Available online at www.sciencedirect.com
SCIENCE DIRECT
Forensic Science International 160 (2006) 90–101
Forensic Science International
www.elsevier.com/locate/forsciint

DNA commission of the International Society of Forensic Genetics: Recommendations on the interpretation of mixtures

P. Gill [a,*], C.H. Brenner [b], J.S. Buckleton [c], A. Carracedo [d], M. Krawczak [e], W.R. Mayr [f], N. Morling [g], M. Prinz [h], P.M. Schneider [i], B.S. Weir [j]

[a] Forensic Science Service, Trident Court, 2960 Solihull Parkway, Birmingham, UK
[b] Forensic Science Group, School of Public Health, University of California, Berkeley, CA 510-339-1911, USA
[c] ESR, Private Bag 92021, Auckland, New Zealand
[d] Institute of Legal Medicine, Faculty of Medicine, University of Santiago de Compostela, 15705 Santiago de Compostela, Spain
[e] Institute of Medical Informatics and Statistics, Kiel, Germany
[f] Division of Blood Group Serology, Medical University of Vienna, Austria
[g] Department of Forensic Genetics, Institute of Forensic Medicine, University of Copenhagen, Copenhagen, Denmark
[h] Office of the Chief Medical Examiner, Department of Forensic Biology, 520 First Avenue, New York, NY 10016, USA
[i] Institute of Legal Medicine University Clinic of Cologne, Melatengürtel, 60-62 D 50823 Köln, Germany
[j] University of Washington, Department of Biostatistics, Box 357232, Seattle, WA 98195, USA

Received 4 April 2006; accepted 10 April 2006
Available online 5 June 2006

Abstract

The DNA commission of the International Society of Forensic Genetics (ISFG) was convened at the 21st congress of the International Society for Forensic Genetics held between 13 and 17 September in the Azores, Portugal. The purpose of the group was to agree on guidelines to encourage best practice that can be universally applied to assist with mixture interpretation. In addition the commission was tasked to provide guidance on low copy number (LCN) reporting. Our discussions have highlighted a significant need for continuing education and research into this area. We have attempted to present a consensus from experts but to be practical we do not claim to have conveyed a clear vision in every respect in this difficult subject. For this reason, we propose to allow a period of time for feedback and reflection by the scientific community. Then the DNA commission will meet again to consider further recommendations.

Keywords: STR typing; Biostatistical analysis; Likelihood ratio; Probability of exclusion; Mixtures; ISFG DNA commission

1. The general approaches used to interpret DNA profiles

There are two different methods in common use to report DNA profiles: these are the classical profile probability approach and the likelihood ratio approach. See Buckleton [1] and Balding [2] for a full discussion of the various methods to interpret evidence.

1.1. The profile probability approach

In the forensic context the profile probability approach presents the probability of the evidentiary DNA profile (E) under a stated hypothesis (H_o). This hypothesis may be as simple as saying that the DNA profile is from a person unrelated to the suspect. The probability is written formally as $\Pr(E|H_o)$, where Pr is an abbreviation for 'probability' and the vertical line, or conditioning bar, is an abbreviation for 'given'. For a single-contributor stain, under the approximation that profiles from unrelated people are independent, this probability is the frequency of occurrence of the profile in the population.

1.2. The likelihood ratio

An extension of the profile probability approach works with the probabilities of the evidence under two or more alternative hypotheses about the source(s) of the profile. A typical analysis of a crime sample has the prosecution hypothesis (H_p) and the defence hypothesis (H_d). For a profile with more than one

* Corresponding author.
E-mail address: dnapgill@compuserve.com (P. Gill).

0379-0738/$ – see front matter © 2006 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.forsciint.2006.04.009

contributor, the prosecution may hypothesise that the suspect (S) and one unknown (U) person were the contributors, whereas the defence may hypothesise that there were two unknown contributors U_1 and U_2. The likelihood ratio (LR) compares the probabilities of the evidence under these alternative hypotheses:

$$\mathrm{LR} = \frac{\Pr(E|H_\mathrm{p})}{\Pr(E|H_\mathrm{d})}$$

If the LR is greater than one, then the evidence favours H_p but if it is less than one then the evidence favours H_d.

In the single-contributor case, the probability of the evidence profile under H_p (the suspect is the contributor) is one and the LR reduces to the reciprocal of the probability of the stain profile if it did not come from the suspect. Ignoring the possibility of relatives and population structure this is just the population frequency of the profile as would have been given by the profile probability approach.

But, it is worth noting that under certain easily defined circumstances, involving low level crime stain profiles, the probability of the numerator $\Pr(E|H_\mathrm{p})$ is less than one. When this happens the LR gives a number that is less than that obtained using the profile probability approach. Examples are given in Appendix A (stutter) and Appendix B (drop-out).

To evaluate mixtures population genetics principles are applied—to the extent that the suspect (if innocent) and the perpetrator are suspected to be from the same sub-population then an F_St correction is desirable.

1.3. Types of alleles

There are three kinds of alleles in a crime stain profile:

A. alleles that are unmistakeable;
B. alleles that may be masked by an artefact such as a stutter;
C. alleles that have dropped out completely and are therefore not detected.

We emphasise the need to carry out appropriate biochemical and genetic tests—e.g. the analysis of multiple stains in order to obtain the best results possible before carrying out the statistical analysis.

2. A comparison of the probability of exclusion method versus the LR method

The probability of exclusion $\Pr(Ex)$, or random man not excluded (RMNE) [3,4] or the complementary probability of inclusion $\Pr(I)$ entails a binary view of alleles, meaning that alleles are only present or absent, and further if present are observed. Using the method therefore entails the implicit assumption that all alleles are either in category A or at least – and this necessitates counting all artefacts that might mask an allele in the RMNE calculation – in category A or B. In particular it is problematical to apply the method when there are loci which, under the hypothesis being considered of the suspect at hand, appear to have alleles in category C. We have seen many instances in which laboratories do just this, usually by omitting from the RMNE calculation the inconvenient loci. Such a calculation implies, certainly incorrectly, that among the ''random men'' considered for comparison by the calculation only the same loci would be used for inculpation/exculpation as those being considered for the present suspect. It fails to acknowledge that choosing the omitted loci is suspect-centric and therefore prejudicial against the suspect. (If, on the other hand, a locus is eliminated from analysis simply because it is a poor result showing no alleles at all, then of course there is no prejudice in ignoring it.)

Consequently the exclusion method may be justified under the following circumstances:

1. It is known that all relevant alleles are in category A.
   Or:
2. It is known that all relevant alleles are in category A or B.
3. All of the suspect's alleles are present and the report is conditional, e.g. ''The suspect is not excluded as being a major (or salient) contributor, whereas x% of random men would be''.

The method is usually quite conservative provided it is properly applied as described above.

The advantage of the LR framework is that stutter and drop-out can be assessed probabilistically [5–7] (Appendices A and B), and it is the only way to provide a meaningful calculation based on the probability of the evidence under H_p and H_d. The RMNE method has considerable intuitive appeal but usually entails an unrealistically simple model of DNA evidence and is therefore restricted in its use to unambiguous profiles. Even in those cases RMNE has the further shortcomings as it does not make full use of the evidence.

A likelihood ratio approach is therefore preferred. There is a broad consensus view on this point that originates from the original recommendation of the NRC II report [8].

Various advantages and disadvantages have been suggested in relation to the LR and RMNE approaches; summarised by Clayton and Buckleton [9]. In particular, Weir [10] states that exclusion probabilities ''often rob the items of any probative value'' and Brenner [11] states ''the exclusion probability usually discards a lot of information compared to the correct likelihood ratio approach''. Michael Krawczak states: ''In my view, this is not a question of 'correct' and 'incorrect', but of 'efficient' and 'inefficient'. The RMNE simply does not use as much of the information included in the data as the LR approach but, conceptually, they are equivalent. The RMNE is based on a different statistical model than the LR approach. So the two methods are bound to give different answers in one and the same case. The RMNE result is still correct, given the model, but is not an optimal result since the model does not make efficient use of the available information''.

