (The possibility of unseen alleles will be addressed later.) Assuming independence of alleles within loci, the probability that one unknown contributor has allele a, and no allele other than a, is

$$P_1(a|a) = p_a^2$$

The probability of a, and only a, from x contributors is

$$P_x(a|a) = p_a^{2x}$$

If a known contributor to the profile is aa, then there are no alleles that need to be accounted for. The probability needed, for x unknowns, is $P_x(\phi|a)$, where ϕ indicates the empty set. Although these unknown people need not have a specific allele, they may have alleles (a) in the profile other than ϕ, so

$$P_x(\phi|a) = p_a^{2x}$$

Profiles with two alleles

If the profile has alleles a, b, then all contributors must be of genotype aa, ab or bb. The probability that one unknown contributor has allele a, and no allele other than a or b, is

$$\begin{aligned} P_1(a|ab) &= p_a^2 + 2p_a p_b \\ &= (p_a + p_b)^2 - p_b^2 \end{aligned}$$

The generalization to x contributors is

$$P_x(a|ab) = (p_a + p_b)^{2x} - p_b^{2x}$$

since the case of all x contributors being homozygous bb is not possible. If known contributors already have alleles a, b, then unknowns need not have any specific alleles but cannot have alleles other than a, b, so

$$P_x(\phi|ab) = (p_a + p_b)^{2x}$$

Finally, if there are no known contributors, then both alleles a, b must be carried by the unknown contributors

$$P_x(ab|ab) = (p_a + p_b)^{2x} - p_a^{2x} - p_b^{2x}$$

since not all contributors can be of the same homozygous type.

Profiles with three alleles

If the profile has alleles a, b, c, then all contributors must be of genotype aa, ab, bb, ac, bc or cc. If known contributors already have alleles a, b, c, then unknowns need not have any specific alleles but cannot have alleles other than a, b, c, so

$$P_x(\phi|abc) = (p_a + p_b + p_c)^{2x}$$

The probability that x unknown contributors have allele a, and no allele other than a, b or c, is

$$P_x(a|abc) = (p_a + p_b + p_c)^{2x} - (p_b + p_c)^{2x}$$

The probability that x unknown contributors have alleles a, b, and no allele other than a, b or c, is

$$P_x(ab|abc) = (p_a + p_b + p_c)^{2x} - (p_b + p_c)^{2x} - (p_a + p_c)^{2x} + p_c^{2x}$$

Finally, if there are no known contributors, then all alleles a, b, c must be carried by the unknown contributors, which means that $x > 1$, and

$$P_x(abc|abc) = (p_a + p_b + p_c)^{2x} - (p_a + p_b)^{2x} - (p_b + p_c)^{2x} - (p_a + p_c)^{2x} + p_a^{2x} + p_b^{2x} + p_c^{2x}$$

since not all contributors can be of the same homozygous type.

Profiles with n alleles.

The general pattern is now clear. For contributors to have at least $U = \{u_i\}$ but not more than $E = \{e_i\}$

$$P_x(U|E) = (T_0)^{2x} - \sum_j (T_{1_j})^{2x} + \sum_{j,k} (T_{2_{j,k}})^{2x} - \sum_{j,k,l} (T_{3_{j,k,l}})^{2x} + \ldots$$

where T_0 is the sum of probabilities of all elements in E, T_{1_j} is the sum of probabilities of all elements in E except for the jth element in U, $T_{2_{j,k}}$ is the sum of probabilities of all elements in E except for the jth and kth elements of U, and so on.

It is unlikely that actual profiles will involve more than six alleles, so explicit equations for the case of 4, 5 or 6 alleles are given in Appendix 1.

Conditional Probabilities

The machinery is now in place to derive the likelihood ratio for any situation. A set of conditional probabilities $\Pr(E|C)$ is listed in Table 1 for some common profiles E and common explanations C. As before, K's indicate known contributors and U's unknown contributors. Subscripts distinguish between different knowns or unknowns.

As an illustration of the use of Table 1, suppose an evidentiary sample E has only three alleles a, b, c and includes contributions from a victim and a perpetrator. If the victim K_1 has genotype ab and a suspect K_2 has profile aa, then the explanation C including the suspect still requires a contributor U_1 for allele c. The alternative explanation $\bar{C}$ may still include the victim but excludes the suspect, and so includes some unknown contributors $U_1, \ldots, U_x$.

The number of contributors to the sample must be specified before conditional probabilities can be evaluated. It may be that the circumstances of the crime specify the number of contributors. If the number was $x = 3$, then in this example

$$\begin{aligned} L &= \frac{P_1(c|abc)}{P_2(c|abc)} \\ &= \frac{(p_a + p_b + p_c)^2 - (p_a + p_b)^2}{(p_a + p_b + p_c)^4 - (p_a + p_b)^4} \\ &= \frac{1}{(p_a + p_b + p_c)^2 + (p_a + p_b)^2} \end{aligned}$$

If x is not specified, there appears to be two possible approaches. The first notices that the number of contributors to the sample is just one number, even if it is unknown. This means that both $\Pr(E|C)$ and $\Pr(E|\bar{C})$ are evaluated with the same number of contributors (known and unknown), and several values of x can be used. In the absence of mutually acceptable priors on values of x, these calculations need to be kept separate. The other approach allows the numbers of contributors to be set independently for C and $\bar{C}$. In this example, the prosecution may set the number of contributors to 3, and the defense may make a plausible, and parsimonious, argument that there was only one unknown contributor since that lowers the likelihood ratio to one:

$$L = \frac{P_1(c|abc)}{P_1(c|abc)}$$

In cases like this, the profiles at other loci are likely to help resolve the issue of the number of contributors.

