Does the use of probabilistic genotyping change the way we should view sub-threshold data?

Duncan Taylor[1,2*], John Buckleton[3,4], Jo-Anne Bright[3]

[1] *Biology Department, Forensic Science South Australia, Adelaide, SA, Australia*

[2] *School of Biological Sciences, Flinders University, Adelaide SA, Australia*

[3] *Institute of Environmental Science and Research Limited, Private Bag 92021 Auckland 1142, New Zealand*

[4] *National Institute of Standards and Technology, Gaithersburg, MD, United States of America*

*Corresponding author. Email: Duncan.Taylor@sa.gov.au

Key words

DNA profile interpretation; sub-threshold; likelihood ratios; analytical threshold.

Abstract

The sensitivity and resolution of modern DNA profiling hardware is such that forensic laboratories generate more data than they have resources to analyse. One coping mechanism is to set a threshold, above the minimum required by instrument noise, so that weak peaks are screened out. In binary interpretations of forensic profiles, the impact of this threshold (sometimes called an analytical threshold) was minimal as interpretations were often limited to a clear major component. With the introduction of continuous typing systems, the interpretation of weak minor components of mixed DNA profiles is possible and consequently the consideration of peaks just above or just below the analytical threshold becomes relevant. We investigate here the occurrence of low-level DNA profile information, specifically that which falls below the analytical threshold. We investigate how it can be dealt with and the consequences of each choice in the framework of continuous DNA profile interpretation systems. Where appropriate we illustrate how these can be implemented using the probabilistic interpretation software STRmix™. We demonstrate a feature of STRmix™ that allows the analyst to guide the software using human observation that there is a low-level contributor present through user-designated prior distributions for contributor mixture proportions.

1. Introduction

The primary method for the analysis of a DNA sample is amplification by polymerase chain reaction (PCR) which incorporates a flurophore. This is followed by separation of the fragments by capillary electrophoresis. The output is a trace of fluorescence versus time that is referred to as an electropherogram (epg). Most laboratories set an analytical threshold, AT, above which peaks are labelled at analysis. The AT is often set well above the level of electronic noise. Peaks in the epg may be artefactual or allelic. Epg analysis software can recognise and filter some of the well characterised artefacts, but many still require the judgement of a human analyst. Many of these remaining artefactual peaks can be recognised by position or morphology. In binary interpretations the impact of these weak peaks was minimal as interpretations were often limited to the interpretation of a clear major component. With the introduction of continuous typing systems the interpretation of weak minor components of mixed DNA profiles is possible and consequently the consideration of peaks just above or just below analytical threshold becomes important.

There have been numerous published methods that describe how the AT could be determined. For a review the reader is referred to the work of Bregu *et al.*[1]. Some recognise that there are different factors that affect the AT, such as dye colour, input DNA amount or instrument [1, 2]. The ideal situation is that these factors are considered on a sample by sample (and even locus by locus) basis and applied to the profile [3]. However, in order to balance the laboratory's sample processing capability with interpretation needs, the laboratories may need to apply a single AT that applies to all profiles, or an AT that is based on dye label, and is set at a level designed to screen out much low level artefactual fluorescence. Thus, it is of value to address the issue of sub-AT information from a standpoint that continues to address the balance between sample processing and interpretation. As such, the purpose of this work is to examine effects of using sub-AT threshold signal on interpretation rather than investigate methods to determine the AT. This work considers that no matter where the AT is set, peaks will exist below it that appear allelic and may affect interpretation.

This work evaluates some options for analysts to deal with sub-threshold information and the risks or benefits associated with each in the context of analysis within a continuous DNA interpretation system. We introduce a novel method for dealing with sub-threshold data implemented within the STRmix™ program that allows the user to specify a prior belief in mixture proportions.

Much of the discussion will be dominated by the topic of choosing a number of contributors for analysis, which is where the sub-AT peaks will have their biggest impact on interpretations.

There have been various works that look at the consequences of overestimation or underestimation of the number of contributors [4, 5]. In general, the consequences of underestimation are that known contributors are excluded due to the forced pairing of peaks that in reality do not pair. The consequence of overestimation is more complex; doing so can have very little effect on a major contributor to a DNA profile and a more marked effect on a minor contributor. This is only true for continuous systems that take peak heights into account. For a semicontinuous system the effect of overestimation will have an effect on all contributors to a mixture as more genotype sets are considered for all contributors to the mixture (see Benschop *et al.*[6]). There is also a greater number of non-contributors that are given relatively neutral likelihood ratios (*LR*s) as the analysis is accounting for more potential dropout.

The Scientific working Group on DNA Analysis methods (SWGDAM) guidelines for the validation of probabilistic genotyping systems [7] advise a study of over and underestimation of contributor numbers (at 4.1.6.4) so that the impact of the above mentioned issues are known for the system being validated. There are methods available that do not require a number of contributors to be assigned [8], however the majority of current probabilistic software programmes do require a choice of number of contributors.