However, if the model is used outside the constraints of its working limitations, then there is no reason to suppose that the concept of 'conservativeness' still applies. An example follows:

Consider a genetic marker, such as a SNP that has only two alleles *a* and *b* in the population. For a two-contributor stain with both alleles (*ab*), no-one in the population is excluded so the RMNE probability is one. However, if the suspect is of type *aa*, and it is a common type, then the LR assuming two contributors is less than one. Although unlikely to concern STR multiplexes in current use, this would extend to the multi-allele case when nearly all of the allele types at the locus are present in the stain.[1]

Clayton and Buckleton [9] report two advantages for the RMNE approach: (a) it does not require an assumption of the number of contributors to a mixture and (b) it is easier to explain in court. Otherwise the RMNE usually results in an underestimate of the strength of evidence in numerical terms (except for unusual situations where all or most alleles are present at a locus). Nevertheless, this may be an important consideration. The US DNA Advisory Board [3] states: ''The calculation is particularly useful in complex mixtures, because it requires no assumptions about the identity or number of contributors to a mixture''.

- *Recommendation* 1: The likelihood ratio is the preferred approach to mixture interpretation. The RMNE approach is restricted to DNA profiles where the profiles are unambiguous. If the DNA crime stain profile is low level and some minor alleles are the same size as stutters of major alleles, and/or if drop-out is possible, then the RMNE method may not be conservative.

3. Court acceptance of the LR approach

In addition, an argument that may be put forward is that courts are unwilling to accept the LR method. Whereas we recognise that there are restrictions that are placed upon scientists by legal systems, we recommend that the scientist should always prepare his/her evidence using the LR method wherever possible. We accept that the court may not wish to hear the evidence presented in this way, but this does not preclude it from being present on the case-file. Neither is the scientist precluded from drawing the courts attention to the preferred method before presenting evidence in line with the court requirements. The court may be unaware of the method if the scientist does not attempt to introduce it. In the O.J. Simpson case [12], the prosecution wished to use LRs, the defence advocated use of RMNE but the final result was that the court heard both methods—the judge finally ruled that the LR method was preferable.

- *Recommendation* 2: Even if the legal system does not implicitly appear to support the use of the likelihood ratio, it is recommended that the scientist is trained in the methodology and routinely uses it in case notes, advising the court in the preferred method before reporting the evidence in line with the court requirements. The scientific community has a responsibility to support improvement of standards of scientific reasoning in the court-room.

4. The likelihood ratio method using the unrestricted combinatorial approach (not taking account of peak height/areas)

This method examines all possible sets of genotypes consistent with the alternative hypotheses of H_p and H_d [13,14]. We assume uniform assumptions (such as number of contributors) across loci. For example, suppose we have four alleles *a*, *b*, *c* and *d* at a locus. If we assume that there are two contributors, then an exhaustive list of all of the possible genotype combinations is given in Table 1. The probabilities are calculated for each combination, e.g. in the first row the probability of genotype *ab* (assuming Hardy–Weinberg equilibrium) is assigned as $2p_ap_b$ and the probability of *cd* is $2p_cp_d$. Multiplying the two together to calculate the probability of *ab* and *cd* gives $4p_ap_bp_cp_d$. This is repeated for each row, then all of the probabilities are summed together to give $\Pr(E|H_d) = 24p_ap_bp_cp_d$.

$\Pr(E|H_p)$ is calculated separately. If the suspect (*S*) is *ab*, the unknown individual (*U*) must be *cd*, then $\Pr(E|H_p) = 2p_cp_d$, hence:

$$\mathrm{LR} = \frac{2\,p_c\,p_d}{24\,p_a\,p_b\,p_c\,p_d} = \frac{1}{12\,p_a\,p_b}$$

The evaluation of two- or three-banded loci is more complex but follows the same rationale [13,14].

- *Recommendation* 3: The methods to calculate likelihood ratios of mixtures (not considering peak area) described by Evett et al. [13] and Weir et al. [14] are recommended.

5. The likelihood ratio method using the restricted combinatorial approach (taking account of peak height/areas)

A typical mixture may consist of major/minor components (Fig. 1). Provided that there is sufficient difference in peak height between the two pairs of alleles and the major components are sufficiently represented so that stochastic effects leading to substantial heterozygous imbalance can be discounted, then they may be separated according to size. Hence in the example above, it may be appropriate to designate *ab* major and *cd* minor components if the profile is derived from a two person mixture.

Interpretation is easiest if the genotype of interest (attributed to the suspect under H_p) corresponds to the major alleles *ab* of the mixture. If the genotype of interest is the minor component

[1] For a two-allele locus with allele frequencies p_a and $p_b = 1 - p_a$ the probability of *ab* under H_p that the contributors were an *aa* suspect and one unknown person is $1 - p_a^2$. The probability of *ab* under H_d that the contributors were two unknown people is $1 - p_a^4 - (1 - p_a)^4$. The LR is less than one when p_a is greater than about 0.4. The RMNE probability is 1 since no-one is excluded from the mixture. For a four-allele locus with allele frequencies p_a, p_b, p_c, p_d, suspect *ab* and crime profile evidence *abcd*, then the LR for H_p: suspect and one unknown versus H_d: two unknowns is $1/(12p_ap_b)$. This is less than one when *ab* is a common genotype, whereas the RMNE probability is one since no-one is excluded from the mixture. The probability of the DNA profile evidence increases with the number of contributors in this case.

Table 1
Evaluation of $\Pr(E|H_d)$; two person mixture with four discrete alleles present

Individual 1	Individual 2	Genotype probability
ab	*cd*	$4p_ap_bp_cp_d$
ac	*bd*	$4p_ap_bp_cp_d$
ad	*bc*	$4p_ap_bp_cp_d$
cd	*ab*	$4p_ap_bp_cp_d$
bd	*ac*	$4p_ap_bp_cp_d$
bc	*ad*	$4p_ap_bp_cp_d$
Sum		$24p_ap_bp_cp_d$

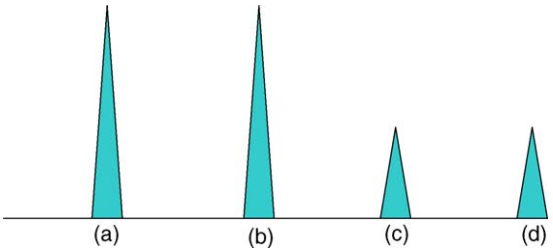

Fig. 1. A four-allele mixture, showing major *ab* and minor *cd* contributions.

cd, then interpretation is more complex since other considerations include drop-out, stutter and masking by major alleles. If the mixture is composed from two contributors, and the suspect is *ac* (i.e. one major and one minor allele), and the unknown contributor is *bd*, this combination would be accepted using the unrestricted combinatorial approach, but would be rejected, or excluded using the restricted approach under H_p. This means that defaulting to the unrestricted combinatorial approach is not necessarily conservative. If the profile is a mixture and the contributors are roughly 50:50 then the restricted approach converges to the unrestricted approach at all four peak loci and approaches it at the other loci. This convergence is most marked if the crime profile is low level as more combinations must be allowed under H_d.

A good understanding of the characteristics of H_b (heterozygote balance) and M_x (the mixture proportion) are needed to properly implement either approach [9,15–18].

5.1. An example of masking—three alleles at a locus

If the genotypes of two persons are *ab* and *bc*, then they share the *b* allele. The contributions are assumed to be additive. Given a mixture ratio of 2:1 as an example, we expect the proportions of $a{:}b{:}c = 2{:}3{:}1$ (Fig. 2). The mixture ratio is approximately the same across loci.