Profiles with Unseen Alleles

An RFLP locus can give rise to a single band c on an electrophoretic gel if it is homozygous cc, or if it is heterozygotes cn for c and some unseen allele n. For evidentiary profile a, b, c, suppose the victim K_1 has type ab and a suspect K_2 has type c. The evidence is certain under C that the contributors were K_1, K_2. However, if $\bar{C}$ includes K_1 but excludes K_2, there are complications with typing systems that may produce "unseen" or "null" alleles. It will need to be assumed that if n is unseen in a reference sample, it may also be unseen in an evidentiary sample. The evidentiary

profile should therefore be written a, b, c, n. Although it still the case that $\Pr(E|C) = 1$, if some person unrelated to the victim contributed to E, then that person must possess allele c and not possess an allele, other than a or b, that would be "seen." The possible genotypes for this person are ac, bc, cc or nc and

$$\begin{aligned}
\Pr(E|\bar{C}) &= P_1(c|abcn) \\
&= (p_a + p_b + p_c + p_n)^2 - (p_a + p_b + p_n)^2 \\
&= p_c(2p_a + 2p_b + 2p_n + p_c) \\
&< p_c(2p_a + 2p_b + 2p_c + 2p_n) \\
&\leq 2p_c
\end{aligned}$$

The strict inequality sign is needed since allele c is known not to have a zero frequency. The likelihood ratio is

$$\begin{aligned}
L &= \frac{1}{P_1(c|abcn)} \\
&= \frac{1}{p_c(2p_a + 2p_b + p_c + 2p_n)} \qquad (1) \\
&> \frac{1}{2p_c} \qquad (2)
\end{aligned}$$

This may be overly conservative. There are often upper bounds on the values of p_n, and these could be used in place of $1 - p_a - p_b - p_c$. It is unusual for RFLP frequencies to have frequencies greater than 0.1, and estimates of null frequencies have all been less than this (Chakraborty et al., 1994). Setting $p_a = p_b = p_c = p_n = 0.1$ gives a value of $L = 14$ from Equation 1 but $L = 5$ from Equation 2.

Equation 2 was given by Evett et al. (1991). The important feature of this result is that it is not the reciprocal of the frequency of the suspect's profile, as in the four-allele case. Proper analyses of mixed profiles require the use of likelihood ratios.

Although null alleles require care, the machinery previously established is adequate provided the null n is added to the evidentiary profile. It need not appear in the set U in the expression $P_x(U|E)$.

Ethnicity of Unknown People

To the extent that profile frequencies differ between different ethnic groups, calculations involving single unknown contributors to a profile need to consider the ethnicity of that person, or to give values for different ethnicities. Presenting several values can convey to a jury the dependence of L on ethnicity. For more than one unknown contributor, a full analysis will consider all possible combinations of ethnicities for these people.

If g_i is the probability that a random person in the population relevant to a crime is from ethnic group G_i, then the probability of the evidentiary profile when victim K and unknown U are the contributors is

$$\Pr(E|K, U) = \sum_i \Pr(E|K, U \epsilon G_i) \Pr(U \epsilon G_i)$$

$$= \sum_i g_i E_i$$

Here, ϵ means "belongs to" and E_i is the conditional probability using frequencies for group G_i. Unless the g_i are known, it will be necessary to present separate E_i values or maybe a range of these values.

Suppose two unknown people U_1, U_2 have ethnicities i, j, and this occurs with frequency g_{ij}. The probability needed is now

$$\begin{aligned} \Pr(E|K, U_1, U_2) &= \sum_i \sum_j \Pr(E|K, U_1 \epsilon G_i, U_2 \epsilon G_j) \Pr(U_1 \epsilon G_i, U_2 \epsilon G_j) \\ &= \sum_i \sum_j g_{ij} E_{ij} \end{aligned}$$

Here E_{ij} is the conditional probability of the profile calculated with frequencies from groups G_i and G_j. There is no need to distinguish between the orders G_i, G_j and G_j, G_i for the two unknown people, and the probabilities are equal in either case: $E_{ij} = E_{ji}$. Furthermore, it can be shown that

$$\Pr(E|K; U_1 \epsilon G_i, U_2 \epsilon G_j \text{ or } U_1 \epsilon G_j, U_2 \epsilon G_i) = E_{ij} = E_{ji}$$

Even if the group probabilities g_i were known, it is unlikely that g_{ij} values would be known and there seems to be no alternative to reporting all separate E_{ij} values, or their range.

The complications increase when both explanations C and $\bar{C}$ involve unknown contributors. If C involves one and $\bar{C}$ involves two unknowns, then a range of ratios E_i/E_{ij} will be needed. The only exception will be when the circumstances of the crime, such as location or a reliable determination by the victim or an eyewitness, specify the ethnicities of the assailant(s). Then only those ethnicities need be used.

For more than one unknown contributor, having to consider more than one ethnicity requires a modification to the computing algorithm. For $x = 2$ contributors from ethnic groups i and j, let p_{a_i} be the frequency of allele a in ethnic group i.

$$\begin{aligned} P_2(\phi|a) &= p_{a_i}^2 p_{a_j}^2 \\ P_2(a|a) &= p_{a_i}^2 p_{a_j}^2 \\ P_2(a|ab) &= (p_{a_i} + p_{b_i})^2 (p_{a_j} + p_{b_j})^2 - p_{b_i}^2 p_{b_j}^2 \\ P_2(a|abc) &= (p_{a_i} + p_{b_i} + p_{c_i})^2 (p_{a_j} + p_{b_j} + p_{c_j})^2 - (p_{b_i} + p_{c_i})^2 (p_{b_j} + p_{c_j})^2 \end{aligned}$$

and so forth.

Population Substructure

It has been assumed that allele frequencies, within and between loci, are independent. This assumption may break down when the population, or ethnic group, for which frequencies are calculated

has substructure. If subgroups within the population have different allelic frequencies, then the frequency of a set of alleles in the whole population differs from the product of each of the frequencies even if this product rule holds within each of the subgroups. This is the Wahlund effect of population genetics.