The profile is no longer balanced and consequently the interpretation is more difficult but more informative. The major component (*ab*) can be identified. The minor component is *bc*. Other combinations might be considered reasonable, such as *bb*, *ac*. The principle followed is to assess the combinations that would be expected to give a reasonable fit to the peak areas, eliminating those that are unreasonable. To do this it is necessary to make an assessment in relation to the heterozygote balance (H_b) and mixture proportion (M_x) [9,15–17].

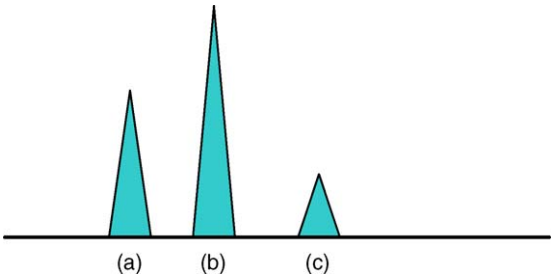

Fig. 2. A three-allele profile showing masking.

5.2. The restricted combinatorial (binary) model

The restricted combinatorial (binary) model [9,16] starts from the position that all alternatives (Table 1) are considered possible unless the combination gives a poor fit to the peak height/areas. For example, in Fig. 1, the combination of a minor allele <60% the peak height/area of the major allele when there is reasonable quantity of DNA analysed (at least 500 pg) is unrealistic given experimental data on heterozygote balance (H_b) [19]. Consequently, the peak height/areas are unlikely given a combination such as (*ac*, *bd*), hence $\Pr(E|ac, bd) \approx 0$. All of the alternatives that give low probabilities for the areas are discounted based on an assessment of whether the genotype combinations are explicable in relation to mixture proportion (M_x) and heterozygote balance (H_b). This assessment is easiest when the loci are four-banded, but can also be carried out when there is masking of alleles, i.e. three- and two-allele mixtures where there are two contributors [20]. The implementation of such an approach in routine casework, in particular when using a computer-based expert system for mixture interpretation, requires an extensive validation of the variable parameters such as H_b and M_x, as well as appropriate guidelines for all laboratory procedures.

Clayton and Buckleton [9] assess the limitations of the restricted combinatorial (binary) model. The method is robust provided that the H_p propositions give a reasonable fit to the peak heights/areas. From the example above, if the suspect was *ac* then this would not give a good fit to the data. Both numerator and denominator need to be separately assessed and this is linked to the formulation of propositions and the number of contributors (Appendix C).

5.3. The steps to interpret a mixture

These guidelines are modified from Clayton et al. [17]. They are widely used and are summarised here as a way to interpret mixture profiles.

5.3.1. Step 1: Identify the presence of a mixture

If more than two allelic bands per locus are present, a mixture may be inferred. Note extra bands may also be present because of somatic/genetic polymorphism and stutters. In addition, allele asymmetry occurs because shared alleles result in ‘masking’. The profile appears unbalanced as a result.

5.3.2. Step 2: Designation of allelic peaks

(1) Alleles should be within ±0.5 bp of the designated control allele ladder marker.
(2) The band shift for each allele, relative to the control allelic ladder marker, should be approximately constant.

5.3.3. Step 3: Identify the number of contributors in the mixture

The number of alleles observed per locus, circumstances of the case, and the possibility of related contributors go into deciding how many contributors to condition on.

When all loci of the crime stain profile (from a cosmopolitan population) are taken into consideration to calculate the LR, often, but not always, the probability of the evidence under H_p and H_d is maximised when the number of contributors is minimised. This applies to STR multiplexes in current use but cannot be applied to SNPs.[2]

5.3.4. Step 4: Estimation of the mixture proportion or ratio of the individuals contributing to the mixture

At this stage, it may be possible to separate major/minor contributors to the mixture. If DNA templates are mixed, then the ratio/proportion of contributors are approximately preserved throughout the mixture at each locus. The mixture proportion (M_x) or ratio (M_r) can be approximately assessed [16,20]. For example, the approximate value of M_x for a four-banded profile conditioned on two contributors, where two minor component alleles *a* and *b* are present with two major component alleles *c* and *d* is:

$$M_x = \frac{\phi_a + \phi_b}{\phi_a + \phi_b + \phi_c + \phi_d}$$

where ϕ_i is the peak height or peak area of the *i*th allele.

More robust methods have been developed that calculate a single $\hat{M}_x$ across all loci by calculating least squares residuals [20]. Experimentation has shown that the error in the estimation of $\hat{M}_x$ is within ±0.35 [9]. Note that the variance of this parameter may differ between processes, e.g. when different STR multiplexes, DNA amounts, and PCR conditions are used—it is given here as an example only.

[2] Other things being equal, the aim of the defense is to maximize the probability of the evidence under H_d. Similarly, the prosecution aims is to maximize the probability of the evidence under H_p, consistent with their theory of the case. The number of contributors is a secondary consideration; usually, but not always, this coincides with the fewest number of contributors required to explain the crime stain profile. It does not assist the defense case to postulate more contributors than necessary, if it reduces $\Pr(H_d)$—but exceptions are possible: consider a crime stain profile $E = a, b, c, d$; for simplicity we assume that the allele frequencies are the same (p_x). The probability given two individuals ($n_c = 2$) under H_d: two unknown individuals is $24x^4$ whereas for three individuals this probability equals $1560x^6$. The latter ($n_c = 3$) is larger than the former ($n_c = 2$) when $p_x > 0.124$. Whereas it is easy to show an exception to the generalisation at a single locus, when it does occur: (a) the impact on the LR of very common alleles on a single locus, is minimal (b) it is unlikely to have any impact when all other loci in the crime stain profile are taken into consideration since much rarer alleles will be prevalent in STR multiplexes in standard use. The overall effect will be to maximize $\Pr(H_d)$ concurrent with minimizing the number of contributors.

The second parameter under consideration is heterozygote balance (H_b)

$$H_b = \frac{\phi_a}{\phi_b} \quad (\text{where } \phi_a \text{ is the smallest peak in height or area}).$$

Experimental observation showed that under conditions where the DNA was undegraded and present in quantities >500 pg, $H_b > 0.6$ [19], hence a genotype where $H_b < 0.6$ would not be supported (we denote the threshold as $H_{b\ min} = 0.6$). Note that for low levels of DNA, stochastic effects reduce the $H_{b\ min}$ threshold. Degradation disproportionately affects high molecular weight alleles more than low molecular weight alleles, this can have a substantial effect in reducing H_b when alleles differ greatly in molecular weight (such as the HUMFIBRA/FGA locus).

5.3.5. Step 5: Consideration of all possible genotype combinations

The next step is to consider all combinations of the unrestricted combinatorial list of genotypes (Table 1) in relation to the mixture proportion (M_x) and the heterozygote balance (H_b) across all loci and their verified experimental tolerances [9]. Taking the example in Fig. 1 where there are two major alleles *ab* and two minor alleles *cd*: if the estimated $\hat{M}_x = 0.7 \pm 0.35$ across loci and $H_{b\ min} = 0.6$, a mixture can be assessed by considering each of the possible genotype combinations, per locus, with respect to these two parameters (Table 2).

Those combinations that are not supported by guidelines formulated around these two parameters are considered to have a low posterior probability and are removed. The final list of genotypes comprises those allelic combinations that are well supported by experimental observations. For example, to explain the combination *ac*, *bd*, this would require a low heterozygous balance that has not been observed in experimental data. In Table 2, only *ab*, as the major contributor, and *cd*, as the minor contributor, are feasible combinations.

These guidelines are not 'all or nothing'. If a genotype combination is borderline or uncertain, then it should be included under H_d since this will increase $\Pr(E|H_d)$, but inclusion of a borderline result is problematic under H_p because the restricted combinatorial (binary) model assumes that conditional genotypes are reasonable fits to the peak height/

Table 2
Assessment of major (*ab*)/minor (*cd*) genotypes of a mixture of two contributors relative to $\hat{M}_x$ and H_b calculated using $\phi_a = 1200$ rfu, $\phi_b = 100$ rfu, $\phi_c = 400$ rfu, $\phi_d = 380$ rfu, where rfu is relative fluorescence units (allele peak height)

Genotypes		M_x|major, minor genotypes	Heterozygous balance		Comment
Major	Minor		$H_{b\ major}$	$H_{b\ minor}$	
ab	*cd*	0.70	0.9	0.9	Passes H_b, $\hat{M}_x$
ac	*bd*	0.53	0.3	0.3	Fails H_b
ad	*bc*	0.51	0.3	0.3	Fails H_b
cd	*ab*	0.30	0.9	0.9	Fails $\hat{M}_x$
bd	*ac*	0.48	0.3	0.3	Fails H_b
bc	*ad*	0.49	0.3	0.3	Fails H_b

areas under this hypothesis. In this extreme example, if the suspect is *ac* and the unknown genotype is *bd* then the H_p propositions are unreasonable.

5.3.6. Step 6: Compare reference samples

It is important that steps 1–5 take place without considering the reference samples. This is because we demonstrably avoid the possibility of bias. If the genotype of a suspect matches a well-supported combination in the list, then there is evidence to suggest that the individual has contributed to the mixture. When the comparisons of the crime profiles and the reference samples are made, it may be necessary to refine the propositions [21]. For example, if the initial propositions suggest H_p: the stain contains the DNA of the suspect (S) and two victims (V_1, V_2), and comparison of the profile with reference samples suggests H_p: the suspect (S), one of the victims (V_1) and one unknown (U), then additional propositions may be considered.

The calculation of the likelihood ratio is exactly the same as described above (Table 1) except that in the summation of probabilities, only those that are well supported are included under H_p and H_d.

Irrespective of the principles outlined in step 3, where conditioning on the minimum number of contributors, maximises $\Pr(E|H_p)$ and $\Pr(E|H_d)$ it may still be necessary to consider multiple propositions at the final stage of analysis. It will be for the court to decide those that are relevant for consideration, bearing in mind that perhaps several different LR calculations are relevant.

- *Recommendation* 4: If peak height or area information is used to eliminate various genotypes from the unrestricted combinatorial method, this can be carried out by following a sequence of guidelines based on Clayton et al. [17].
- *Recommendation* 5: The probability of the evidence under H_p is the province of the prosecution and the probability of the evidence under H_d is the province of the defence. The prosecution and defence both seek to maximise their respective probabilities of the evidence profile. To do this both H_p and H_d require propositions. There is no reason why multiple pairs of propositions may not be evaluated (Appendix C).

6. Treatment of stutter

The characteristics of stutter bands (one tandem repeat less than the parent allele) have been evaluated in relation to the size of the associated parent allele [22,23]. The stutter peak area or height is measured as a proportion (St_p) of the parent allele peak area or height.

$$St_p = \frac{\phi_{stutter}}{\phi_{allele}}$$

In general $St_p < 0.15$.

Suppose there are minor alleles *ab* and two major alleles *cd* where *b* is in a stutter position and is within the range of experimental observations of St_p (Fig. 3). It is not known if the band in the stutter position is an allele, a stutter, or a mixture of both. The genotypes of the minor contributor to consider are *ab* (if *b* is not a stutter, or an allele with a stutter) and *ac*, *ad* and *aa* (if *b* is a stutter). If the suspect is *ab* and the victim is *cd*, then calculation of the LR is conservative if genotype combinations include bands in stutter positions under H_d. However, if the suspect is *aa* and the victim is *cd* such that the explanation under H_p is conditional upon *b* being a stutter, then the probability of stutter must be considered in the numerator. Further advice and examples are given in Appendix A.

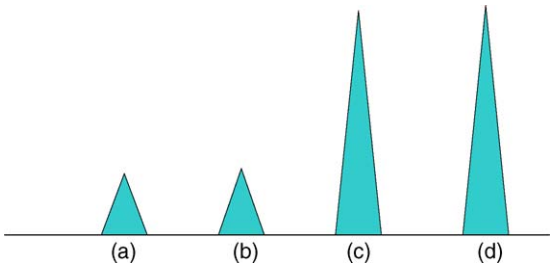

Fig. 3. Two minor alleles, *a* and *b*, with two major alleles, *c* and *d*, where allele *b* is in a stutter position.

- *Recommendation* 6: If the crime profile is a major/minor mixture, where minor alleles are the same size (height or area) as stutters of major alleles, then stutters and minor alleles are indistinguishable. Under these circumstances alleles in stutter positions that do not support H_p should be included in the assessment.

7. Drop-out

The consideration of drop-out is analogous to stutter. Suppose an allele *a* is present in a mixture at close to background level, indicating a contributor who made a small contribution. There is a substantial probability that *a*'s partner allele has dropped out completely. This has implications for an *ab* suspect when *b* is not seen. It may be net evidence against the suspect of strength approximately $1/2p_a$. But as the intensity of the *a* allele increases, the probability of drop-out $p(D)$ continually decreases until the point at which the $p(D)$ is zero and the suspect is excluded and the LR at the locus is zero [7]. Consequently, for slightly lesser *a* intensities, the net evidential value of the locus must be in favour of the suspect, i.e. LR is less than one. Therefore, it would be prejudicial to calculate a likelihood ratio of one or greater or to omit the locus because that amounts to taking LR = 1. If the hypothesised genotype is *ab* and the crime stain profile includes *a* but not *b*, then drop-out is very plausible if allele *a* is close to the background level. If allele *a* is significant in size (i.e. at a level where drop-out would not be expected), then the probability of drop-out is less likely, i.e. the possibility that the source is *aa* is more likely. See Appendix B for further considerations.

A point is reached where the background noise of the electropherogram is equivalent to the signal strength of the DNA profile. The negative controls will be particularly useful to ascertain this level. A biostatistical interpretation of an evidential

profile that is dominated by background noise is inadvisable—in the case of a major/minor mixture, only the contribution by the low level minor contributor(s) is compromised, while the major contributor is unaffected and the interpretation of this component of the mixture is not compromised.

- *Recommendation* 7: If drop-out of an allele is required to explain the evidence under H_p: ($S = ab$; $E = a$), then the allele should be small enough (height/area) to justify this. Conversely, if a full crime stain profile is obtained where alleles are well above the background level, and the probability of drop-out approaches $\Pr(D) \approx 0$, then H_p is not supported.
- *Recommendation* 8: If the alleles of certain loci in the DNA profile are at a level that is dominated by background noise, then a biostatistical interpretation for these alleles should not be attempted.

8. Low copy number

The operational definition of low copy number PCR is the manifestation of stochastic effects leading to allelic imbalance, drop-out and increased prevalence of laboratory-based contamination. Consequently, the conventional rules of heterozygous balance and other characteristics of DNA profiling do not apply [6] in the same way.

Low copy number is usually associated with a low amount of DNA (less than 200 pg). The method is typically associated with an elevated PCR cycle number, but it is important to realise that the effects may occur at 28 PCR cycles, typically with a major/minor mixture where the minor component alleles are subject to drop-out and may be the same size as stutter alleles. There are a number of caveats associated with LCN reporting [24]. LCN alleles are not necessarily category A (unambiguous). Therefore, LCN mixture analysis will have to allow for stochastic events (drop-out, heterozygous imbalance and contamination) [6].