It is necessary to restate the probabilities in the likelihood ratio. Return to the case of an evidentiary profile E with alleles a, b, c, d where the victim K_1 of type ab is known to be a contributor so that the perpetrator P must be of type cd. The evidence E can be regarded as the profiles of K_1 and P. A suspect K_2 of type cd is not excluded as a contributor. Explanation C has K_1, K_2 as contributors, whereas $\bar{C}$ has K_1, U as contributors for some unknown person U, and U must be the perpetrator. Let expressions such as $K_1 = ab$ mean that K_1 has genotype ab and $K_2 \equiv P$ mean that K_2 and P are the same person. Then the likelihood ratio is

$$\begin{aligned} L &= \frac{\Pr(E|C)}{\Pr(E|\bar{C})} \\ &= \frac{\Pr(K_1, P|K_2; P \equiv K_2)}{\Pr(K_1, P|K_2; P \equiv U)} \\ &= \frac{1}{\Pr(P = cd|K_2 = cd; P \equiv U)} \end{aligned}$$

Under the independence assumption, $\Pr(P = cd|K_2 = cd; P \equiv U)$ is just $\Pr(P = cd) = 2p_c p_d$, but in a substructured population

$$\Pr(P = cd|K_2 = cd; P \equiv U) = \frac{2[p_c + \theta(1 - p_c)][p_d + \theta(1 - p_d)]}{(1 + \theta)(1 + 2\theta)]}$$

This expression was given by Balding and Nichols (1994) and is an approximation to a more detailed value given by Weir (1994). The quantity θ, also written as F_{ST}, is a measure of population structure. It provides the component of variance of allele frequencies among subpopulations, and is defined as the probability that two alleles in the same subpopulation are identical by descent. For large human populations, θ is unlikely to exceed 0.01 (Morton, 1992).

In the present context, allowing for substructure is needed only when P and K_2 belong to the same population or ethnic group. This is the context in which the ethnic background of the suspect is material. For situations with multiple perpetrators, and the need to calculate probabilities like the E_{ij} discussed above, allowing for substructure will be necessary whenever calculations involve more than one person from the same ethnic group.

Effects of Relatives

Conditional frequencies of the kind $\Pr(P = cd|K_2 = cd; P \equiv U)$ are different from marginal frequencies $\Pr(P = cd)$ when P, K_2 belong to the same substructured population. They are also different when P and K_2 are related. A general formulation was given by Weir (1994), when the relatives belonged to a substructured population. As the effects of relatedness outweigh the effects of substructuring, however, the latter can usually be ignored.

To illustrate the effects of relatives for mixed stains, consider the case where explanation $\bar{C}$ is that the suspect is the brother of one of two perpetrators. Suppose victim $K_1 : ab$ reports

being raped by two men. Two suspects are identified, $K_2 : cd$ and $K_3 : ef$, whose genotypes are included in the evidentiary profile a, b, c, d, e, f. Explanation C is that K_2, K_3 are the perpetrators and explanation $\bar{C}$ is that the perpetrators are a brother U_1 of K_2 and an unknown person U_2. Evidently $\Pr(E|C) = 1$. For $\Pr(E|\bar{C})$, account must be taken of all possible pairs of genotypes for U_1 and U_2. These two people must have alleles c, d, e, f between them. The conditional frequencies needed for two brothers are:

$$\begin{aligned}
\Pr(U_1 = cd|K_2 = cd) &= (1 + p_c + p_d + 2p_cp_d)/4 \\
\Pr(U_1 = ce|K_2 = cd) &= p_e(1 + 2p_c)/4 \\
\Pr(U_1 = de|K_2 = cd) &= p_e(1 + 2p_d)/4 \\
\Pr(U_1 = ef|K_2 = cd) &= p_ep_f/2
\end{aligned}$$

Corresponding values for other pairs of relatives are given in Table 2. The likelihood ratio for C versus $\bar{C}$ is

$$L = \frac{2}{p_ep_f(1 + 2p_c + 2p_d + 8p_cp_d)}$$

Discussion

Likelihood ratios for a series of common situations are shown in Appendix 2. In each case, the Ks indicate people whose genotype is known and who are believed to have contributed to the profile under the stated explanation. These people may be victims or suspects. The U's indicate people who have not been typed, and will generally indicate unknown perpetrators. All the expressions involving allele frequencies are for contributors from the same ethnic group. Frequencies are assumed to be independent.

The interpretation of mixed stains is straightforward in the likelihood ratio context. Alternative explanations for the mixed stain profiles need to be specified, and then compared on the basis of the probabilities of the profile under those explanations. Calculations that consider only single contributors are without a logical foundation. The 1992 report of the United States National Research Council recommended "If a suspect's pattern is found within the mixed pattern, the appropriate frequency to assign such a ' match' the sum of frequencies of all the genotypes that are contained within (i.e. that are a subset of) the mixed pattern." To see why this statement has no merit, consider the four-allele mixed profile a, b, c, d for a crime in which the victim has type ab and both a suspect and the perpetrator have type cd. As shown above, the likelihood ratio is $1/2p_cp_d$. The NRC method, in essence, uses the smaller ratio $1/(p_a + p_b + p_c + p_d)^2$. For a locus with only four alleles, this last ratio is 1, and there is no probative value to the inclusion of the suspect's profile in the mixed stain. There is no probative value even if alleles c, d each occur only with frequency 0.01 and alleles a, b each occur with frequency 0.49. The correct likelihood ratio is 5,000. If the alternative explanation $\bar{C}$ is that both contributors are unknown, the NRC recommendation still leads to $1/(p_a + p_b + p_c + p_d)^2$, whereas the true likelihood ratio is $1/12p_ap_bp_cp_d$. The NRC ignores the fact that not all pairs of people have alleles a, b, c, d, and only these alleles, between them. Similar comments apply to calculations for "Random Men Not Excluded" which are based on analogous discussions in the paternity testing literature (e.g. Walker et al., 1983).

When there are only two explanations for an item of evidence, the likelihood ratio is separate from prior probabilities, which are rightly within the province of the jury. For more than two explanations, however, the general form of Bayes' theorem means that priors and likelihoods become intermixed in the same expression. Probabilities involving x unknown contributors, where x is not known, are needed. Rather than this being a limitation of the Bayesian approach, it is a fundamental problem of legal inference to which there is no simple solution. Difficult questions should not be replaced with deceptively simple questions such as those answered by the NRC rule. The present discussion has avoided the issue by considering alternative explanations in pairs. The resulting analysis does not provide a single number, but the direction of the evidence can be explained with care to a jury. Prosecution and defense may well cast different interpretations on the magnitude of the numbers, but there should be as little disagreement on which numbers to present as there is for stains from single contributors.

Acknowledgements

The wise counsel of Dr. Ian Evett is greatly appreciated. This work was supported in part by NIH grant GM45344, and by award 95-IJ-CX-0007 from the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. Points of view in this document are those of the authors and do not necessarily represent the official position of the U.S. Department of Justice.

References

Aitken, C.G.G. 1995. Statistics and the Evaluation of Evidence for Forensic Scientists. Wiley, New York

Balding, D.J. and R.A. Nichols. 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, J.S., K.A.J. Walsh, G.A.F. Seber and D.G. Woodfield. 1987. A stratified approach to the compilation of blood group frequency surveys. J. Forensic Sci. Soc. 27:103–112.

Chakraborty, R., Y. Zhong, L. Jin and B. Budowle. 1994. Nondetectability of restriction fragment sizes within and between loci in RFLP typing of DNA. Am. J. Human Genet. 55:391–401.

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

Gill, P., A.J. Jeffreys and D.J. Werrett. 1985. Forensic application of DNA 'fingerprints.' Nature 318:577–579.

Morton, N.E. 1992. Genetic structure of forensic populations. Proc. Natl. Acad. Sci. USA 8:2556–2560.

Walker, R.H., R.J. Duquesnoy, E.R. Jennings, H.D. Krause, C.L. Lee and H.F. Polesky. (Editors). 1983. Inclusion Probabilities in Parentage Testing. American Association of Blood Banks, Arlington, VA.

Weir, B.S. 1994. The effects of inbreeding on forensic calculations. Ann. Rev. Genet. 28:597–621.

Appendix 1

Four-allele profile

$$
\begin{aligned}
P_x(\phi|abcd) &= (p_a+p_b+p_c+p_d)^{2x} \\
P_x(a|abcd) &= (p_a+p_b+p_c+p_d)^{2x} - (p_b+p_c+p_d)^{2x} \\
P_x(ab|abcd) &= (p_a+p_b+p_c+p_d)^{2x} - (p_b+p_c+p_d)^{2x} - (p_a+p_c+p_d)^{2x} + (p_c+p_d)^{2x} \\
P_x(abc|abcd) &= (p_a+p_b+p_c+p_d)^{2x} - (p_b+p_c+p_d)^{2x} - (p_a+p_c+p_d)^{2x} - (p_a+p_b+p_d)^{2x} \\
&\quad + (p_c+p_d)^{2x} + (p_b+p_d)^{2x} + (p_a+p_d)^{2x} - p_d^{2x}, x>1 \\
P_x(abcd|abcd) &= (p_a+p_b+p_c+p_d)^{2x} - (p_b+p_c+p_d)^{2x} - (p_a+p_c+p_d)^{2x} - (p_a+p_b+p_d)^{2x} \\
&\quad - (p_a+p_b+p_c)^{2x} \\
&\quad + (p_c+p_d)^{2x} + (p_b+p_d)^{2x} + (p_b+p_c)^{2x} + (p_a+p_d)^{2x} + (p_a+p_c)^{2x} + (p_a+p_b)^{2x} \\
&\quad - p_a^{2x} - p_b^{2x} - p_c^{2x} - p_d^{2x}, x>1
\end{aligned}
$$