- *Recommendation* 9: In relation to low copy number, stochastic effects limit the usefulness of heterozygous balance and mixture proportion estimates. In addition, allelic drop-out and allelic drop-in (contamination) should be taken into consideration of any assessment.

9. Definition of contamination

DNA introduced after the crime has happened and from a source that is unrelated to the crime scene: for example, the investigating officer, laboratory technicians, laboratory plasticware [25,26].

10. Training

We recognise that scientists should be trained to a level appropriate to carry out the necessary calculations. Training schedules are required for accreditation under standards such as ISO17025. There is clearly a need for comprehensive training schedules to become widely available.

11. Future

A future approach would elaborate the combinatorial approaches by taking into account all aspects including stutter, contamination and other artefacts, allelic drop-out, such as by using a probabilistic weighting for each possible genotype rather than just using a weighting of zero or one, as is inherent in the restricted combinatorial (binary) approach.

12. Accreditation and expert systems

We note that accrediting standards such as ISO17025 require traceability, which may be interpreted as excluding ''black boxes''. This is a consideration in using expert system computer programs.

Acknowledgement

The authors are grateful to James Curran for clarifying a number of issues in this paper.

Appendix A

A.1. Stutters

The interpretation of allelic components of the minor component of a mixture can be compromised:

Stutters (from a major contributor) may be the same height/peak area as the minor contributor to the mixture. This means (Fig. 4) that those bands in stutter positions may be allele only, allele plus stutter, or stutter only. In Fig. 4, bands *a*, *b* are minor alleles that are very similar in height/area. Band *b* is in a stutter position and we must assume that it could be from an unknown contributor under H_d. Consequently, if we condition on the number of contributors = 2, then the possible minor contributor genotypes are *aa*, *ac*, *ad* (where *b* is a stutter), or *ab* (where *b* is an allele either with or without a stutter).

Taking a simple scenario H_p: the stain contains the DNA of the suspect and the victim versus H_d: the stain contains the DNA of the victim and an unknown individual. If the genotype of $V = cd$, then under H_d, the possible genotypes for U include

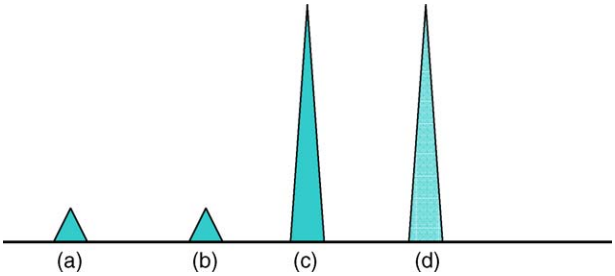

Fig. 4. *c* and *d* are unambiguous alleles, *b* is a minor allele in a stutter position and *a* is an unambiguous minor allele.

ab, since the stutter *b* is a possible allele, and $\Pr(E|H_d) = 2p_a p_b + 2p_a p_c + 2p_a p_d + p_a^2$, where p_i is the allele probability for the *i*th allele. If $S = ab$, then the LR is computed conservatively by including *ab* among the choices for *U* in the denominator, whereas if $S = aa$ or *ac* (i.e. does not have an allele in a stutter position) then it may not be conservative to include *ab* among the choices for *U* [5]. This is because $\Pr(E|H_p)$ has the probability of stutter Pr(St) as a factor, i.e. the numerator is less than one. Under H_d, we multiply by Pr(St) the combinations that can be explained if a stutter has occurred. If stutter has not occurred, then the only possibility is *ab* but we must multiply by the probability $\Pr(\overline{\text{St}})$ that stutter has not occurred where $\Pr(\overline{\text{St}}) = 1 - \Pr(\text{St})$. The formula is now:

$$\text{LR} = \frac{\Pr(\text{St})}{[p_a^2 + 2p_a p_b + 2p_a p_c + 2p_a p_d]\Pr(\text{St}) + [2p_a p_b]\Pr(\overline{\text{St}})}$$

$$\text{LR} = \frac{\Pr(\text{St})}{[p_a^2 + 2p_a p_c + 2p_a p_d]\Pr(\text{St}) + [2p_a p_b]}$$

$$\text{LR} = \frac{1}{p_a^2 + 2p_a p_c + 2p_a p_d + (2p_a p_b / \Pr(\text{St}))}$$

Which obviously approaches zero monotonically as Pr(St) approaches zero (Fig. 5).

The probability Pr(St) can be determined experimentally from a known population of samples where the proportion ϕ_{St}/ϕ_a is calculated; ϕ_{St} is the peak area/height of a stutter and ϕ_a is the peak area/height of an allele.

If ϕ is either peak area or height (it does not matter which so long as we are consistent throughout), then we can calculate the probability from data of experimental observations of probability of observing a stutter of a given proportion conditioned on the size of the 'parent' allele.

It is possible to generalise that stutters are rarely observed when $\phi_{St}/\phi_a > 0.15$ [22,23]. This means that when the allele in the stutter position is larger than this, $\Pr(\text{St}) \approx 0$.

To summarise, if the suspect is *aa*, and there is an allele *b* present, which is in a stutter position, and allele *b* is too large to be only a stutter, then $\Pr(\text{St}) \approx 0$ (from experimental observations). This means that the LR is close to zero and the H_p proposition is unsupported.

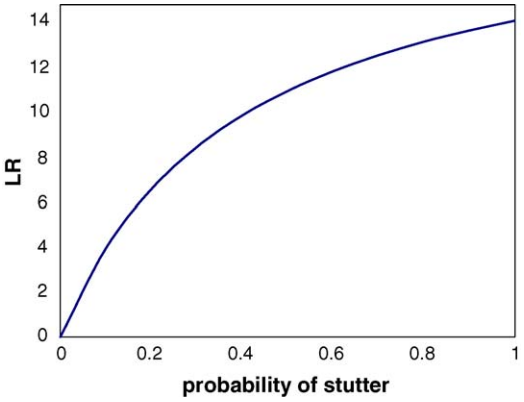

Fig. 5. Plot of Pr(St) vs. $1/(p_a^2 + 2p_a p_c + 2p_a p_d + (2p_a p_b/\Pr(\text{St})))$, where $p_a = p_b = p_c = p_d = 0.1$. The suspect is a minor contributor *aa*, the victim is (major) *cd* and allele *b* is present at the stutter position.

If ϕ is either peak area or height (it does not matter which so long as we are consistent throughout), then we can calculate the probability from data of experimental observations of probability of observing a stutter of a given proportion conditioned on the size of the 'parent' allele.

Appendix B

B.1. Further considerations of drop-out

Allele drop-out is an important consideration whenever a homozygote is observed in a DNA profile. Is the genotype of the contributor homozygous, or is it heterozygous and an allele has dropped out, giving a 'false' homozygote? Many laboratories have carried out experimentation to determine a threshold, T_{rfu} (either peak height or peak area) to signify the upper limit where allele drop-out has been observed in a heterozygote (Fig. 6). Provided that $\phi_a > T_{rfu}$ (ϕ_a is the peak height/area of allele *a*) then the probability of drop-out $\Pr(D) \approx 0$. If a homozygote is observed where $\phi_a < T_{rfu}$ then $\Pr(D) < 1$. Furthermore, the smaller ϕ_a then the greater Pr(*D*) becomes (Fig. 6).

If the suspect (*S*) is *ab* and the crime stain profile is *a*, then under H_p we must consider the probability of drop-out Pr(*D*). If the $\Pr(D) \approx 0$, as $\phi_a > T_{rfu}$, then the proposition that the suspect is a donor is not supported and an exclusion is likely to be the best conclusion. If $\Pr(D) < 1$, then the term Pr(*D*) must appear in the numerator of the likelihood ratio:

$$\text{LR} \approx \frac{\Pr(D)}{p_a(p_a + 2\Pr(D)(1 - p_a))} \quad \text{from [7]} \tag{1}$$

The correct formulae have been described for non-mixtures [6,7], but their complexity has led to the use of approximations; an example is the '*F*' designation which represents the situation where an allele may have failed to amplify. In such a circumstance the genotype may be signified by *aF* which describes a genotype containing the *a* allele and any other allele. It is customary to assign the probability of the profile as $2\Pr(a)$. This is often termed the $2p$ rule.