Five-allele profile

$$
\begin{aligned}
P_x(\phi|abcde) &= (p_a+p_b+p_c+p_d+p_e)^{2x} \\
P_x(a|abcde) &= (p_a+p_b+p_c+p_d+p_e)^{2x} - (p_b+p_c+p_d+p_e)^{2x} \\
P_x(ab|abcde) &= (p_a+p_b+p_c+p_d+p_e)^{2x} - (p_b+p_c+p_d+p_e)^{2x} - (p_a+p_c+p_d+p_e)^{2x} \\
&\quad + (p_c+p_d+p_e)^{2x} \\
P_x(abc|abcde) &= (p_a+p_b+p_c+p_d+p_e)^{2x} - (p_b+p_c+p_d+p_e)^{2x} - (p_a+p_c+p_d+p_e)^{2x} \\
&\quad - (p_a+p_b+p_d+p_e)^{2x} \\
&\quad + (p_c+p_d+p_e)^{2x} + (p_b+p_d+p_e)^{2x} + (p_a+p_d+p_e)^{2x} - (p_d+p_e)^{2x}, x>1 \\
P_x(abcd|abcde) &= (p_a+p_b+p_c+p_d+p_e)^{2x} - (p_b+p_c+p_d+p_e)^{2x} - (p_a+p_c+p_d+p_e)^{2x} \\
&\quad - (p_a+p_b+p_d+p_e)^{2x} - (p_a+p_b+p_c+p_e)^{2x} \\
&\quad + (p_c+p_d+p_e)^{2x} + (p_b+p_d+p_e)^{2x} + (p_b+p_c+p_e)^{2x} + (p_a+p_d+p_e)^{2x} \\
&\quad + (p_a+p_c+p_e)^{2x} + (p_a+p_b+p_e)^{2x} \\
&\quad - (p_d+p_e)^{2x} - (p_c+p_e)^{2x} - (p_b+p_e)^{2x} - (p_a+p_e)^{2x} + p_e^{2x}, x>1 \\
P_x(abcde|abcde) &= (p_a+p_b+p_c+p_d+p_e)^{2x} - (p_b+p_c+p_d+p_e)^{2x} - (p_a+p_c+p_d+p_e)^{2x} \\
&\quad - (p_a+p_b+p_d+p_e)^{2x} - (p_a+p_b+p_c+p_e)^{2x} - (p_a+p_b+p_c+p_d)^{2x} \\
&\quad + (p_c+p_d+p_e)^{2x} + (p_b+p_d+p_e)^{2x} + (p_b+p_c+p_e)^{2x} + (p_b+p_c+p_d)^{2x} \\
&\quad + (p_a+p_d+p_e)^{2x} + (p_a+p_c+p_e)^{2x} + (p_a+p_c+p_d)^{2x} + (p_a+p_b+p_e)^{2x} \\
&\quad + (p_a+p_b+p_d)^{2x} + (p_a+p_b+p_c)^{2x} \\
&\quad - (p_a+p_b)^{2x} - (p_a+p_c)^{2x} - (p_a+p_d)^{2x} - (p_a+p_e)^{2x} - (p_b+p_c)^{2x} \\
&\quad - (p_b+p_d)^{2x} - (p_b+p_e)^{2x} - (p_c+p_d)^{2x} - (p_c+p_e)^{2x} - (p_d+p_e)^{2x} \\
&\quad + p_a^{2x} + p_b^{2x} + p_c^{2x} + p_d^{2x} + p_e^{2x}, x>2
\end{aligned}
$$

Six-allele profile

$$
\begin{aligned}
P_x(\phi|abcdef) &= (p_a+p_b+p_c+p_d+p_e+p_f)^{2x} \\
P_x(a|abcdef) &= (p_a+p_b+p_c+p_d+p_e+p_f)^{2x} - (p_b+p_c+p_d+p_e+p_f)^{2x} \\
P_x(ab|abcdef) &= (p_a+p_b+p_c+p_d+p_e+p_f)^{2x} \\
&\quad - (p_b+p_c+p_d+p_e+p_f)^{2x} - (p_a+p_c+p_d+p_e+p_f)^{2x} \\
&\quad + (p_c+p_d+p_e+p_f)^{2x} \\
P_x(abc|abcdef) &= (p_a+p_b+p_c+p_d+p_e+p_f)^{2x} - (p_b+p_c+p_d+p_e+p_f)^{2x} \\
&\quad - (p_a+p_c+p_d+p_e+p_f)^{2x} - (p_a+p_b+p_d+p_e+p_f)^{2x} \\
&\quad + (p_c+p_d+p_e+p_f)^{2x} + (p_b+p_d+p_e+p_f)^{2x} + (p_a+p_d+p_e+p_f)^{2x} \\
&\quad - (p_d+p_e+p_f)^{2x}, x>1 \\
P_x(abcd|abcdef) &= (p_a+p_b+p_c+p_d+p_e+p_f)^{2x} \\
&\quad - (p_b+p_c+p_d+p_e+p_f)^{2x} - (p_a+p_c+p_d+p_e+p_f)^{2x} \\
&\quad - (p_a+p_b+p_d+p_e+p_f)^{2x} - (p_a+p_b+p_c+p_e+p_f)^{2x} \\
&\quad + (p_c+p_d+p_e+p_f)^{2x} + (p_b+p_d+p_e+p_f)^{2x} + (p_b+p_c+p_e+p_f)^{2x} \\
&\quad + (p_b+p_c+p_d+p_e)^{2x} + (p_a+p_d+p_e+p_f)^{2x} + (p_a+p_c+p_e+p_f)^{2x} \\
&\quad - (p_d+p_e+p_f)^{2x} - (p_c+p_e+p_f)^{2x} - (p_b+p_e+p_f)^{2x} - (p_a+p_e+p_f)^{2x} \\
&\quad + (p_e+p_f)^{2x}, x>1 \\
P_x(abcde|abcdef) &= (p_a+p_b+p_c+p_d+p_e+p_f)^{2x} - (p_b+p_c+p_d+p_e+p_f)^{2x} \\
&\quad - (p_a+p_c+p_d+p_e+p_f)^{2x} - (p_a+p_b+p_d+p_e+p_f)^{2x} \\
&\quad - (p_a+p_b+p_c+p_e+p_f)^{2x} - (p_a+p_b+p_c+p_d+p_f)^{2x} \\
&\quad + (p_c+p_d+p_e+p_f)^{2x} + (p_b+p_d+p_e+p_f)^{2x} + (p_b+p_c+p_e+p_f)^{2x} \\
&\quad + (p_b+p_c+p_d+p_f)^{2x} + (p_a+p_d+p_e+p_f)^{2x} + (p_a+p_c+p_e+p_f)^{2x} \\
&\quad + (p_a+p_c+p_d+p_f)^{2x} + (p_a+p_b+p_e+p_f)^{2x} + (p_a+p_b+p_d+p_f)^{2x} \\
&\quad + (p_a+p_b+p_c+p_f)^{2x} \\
&\quad - (p_d+p_e+p_f)^{2x} - (p_c+p_e+p_f)^{2x} - (p_c+p_d+p_f)^{2x} - (p_b+p_e+p_f)^{2x} \\
&\quad - (p_b+p_d+p_f)^{2x} - (p_b+p_c+p_f)^{2x} - (p_a+p_e+p_f)^{2x} - (p_a+p_d+p_f)^{2x} \\
&\quad - (p_a+p_c+p_f)^{2x} - (p_a+p_b+p_f)^{2x} \\
&\quad + (p_e+p_f)^{2x} + (p_d+p_f)^{2x} + (p_c+p_f)^{2x} + (p_b+p_f)^{2x} + (p_a+p_f)^{2x} - p_f^{2x}, x>2
\end{aligned}
$$