However, this formula may overestimate the strength of the evidence. An example where the '*F*' designation is not conservative, for non-mixtures, occurs whenever Pr(*D*) appears in the numerator (as above), i.e. the suspect is *ab*, the stain is *a* and $\Pr(D) < 0.5$ (excluding sub-population corrections).

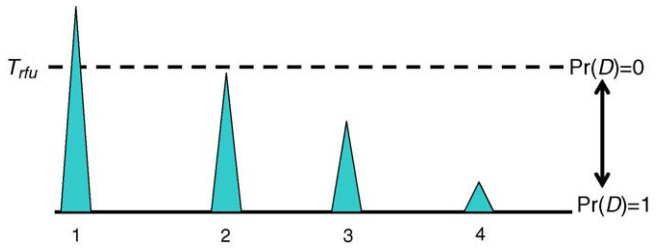

Fig. 6. Alleles 1–4 are phenotypic homozygotes of decreasing size. The probability of drop-out Pr(*D*) increases as the size of the surviving peak decreases.

Table 3
Evaluation of the crime stain profile $E = acd$, $S = ab$ and $U = cd$

Hypothesis	M_j	$\Pr(M_j)$	$\Pr(E\|M_j)$	Comments
H_p	cd	$2p_c p_d$	$p(D)\,p(\bar{D})^3$	One drop-out event
H_{d1}	Any combination that carries $acdQ$	$24p_a p_c p_d p_Q$	$p(D)\,p(\bar{D})^3$	One drop-out event (with Q allele)
H_{d2}	Any combination that carries acd	$12\ p_a p_c p_d(p_a + p_c + p_d)$	$p(\bar{D})^4$	No drop-out event

M_j is a "genotype" or a collection of ordered alleles representing a genetic combination we might wish to consider as having gone into the crime scene stain.

If it is not necessary to invoke drop-out to explain the evidence—if the suspect is a donor under H_p, then the F designation is always conservative (unless F_{St} and $\Pr(D)$ are high).

$$\mathrm{LR} \approx \frac{\Pr(D)}{p_a(p_a + 2\Pr(D)(1 - p_a))} \geq \frac{1}{2p_a} \quad \text{from [7]} \tag{2}$$

Expansion of these concepts to mixtures is complex and this is the reason why they are not generally used. Programmed solutions have recently appeared however that use a modified (improved) concept instead of 'F' [27]. This is called the 'Q' virtual allele concept: if there are n alleles visible in a mixture and drop-out has occurred, we can calculate a ceiling for the frequency of any missing allele:

$$\Pr(Q) = 1 - \sum_{i}^{k_p} p_i \tag{3}$$

where k_p is the number of alleles present at the locus in the crime stain and p_i is the population frequency of the ith allele.

We include below a summary of a further evaluation using the 'F' designation compared to the model incorporating $\Pr(D)$ for a number of scenarios for a simple mixture:

$$\mathrm{LR} = \frac{\Pr(E|S + U)}{\Pr(E|U_1 + U_2)} \tag{4}$$

where S is the suspect, U the unknown and E is the crime stain profile DNA evidence. No sub-population correction is made in this example. We make the simplified assumption that $\Pr(D)$ is the same for S and U.

B.2. Example 1

We assume that the probability of drop-out is the same for all alleles. The crime stain profile $E = acd$, $S = ab$ and $U = cd$. This means that under H_p, allele b has dropped out. To calculate H_d we consider separately the conditions of drop-out and no drop-out. Under H_{d1}, drop-out is invoked. We simultaneously incorporate the virtual allele Q to describe all pairwise combinations (M_j) from alleles a, c, d, Q. Alternatively, under H_{d2} drop-out is not invoked, in which case combinations (M_j) from the visible alleles a, c, d are evaluated. Summing H_{d1} and H_{d2} gives the denominator of the LR (Table 3).

$$\mathrm{LR} = \frac{p(D)}{6p_a\{2p(D)p_Q + p(\overline{D})(p_a + p_c + p_d)\}} \tag{5}$$

Using Eq. (5), if $p_a = p_c = p_d = 0.1$ then the resultant LRs are shown in Table 4.

The evidence favours H_d, unless $p(D) > 0.6$, when it is neutral. If the 'F' designation is used, the numerator = 1, then:

$\mathrm{LR} = (1/12p_a) = 0.83$ which corresponds to $p(D) \approx 0.3$. Note that if $p(D)$ is smaller, this has a relatively minor effect, e.g. LR = 0.41 when $p(D) = 0.1$.

If $p_a = p_c = p_d = 0.02$, then the resultant LRs are shown in Table 5.

The biological model ('F' designation) returns LR = 4.17, consistent with $\Pr(D) \approx 0.3$. The LR is relatively insensitive to $\Pr(D)$ in this example.

Conclusion: The 'F' designation is conservative provided $\Pr(D) > 0.3$ (approximately).

B.3. Example 2

As usual, we assume that the probability of drop-out is the same for all alleles. Consider a low level profile $E = abd$, $S = ab$ and $U = d$. Because the profile is low level, it is possible that allele drop-out has occurred, although both alleles pertaining to S are observed. Under H_d, we should assume that an allele may have dropped out. In such a case we should

Table 4
LRs generated from Eq. (6) where $p_a = p_c = p_d = 0.1$

$\Pr(D)$	$\Pr(E\|H_p)$	$\Pr(E\|H_d)$	LR
0.1	0.1000	0.246	0.41
0.2	0.2000	0.312	0.64
0.3	0.3000	0.378	0.79
0.4	0.4000	0.444	0.90
0.5	0.5000	0.510	0.98
0.6	0.6000	0.576	1.04
0.7	0.7000	0.642	1.09
0.8	0.8000	0.708	1.13
0.9	0.9000	0.774	1.16

Table 5
LRs generated from Eq. (6) where $p_a = p_c = p_d = 0.02$

$\Pr(D)$	$\Pr(E\|H_p)$	$\Pr(E\|H_d)$	LR
0.1	0.1000	0.029	3.44
0.2	0.2000	0.051	3.93
0.3	0.3000	0.073	4.13
0.4	0.4000	0.095	4.23
0.5	0.5000	0.116	4.30
0.6	0.6000	0.138	4.34
0.7	0.7000	0.160	4.37
0.8	0.8000	0.182	4.40
0.9	0.9000	0.204	4.42

Table 6
Evaluation of a low level profile where $E = abd$, $S = ab$ and $U = d$

Hypothesis	M_j	$\Pr(M_j)$	$\Pr(E\|M_j)$	Comments
H_p	ad, bd or dd	$2p_a p_d + 2p_b p_d + p_d^2$	$\Pr(\bar{D})^4$	No drop-out event
H_{d1}	Any combination that carries $abdQ$	$24p_a p_b p_d p_Q$	$\Pr(D)\Pr(\bar{D})^3$	One drop-out event (with Q allele)
H_{d2}	Any combination that carries abd	$12p_a p_b p_d(p_a + p_b + p_d)$	$\Pr(\bar{D})^4$	No drop-out events

Table 7
LRs generated from Eq. (7) where $p_a = p_c = p_d = 0.1$

$\Pr(D)$	$\Pr(E\|H_p)$	$\Pr(E\|H_d)$	LR
0.1	0.4500	0.04920	9.1
0.2	0.4000	0.06240	6.4
0.3	0.3500	0.07560	4.6
0.4	0.3000	0.08880	3.4
0.5	0.2500	0.10200	2.5
0.6	0.2000	0.11520	1.7
0.7	0.1500	0.12840	1.2
0.8	0.1000	0.14160	0.7
0.9	0.0500	0.15480	0.3

Table 8
LRs generated from Eq. (7) where $p_a = p_c = p_d = 0.02$

$\Pr(D)$	$\Pr(E\|H_p)$	$\Pr(E\|H_d)$	LR
0.1	0.0900	0.00116	77.5
0.2	0.0800	0.00204	39.3
0.3	0.0700	0.00291	24.1
0.4	0.0600	0.00378	15.9
0.5	0.0500	0.00466	10.7
0.6	0.0400	0.00553	7.2
0.7	0.0300	0.00640	4.7
0.8	0.0200	0.00728	2.7
0.9	0.0100	0.00815	1.2

invoke Q, where Q is any allele other than a, b, d. Under H_p, it is not necessary to invoke Q to explain S, hence the simplest explanation of U that maximises $\Pr(E|H_p)$ is either ad, bd or dd. Under H_d, $\Pr(E|H_d)$ is the same as in the previous example, hence the LR is calculated (Table 6):

$$\mathrm{LR} = \frac{\Pr(\overline{D})(2p_a + 2p_b + p_d)}{12p_a p_b\{2\Pr(D)p_Q + \Pr(\overline{D})(p_a + p_b + p_d)\}}. \quad (6)$$

When $p_a = p_b = p_d = 0.1$, then the resultant LRs are shown in Table 7.

Note that under H_p, drop-out is not invoked. Under H_d, there are two scenarios—one assumes drop-out, whereas the other does not. The LR is greatest when $p(D)$ is low. If $p(D)$ is high, then the LR is low since it is more likely that two bands will survive.

Invoking the 'F' designation produces:

$$\mathrm{LR} = \frac{2p_a + 2p_b + p_d}{12p_a p_b(2\Pr(F) + p_a + p_b + p_d)} = 1.8 \quad (7)$$

The LR corresponds approximately to $p(D) \approx 0.6$.

We now calculate (Table 8) using a rare allele probability ($p_a = p_b = p_d = 0.02$):

The 'F' designation gives LR = 10.11, corresponding to $\Pr(D) \approx 0.5$.

Conclusion: Although both S alleles are present, it is reasonable to postulate drop-out under H_d if $\phi_a < T_{rfu}$. The 'F' designation is conservative if $\Pr(D) < 0.5$. If $\phi_d > T_{rfu}$, then there is no need to use 'F' under H_d since the best supported explanation for U is homozygote dd.

B.4. Example 3

The profile is cd and $S = ab$; both S alleles have dropped out. Under H_p, $U = cd$, but under H_{d1}, U_1 and U_2 incorporate any combination of alleles Q, c and d where Q is any allele except for c and d. In addition, H_{d2} can invoke any combination of two alleles c, d without Q. However, the probability of a two-allele model is several orders of magnitude lower than the Q model and is consequently not included in this example (Table 9).

If $p_c = p_d = 0.1$, then:

$$\mathrm{LR} \approx \frac{1}{6p_Q} = 0.21 \quad (\text{independent of } p(D)) \quad (8)$$

The LR always favours H_d, independent of $p(D)$. Substituting with the 'F' designation results in:

$$\mathrm{LR} = \frac{1}{6\Pr(F)(p_c + p_d + \Pr(F))} = 0.14 \quad (9)$$

If the scenario changes so that U has dropped out, then the numerator ≈ 1, as U could be any allelic combination. The LR is:

$$\mathrm{LR} \approx \frac{1}{12p_a p_b p_Q(p_a + p_b + p_Q)} \quad (10)$$

LR $\approx$ 10.4 (when $p_c = p_d = 0.1$).

Substituting 'F' instead of Q gives LR = 6.9.

Table 9
Evaluation of a DNA profile where $E = cd$ and $S = ab$; both S alleles have dropped out

Hypothesis	M_j	$\Pr(M_j)$	$\Pr(E\|M_j)$	Comments
H_p	cd	$2p_c p_d$	$\Pr(D)^2\Pr(\bar{D})^2$	Two drop-out events
H_{d1}	Any combination that carries cdQ	$12p_c p_d p_q(p_c + p_d + p_Q)$	$\Pr(D)^2\Pr(\bar{D})^2$	Two drop-out events (with Q alleles)

Conclusion: The evidence strongly favours H_d. The 'F' designation gives a slightly lower LR.

B.5. Generalised conclusions

A further generalisation can be made. Whenever a correction factor such as 'F' is used, the effect is to increase the probability. Therefore caution is required whenever this is used in the numerator. For example, if $S = ab$, $U = c$, $E = abc$ and $\phi_{abc} < T_{rfu}$, then drop-out may or may not have happened. Whereas it is reasonable to include 'F' in the denominator to achieve Pr_{max}, it is not necessary in the numerator. U can be conservatively assigned as genotype cc, which is always less then $Pr(cF)$.

Appendix C

C.1. The formulation of propositions

It is not always easy to specify hypotheses in complex cases where multiple perpetrators or victims may be present. The DNA result itself may indicate that different explanations are possible. Furthermore, it is possible that H_p and H_d could be very different from each other. For example, under H_p we might consider a victim and suspect to be the contributors ($V + S$), whereas under H_d we might examine more complex scenarios such as three unknowns being the contributors to the stain ($U_0 + U_1 + U_2$). There is a common misconception that the numbers of contributors under H_p and H_d should be the same. There is no requirement for this.

C.2. Formulation of H_p and H_d

In principle, H_p is the province of the prosecutor and H_d is the province of the defence. Both are constrained by what is known about the circumstances of the case. The forensic scientist usually formulates both H_p and H_d. In a typical example, H_p may propose that the DNA is a mixture of the suspect (S) and an unknown (U_1) individual. Under H_d, S is substituted by U_0. However, the defence may alter H_d (but not H_p), for example, if the number of contributors is contested. Consequently, some dialogue between the forensic scientist and defence is desirable in order to establish H_d. If this cannot be carried out pre-trial, the analyst may acknowledge in the report that the defence may offer alternative propositions which will require additional calculations.

C.3. Number of contributors

The number of contributors under H_p and H_d may be different. The most parsimonious explanations (the smallest number of unknown contributors needed to explain the evidence) are usually the ones that maximise the respective likelihoods [12]. But further research is needed to clarify, hence it may be wise explore options for different numbers of contributors.

C.4. Relevance of propositions

It follows that some propositions may be redundant if they only serve to reduce $Pr(E|H_d)$. This will be especially true in many circumstances where H_d incorporates more unknown individuals than required to maximise this probability.