$$
\begin{aligned}
P_x(abcdef|abcdef) \;=\;& (p_a+p_b+p_c+p_d+p_e+p_f)^{2x} \\
& -(p_b+p_c+p_d+p_e+p_f)^{2x}-(p_a+p_c+p_d+p_e+p_f)^{2x} \\
& -(p_a+p_b+p_d+p_e+p_f)^{2x}-(p_a+p_b+p_c+p_e+p_f)^{2x} \\
& -(p_a+p_b+p_c+p_d+p_f)^{2x}-(p_a+p_b+p_c+p_d+p_e)^{2x} \\
& +(p_c+p_d+p_e+p_f)^{2x}+(p_b+p_d+p_e+p_f)^{2x}+(p_b+p_c+p_e+p_f)^{2x} \\
& +(p_b+p_c+p_d+p_f)^{2x}+(p_b+p_c+p_d+p_e)^{2x}+(p_a+p_d+p_e+p_f)^{2x} \\
& +(p_a+p_c+p_e+p_f)^{2x}+(p_a+p_c+p_d+p_e)^{2x}+(p_a+p_b+p_e+p_f)^{2x} \\
& +(p_a+p_b+p_d+p_f)^{2x}+(p_a+p_b+p_d+p_e)^{2x}+(p_a+p_b+p_c+p_f)^{2x} \\
& +(p_a+p_b+p_c+p_e)^{2x}+(p_a+p_b+p_c+p_e)^{2x}+(p_a+p_b+p_c+p_d)^{2x} \\
& -(p_d+p_e+p_f)^{2x}-(p_c+p_e+p_f)^{2x}-(p_c+p_d+p_f)^{2x}-(p_c+p_d+p_e)^{2x} \\
& -(p_b+p_e+p_f)^{2x}-(p_b+p_d+p_f)^{2x}-(p_b+p_d+p_e)^{2x}-(p_b+p_c+p_f)^{2x} \\
& -(p_b+p_c+p_e)^{2x}-(p_b+p_c+p_d)^{2x}-(p_a+p_e+p_f)^{2x}-(p_a+p_d+p_f)^{2x} \\
& -(p_a+p_d+p_e)^{2x}-(p_a+p_c+p_f)^{2x}-(p_a+p_c+p_e)^{2x}-(p_a+p_c+p_d)^{2x} \\
& -(p_a+p_b+p_f)^{2x}-(p_a+p_b+p_e)^{2x}-(p_a+p_b+p_d)^{2x}-(p_a+p_b+p_c)^{2x} \\
& +(p_e+p_f)^{2x}+(p_d+p_f)^{2x}+(p_d+p_e)^{2x}+(p_c+p_f)^{2x}+(p_c+p_e)^{2x} \\
& +(p_c+p_d)^{2x}+(p_b+p_f)^{2x}+(p_b+p_e)^{2x}+(p_b+p_d)^{2x}+(p_b+p_c)^{2x} \\
& +(p_a+p_f)^{2x}+(p_a+p_e)^{2x}+(p_a+p_d)^{2x}+(p_a+p_c)^{2x}+(p_a+p_b)^{2x} \\
& -p_a^{2x}-p_b^{2x}-p_c^{2x}-p_d^{2x}-p_e^{2x}-p_f^{2x}, x>2
\end{aligned}
$$

Appendix 2

Without Null Alleles

Profile with one allele: $E = a$

$$C : K_1 = aa \;\; \bar{C} : U_1$$
$$L = \frac{1}{P_1(a|a)} = \frac{1}{p_a^2}$$

$$C : K_1 = aa, K_2 = aa, \;\; , \bar{C} : K_1 = aa, U_1$$
$$L = \frac{1}{P_1(\phi|a)} = \frac{1}{p_a^2}$$

Profile with two alleles: $E = a, b$

$$C : K_1 = ab \;\; \bar{C} : U_1$$
$$L = \frac{1}{P_1(ab|ab)} = \frac{1}{2p_a p_b}$$

$$C : K_1 = aa, U_1 \;\; \bar{C} : U_1, U_2$$
$$L = \frac{P_1(b|ab)}{P_2(ab|ab)} = \frac{(p_a+p_b)^2-p_a^2}{(p_a+p_b)^4-p_a^4-p_b^4}$$

$$C : K_1 = aa, K_2 = ab \;\; \bar{C} : K_1 = aa, U_1$$
$$L = \frac{1}{P_1(b|ab)} = \frac{1}{(p_a+p_b)^2-p_a^2}$$

$$C : K_1 = aa, K_2 = bb \;\; \bar{C} : K_1 = aa, U_1$$
$$L = \frac{1}{P_1(b|ab)} = \frac{1}{(p_a+p_b)^2-p_a^2}$$

$$C : K_1 = aa, K_2 = aa, U_1 \;\; \bar{C} : K_1 = aa, U_1, U_2$$
$$L = \frac{P_1(b|ab)}{P_2(b|ab)} = \frac{1}{(p_a+p_b)^2+p_a^2}$$

$$C : K_1 = ab, K_2 = aa \;\; \bar{C} : K_1 = ab, U_1$$
$$L = \frac{1}{P_2(\phi|ab)} = \frac{1}{(p_a+p_b)^4}$$

$$C : K_1 = ab, K_2 = ab \;\; \bar{C} : K_1 = ab, U_1$$
$$L = \frac{1}{P_2(\phi|ab)} = \frac{1}{(p_a+p_b)^4}$$