References

[1] J. Buckleton, A framework for interpreting evidence, in: J. Buckleton, C.M. Triggs, S.J. Walsh (Eds.), Forensic DNA Evidence Interpretation, CRC Press, London, 2005, pp. 27–63.
[2] D.J. Balding, Weight-of-Evidence for Forensic DNA Profiles, Wiley, 2005.
[3] DNA Advisory Board, Evaluation of the frequency of occurrence of DNA profiles calculated from pertinent population databases, Forensic Sci. Commun. 2 (2000), http://www.fbi.gov/hq/lab/fsc/backissu/july2000/dnastat.htm.
[4] C. Ladd, H. Lee, N. Yang, F. Bieber, Interpretation of complex forensic DNA mixtures, Croat. Med. J. 43 (2001) 244–246.
[5] P. Gill, R. Sparkes, J.S. Buckleton, Interpretation of simple mixtures when artefacts such as a stutters are present—with special reference to multiplex STRs used by the Forensic Science Service, Forensic Sci. Int. 95 (1998) 213–224.
[6] P. Gill, J. Whitaker, C. Flaxman, N. Brown, J. Buckleton, An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA, Forensic Sci. Int. 112 (2000) 17–40.
[7] J. Buckleton, C. Triggs, Is the 2p rule always conservative? Forensic Sci. Int. (2005).
[8] National Research Council, The Evaluation of Forensic DNA Evidence, National Academy Press, Washington, DC, 1996, p. 130.
[9] T. Clayton, J. Buckleton, Mixtures, in: J. Buckleton, C. Triggs, S.J. Walsh (Eds.), Forensic DNA Evidence Interpretation, CRC Press, London, 2005, pp. 217–274.
[10] B.S. Weir, Court experiences in the USA: people v. Simpson, in: First International Conference on Forensic Human Identification in the Millenium, London, 1999.
[11] C.H. Brenner, What's Wrong With the ''Exclusion Probability?'', 1997, http://www.dna-view.com/exclusn.htm.
[12] B.S. Weir, DNA statistics in the Simpson matter, Nat. Genet. 11 (1995) 365–368.
[13] I.W. Evett, C. Buffery, G. Willott, D. Stoney, A guide to interpreting single locus profiles of DNA mixtures in forensic cases, J. Forensic Sci. Soc. 31 (1991) 41–47.
[14] B.S. Weir, C.M. Triggs, L. Starling, K.A.J. Stowell, J. Buckleton, Interpreting DNA mixtures, J. Forensic Sci. 42 (1997) 213–222.
[15] P. Gill, R. Sparkes, C. Kimpton, Development of guidelines to designate alleles using an STR multiplex system, Forensic Sci. Int. 89 (1997) 185–197.
[16] P. Gill, R. Sparkes, R. Pinchin, T. Clayton, J. Whitaker, J. Buckleton, Interpreting simple STR mixtures using allele peak areas, Forensic Sci. Int. 91 (1998) 41–53.
[17] T.M. Clayton, J.P. Whitaker, R.L. Sparkes, P. Gill, Analysis and intepretation of mixed forensic stains using DNA STR profiling, Forensic Sci. Int. 91 (1998) 55–70.
[18] M.W. Perlin, B. Szabady, Linear mixture analysis: a mathematical approach to resolving mixed DNA samples, J. Forensic Sci. 46 (2001) 1372–1378.
[19] P. Gill, J. Curran, K. Elliot, A graphical simulation model of the entire DNA process associated with the analysis of short tandem repeat loci, Nucleic Acids Res. 33 (2005) 632–643.
[20] M. Bill, P. Gill, J. Curran, T. Clayton, R. Pinchin, M. Healy, J. Buckleton, PENDULUM—a guideline based approach to the interpretation of STR mixtures, Forensic Sci. Int. 148 (2004) 181–189.
[21] I.W. Evett, G. Jackson, J.A. Lambert, More on the hierarchy of propositions: exploring the distinction between explanations and propositions, Sci. Justice 40 (2000) 3–10.

[22] C.J. Fregeau, K.L. Bowen, B. Leclair, I. Trudel, L. Bishop, R.M. Fourney, AmpFlSTR profiler Plus short tandem repeat DNA analysis of casework samples, mixture samples, and nonhuman DNA samples amplified under reduced PCR volume conditions (25 microL), J. Forensic Sci. 48 (2003) 1014–1034.
[23] J.P. Whitaker, E.A. Cotton, P. Gill, A comparison of the characteristics of profiles produced with the AMPFlSTR SGM Plus multiplex system for both standard and low copy number (LCN) STR DNA analysis, Forensic Sci. Int. 123 (2001) 215–223.
[24] P. Gill, Application of low copy number DNA profiling, Croat. Med. J. 42 (2001) 229–232.
[25] T. Howitt, Ensuring the Integrity of Results: A Continuing Challenge in Forensic DNA Analysis, 2003, http://www.promega.com/geneticidproc/ussymp14proc/oralpresentations/Howitt.pdf.
[26] P. Gill, A. Kirkham, Development of a simulation model to assess the impact of contamination in casework using STRs, J. Forensic Sci. 49 (2004) 485–491.
[27] J.M. Curran, P. Gill, M.R. Bill, Interpretation of repeat measurement DNA evidence allowing for multiple contributors and population substructure, Forensic Sci. Int. 148 (2005) 47–53.

Glossary

Allele drop-in: Contamination from a source unassociated with the crime stain manifested as one or two alleles.

Allele drop-out: Low level of DNA insufficiently amplified to give a detectable signal.

Conservative: 1. An assignment for the weight of evidence that is believed to favour the defence. 2. When the evidence is very powerful in one direction, assigning the weight as less than our belief in that direction. 3. Lack of conservativeness will often result when the assumptions that underpin a statistical model are seriously violated.

Contamination: Extraneous DNA from a source unassociated with the crime stain—e.g. plastic-ware can be contaminated at manufacturing source.

Continuous approach: The allelic intensity information is used to give a variable, probability, weight to the validity of each genotype set as an explanation, rather than merely binary weights as in the combinatorial approaches.

Exclusion: Exclusion from a stain: 1. a decision (by the expert) that a particular reference DNA profile does not represent a contributor to the stain; 2. (jargon) situation in which the reference profile is ''excluded (3)'' from the stain at one or more loci. *Exclusion at a locus*: 3. (jargon) pattern of the assumed genotypes at a locus that some allele seen in a particular reference DNA profile is not observed in a stain.

Exclusion probability: The probability that a randomly selected DNA profile would be excluded (2).

Frequency: Rate at which an event occurs. For example, *sample frequency of an allele* is the number of occurrences of the allele in a population sample, divided by the sample size; *population frequency of a DNA profile* is the (unknown) number of times that the profile occurs in the population, divided by the population size.

Likelihood: Conditional probability of an event, where the event is considered as an outcome corresponding to one of several conditions or hypotheses. An example of an event is the DNA profile evidence from a crime stain. The probability of the event is conditional upon the hypothesis that may vary. If the DNA profile is a mixture, a typical prosecution hypothesis may be suspect and victim. This is written as $\Pr(E|H)$, where E is the event, the vertical bar in between the two terms means ''given'', and H is the hypothesis.

Likelihood ratio: Ratio of two likelihoods, i.e. the ratio of two probabilities of the same event (E) under different hypotheses (H_1, H_2). Written as $\mathrm{LR} = (E|H_1)/(E/H_2)$. Typically H_1 corresponds to the prosecution hypothesis and H_2 corresponds to the defence hypothesis. If H_1 consists of suspect and victim, then the alternative H_2 is unknown and victim.

Probability: Long-term rate of occurrence of an event in a conceptually repeatable experiment. Same as *expected* frequency, the expectation evaluated over cases described by the probability condition. Or: a coherent assignment of a number between zero and one that reflects in a fair and reasonable way our belief that the event is true.

Propositions: The hypothesis of the defence or prosecution arguments that are used to formulate the likelihood ratio.

Restricted combinatorial method: Elaboration of the unrestricted method in which allelic intensity (peak height/area) information is used to restrict the sets of genotypes that are considered plausible explanations.

Stutter: An allelic artefact cause by 'slippage' of the *Taq* polymerase enzyme. It is always four bases less than the allele that causes the stutter. Stutters are always found in allelic positions and can compromise interpretation of minor contributors to mixtures.

Unrestricted combinatorial method: The simple likelihood ratio method of evaluating mixture evidence described in Weir et al. [14] and Clayton and Buckleton [9]. The method assumes a list of all alleles in the mixture, and considers competing hypotheses that various known or unknown profiles are the constituents of the mixture. It uses no information about allelic intensities, hence one set of genotypes whose allele sets are coincident with the mixture is considered to be as valid an explanation of the mixture as any other set.