Profile with three alleles: $E = a, b, c$

$$C : K_1 = aa, U_1 \;\; \bar{C} : U_1, U_2$$
$$L = \frac{P_1(bc|abc)}{P_2(abc|abc)}$$
$$box = \frac{(p_a+p_b+p_c)^2-(p_a+p_c)^2-(p_a+p_b)^2+p_a^2}{(p_a+p_b+p_c)^4-(p_a+p_c)^4-(p_a+p_b)^4-(p_b+p_c)^4+p_a^4+p_b^4+p_c^4}$$

$$C : K_1 = ab, U_1 \;\; \bar{C} : U_1, U_2$$
$$L = \frac{P_1(c|abc)}{P_2(abc|abc)}$$
$$= \frac{(p_a+p_b+p_c)^2-(p_a+p_b)^2}{(p_a+p_b+p_c)^4-(p_a+p_c)^4-(p_a+p_b)^4-(p_b+p_c)^4+p_a^4+p_b^4+p_c^4}$$

$$C : K_1 = aa, K_2 = bc \;\; \bar{C} : K_1 = aa, U_1$$
$$L = \frac{1}{P_1(bc|abc)}$$
$$= \frac{1}{(p_a+p_b+p_c)^2-(p_a+p_c)^2-(p_a+p_b)^2+p_a^2}$$

$$C : K_1 = aa, K_2 = aa, U_1 \;\; \bar{C} : K_1 = aa, U_1, U_2$$
$$L = \frac{P_1(bc|abc)}{P_2(bc|abc)}$$
$$= \frac{(p_a+p_b+p_c)^2-(p_a+p_c)^2-(p_a+p_b)^2+p_a^2}{(p_a+p_b+p_c)^4-(p_a+p_c)^4-(p_a+p_b)^4+p_a^4}$$

$$C : K_1 = aa, K_2 = ab, U_1 \;\; \bar{C} : K_1 = aa, U_1, U_2$$
$$L = \frac{P_1(c|abc)}{P_2(bc|abc)}$$
$$= \frac{(p_a+p_b+p_c)^2-(p_a+p_b)^2}{(p_a+p_b+p_c)^4-(p_a+p_c)^4-(p_a+p_b)^4+p_a^4}$$

$$C : K_1 = ab, K_2 = ac \;\; \bar{C} : K_1 = ab, U_1$$
$$L = \frac{1}{P_1(c|abc)} \frac{1}{(p_a+p_b+p_c)^2-(p_a+p_b)^2}$$

$$C : K_1 = ab, K_2 = cc \;\; \bar{C} : K_1 = ab, U_1$$
$$L = \frac{1}{P_1(c|abc)} \frac{1}{(p_a+p_b+p_c)^2-(p_a+p_b)^2}$$

Profile with four alleles: $E = a, b, c, d$

$$C : K_1 = ab, U_1 \;\; \bar{C} : U_1, U_2$$
$$L = \frac{P_1(cd|abcd)}{P_2(abcd|abcd)}$$

$$C : K_1 = aa, U_1, U_2 \;\; \bar{C} : U_1, U_2, U_3$$
$$L = \frac{P_2(bcd|abcd)}{P_3(abcd|abcd)}$$

$$C : K_1 = aa, K_2 = bc, U_1 \;\; \bar{C} : K_1 = aa, U_1, U_2$$
$$L = \frac{P_1(d|abcd)}{P_2(bcd|abcd)}$$

$$C : K_1 = aa, K_2 = bb, U_1 \;\; \bar{C} : K_1 = aa, U_1, U_2$$
$$L = \frac{P_1(cd|abcd)}{P_2(bcd|abcd)}$$

$$C : K_1 = aa, K_2 = ab, U_1 \;\; \bar{C} : K_1 = aa, U_1, U_2$$
$$L = \frac{P_1(cd|abcd)}{P_2(abcd|abcd)}$$

$$C : K_1 = aa, K_2 = aa, U_1, U_2 \;\; \bar{C} : K_1 = aa, U_1, U_2, U_3$$
$$L = \frac{P_1(bcd|abcd)}{P_3(abcd|abcd)}$$

$$C : K_1 = ab, K_2 = cd \;\; \bar{C} : K_1 = ab, U_1$$
$$L = \frac{1}{P_1(cd|abcd)}$$

$$C : K_1 = ab, K_2 = cc, U_1 \;\; \bar{C} : K_1 = ab, U_1, U_2$$
$$L = \frac{P_1(d|abcd)}{P_2(cd|abcd)}$$

$$C : K_1 = ab, K_2 = ac, U_1 \;\; \bar{C} : K_1 = ab, U_1, U_2$$
$$L = \frac{P_1(d|abcd)}{P_2(cd|abcd)}$$

$$C : K_1 = ab, K_2 = aa, U_1 \;\; \bar{C} : K_1 = ab, U_1, U_2$$
$$L = \frac{P_1(cd|abcd)}{P_2(cd|abcd)}$$

$$C : K_1 = ab, K_2 = ab, U_1 \;\; \bar{C} : K_1 = ab, U_1, U_2$$
$$L = \frac{P_1(cd|abcd)}{P_2(cd|abcd)}$$

With Null Alleles

Known genotypes with single letters, e.g. $K_1 = a$ may be homozygotes aa or heterozygotes an where n is a null allele.

Profile with one allele: $E = a$

$$C : K_1 = a \;\; \bar{C} : U_1$$
$$L = \frac{1}{P_1(a|an)} = \frac{1}{(p_a+p_n)^2-p_n^2}$$

$$C : K_1 = a, U_1 = a \;\; \bar{C} : K_1 = a, U_1$$
$$L = \frac{1}{P_1(\phi|an)} = \frac{1}{(p_a+p_n)^2}$$

Profile with two alleles: $E = a, b$

$$C : K_1 = a, U_1 \;\; \bar{C} : U_1, U_2$$
$$L = \frac{P_1(b|abn)}{P_2(ab|abn)}$$

$$C : K_1 = a, K_2 = ab \;\; \bar{C} : K_1 = a, U_1$$
$$L = \frac{1}{P_1(b|abn)}$$
$$C : K_1 = a, K_2 = b \;\; \bar{C} : K_1 = a, U_1$$
$$L = \frac{1}{P_1(b|abn)}$$
$$C : K_1 = a, K_2 = a, U_1 \;\; \bar{C} : K_1 = a, U_1, U_2$$
$$L = \frac{P_1(b|abn)}{P_2(b|abn)}$$
$$C : K_1 = a, K_2 = ab \;\; \bar{C} : K_1 = ab, U_1$$
$$L = \frac{1}{P_1(\phi|abn)}$$

Profile with three alleles: $E = a, b, c$

$$C : K_1 = a, U_1 \;\; \bar{C} : U_1, U_2$$
$$L = \frac{P_1(bc|abcn)}{P_2(abc|abcn)}$$

$$C : K_1 = a, K_2 = bc \;\; \bar{C} : K_1 = a, U_1$$
$$L = \frac{1}{P_1(bc|abcn)}$$

$$C : K_1 = a, K_2 = a, U_1 \;\; \bar{C} : K_1 = a, U_1, U_2$$
$$L = \frac{P_1(bc|abcn)}{P_2(bc|abcn)}$$

$$C : K_1 = a, K_2 = ab, U_1 \;\; \bar{C} : K_1 = a, U_1, U_2$$
$$L = \frac{P_1(c|abcn)}{P_2(bc|abcn)}$$

$$C : K_1 = ab, K_2 = c \;\; \bar{C} : K_1 = ab, U_1$$
$$L = \frac{1}{P_2(c|abcn)}$$

$$C : K_1 = ab, K_2 = a, U_1 \;\; \bar{C} : K_1 = ab, U_1, U_2$$
$$L = \frac{P_1(c|abcn)}{P_2(c|abcn)}$$

Profile with four alleles: $E = a, b, c, d$

$C : K_1 = a, U_1, U_2 \;\; \bar{C} : U_1, U_2, U_3$
$L = \frac{P_2(bcd|abcdn)}{P_3(abcd|abcdn)}$

$C : K_1 = a, K_2 = bc, U_1 \;\; \bar{C} : K_1 = a, U_1, U_2$
$L = \frac{P_1(d|abcdn)}{P_2(bcd|abcdn)}$

$C : K_1 = a, K_2 = b, U_1 \;\; \bar{C} : K_1 = a, U_1, U_2$
$L = \frac{P_1(cd|abcdn)}{P_2(bcd|abcdn)}$

$C : K_1 = a, K_2 = ab, U_1 \;\; \bar{C} : K_1 = a, U_1, U_2$
$L = \frac{P_1(cd|abcdn)}{P_2(bcd|abcdn)}$

$C : K_1 = a, K_2 = a, U_1, U_2 \;\; \bar{C} : K_1 = a, U_1, U_2, U_3$
$L = \frac{P_1(bcd|abcdn)}{P_3(bcd|abcdn)}$

$C : K_1 = ab, K_2 = c, U_1 \;\; \bar{C} : K_1 = ab, U_1, U_2$
$L = \frac{P_1(d|abcdn)}{P_2(cd|abcdn)}$

$C : K_1 = ab, K_2 = a, U_1 \;\; \bar{C} : K_1 = ab, U_1, U_2$
$L = \frac{P_1(cd|abcdn)}{P_2(cd|abcdn)}$

Table 1 Some common profiles and explanations.

Profile E	Explanation C^*	$\Pr(E\|C)$
aa	$U_1, \ldots, U_x$	$P_x(a\|a)$
	$K_1 = aa$	1
	$K_1 = aa, U_1 \ldots, U_x$	$P_x(\phi\|a)$
ab	$U_1, \ldots, U_x$	$P_x(ab\|ab)$
	$K_1 = ab$	1
	$K_1 = aa, U_1, \ldots, U_x$	$P_x(b\|ab)$
	$K_1 = ab, U_1, \ldots, U_x$	$P_x(\phi\|ab)$
	$K_1 = aa, K_2 = ab, U_1, \ldots, U_x$	$P_x(\phi\|ab)$
abc	$U_1, U_2, \ldots, U_x$	$P_x(abc\|abc), x > 1$
	$K_1 = aa, U_1, \ldots, U_x$	$P_x(bc\|abc)$
	$K_1 = ab, U_1, \ldots, U_x$	$P_x(c\|abc)$
	$K_1 = aa, K_2 = bc, U_1, \ldots, U_x$	$P_x(\phi\|abc)$
	$K_1 = ab, K_2 = bc, U_1, \ldots U_x$	$P_x(\phi\|abc)$
$abcd$	$U_1, U_2, \ldots, U_x$	$P_x(abcd\|abcd), x > 1$
	$K_1 = aa, U_1, U_2, \ldots U_x$	$P_x(bcd\|abcd), x > 1$
	$K_1 = ab, U_1, \ldots, U_x$	$P_x(cd\|abcd)$
	$K_1 = ab, K_2 = cd$	1
	$K_1 = ab, K_2 = cd, U_1, \ldots, U_x$	$P_x(\phi\|abcd)$
	$K_1 = ab, K_2 = ac, U_1, \ldots, U_x$	$P_x(d\|abcd)$

*$U_1, \ldots, U_x$ means x unknowns

PROPERTY OF
National Criminal Justice Reference Service (NCJRS)
Box 6000
Rockville, MD 20849-6000