This leads to the question of how, if at all, should sub-threshold information be taken into account when making the choice of number of contributors. We consider four broad categories for consideration:

1) Ignore the presence of sub-threshold peaks when interpreting DNA profiles
2) Change the method by which data is generated (either by lower the AT or carrying out replicate PCRs)
3) Use informed priors on mixture proportion in a probabilistic system
4) Do not interpret the DNA profile

1.1 Ignore the presence of sub-threshold peaks when interpreting DNA profiles

To consider the performance and consequence of utilising sub-threshold information when carrying out an interpretation we first start by considering the scope of the issue. We do this in

two ways; firstly via a simulation designed to give an indication of how ignoring sub-threshold information will lead to an underestimate of the number of contributors in the most high-risk situations and secondly a demonstration of the practical consequences of ignoring sub-threshold data.

We first start by considering the probability that by ignoring sub-threshold information a low level two person mixture would be assigned as a single source profile. We do this by simulating two contributors with low levels of DNA and different levels of allele sharing and over various analytical thresholds. 21 locus profiles were simulated and the peak heights and AT are intended to be realistic for an Applied Biosystems 3130 capillary electrophoresis (CE) system (Thermo Fisher Scientific, CA). Details of this simulation appear in appendix 1.

Simulation was chosen in this part of the study because it allows for control over the experimental conditions and for a large number of experiments (for example, table 1 give the results of 150,000 simulated mixtures).

Table 1 gives the number of simulations (out of 1,000) of two low-level contributors that when combined collectively gave a profile that looked like a single contributor. Simple allele count per locus was used to assign the number of contributors. Use of peak heights is likely to be superior but at such low-levels this is not likely make a significant difference to the count [9].

Inspection of table 1 suggests that, under the trialled circumstances, there is a high probability of the alleles from two individuals masquerading as a low-level single source profile. The table also shows that this effect is likely to be reduced at lower AT.

This simulation informs the probability of assigning one donor if there are in fact two. It is important not to confuse this with the probability that there are two if we assign one. This latter probability is what we really want. To obtain this probability we need the prior probabilities that there are one or two contributors in a profile. We are allowed to know what type of sample it is and what analysis regime we have employed but we cannot use profile information itself. We will use equal priors for this work, accepting that this was an arbitrary choice. Making this choice will restrict the lower bound probability that a profile is single source given that it appears as single source to 0.5. Using these priors the probabilities in table 2 are obtained (details of the calculation appear in appendix 2).

Table 1. The number of simulations (out of 1000) of two low-level contributors that gave a profile that looked like a single contributor based on allele count at 21 loci.

		Average peak height of Contributor 2 (rfu)										
			20	40	60	80	100	120	140	160	180	200
Average peak height of contributor 1 (rfu)	AT = 100 rfu	20	722	734	705	642	549	436	344	230	199	179
		40	734	947	869	718	559	337	194	118	113	78
		60	705	869	746	530	302	119	52	36	17	9
		80	642	718	530	283	95	22	6	3	0	0
		100	549	559	302	95	19	4	0	0	0	0
		Average peak height of Contributor 2 (rfu)										
			10	20	30	40	50	60	70	80	90	100
	AT = 50 rfu	10	754	694	633	557	448	356	249	201	168	137
		20	694	757	520	378	239	122	71	34	33	15
		30	633	520	305	151	57	19	10	2	1	1
		40	557	378	151	70	19	4	0	0	0	1
		50	448	239	57	19	2	0	0	0	0	0
		Average peak height of Contributor 2 (rfu)										
			10	20	30	40	50	60	70	80	90	100
	AT = 30 rfu	10	709	504	315	227	117	71	57	40	30	40
		20	504	302	110	32	16	5	1	0	2	0
		30	315	110	16	1	0	0	0	0	0	0
		40	227	32	1	0	0	0	0	0	0	0
		50	117	16	0	0	0	0	0	0	0	0

Table 2. The probability that the peaks above AT are from a single source (S) given that they look like a single source on simple allele count (AS), Pr(S|AS).

Masking	0.2		0.5	
Mean peak height in range AT (rfu)	10-50 rfu	10-100 rfu	10-50 rfu	10-100 rfu
	Pr(S|AS)			
30	0.91	0.98	0.7	0.82
50	0.66	0.87	0.56	0.67
100	0.56	0.61	0.59	0.56

For the CE system that we are simulating here it is likely that peaks above 30 rfu that have passed expert inspection are all allelic. This suggests that for an AT = 100 or 50 rfu there is a possible strategy of using peaks below the threshold to help improve the assignment of the number of contributors.

These results suggest that ignoring sub-threshold peaks when interpreting low level putatively mixed DNA profiles is likely to lead to underestimation of the number of contributors and thereby has the potential to lead to incorrect interpretations. It is unlikely that a blanket rule to ignore such information would be sustainable. There may be concern that these *in silico* mixtures ignore the effect of stutters. Any stutters miss-assigned as allelic tends to increase the allele count and hence have no effect at all in the direction of underestimation.

We do however look at a number of *in vitro* mixtures. A range of four person mixtures were amplified using GlobalFiler™ (Thermo Fisher Scientific, CA), as per the manufacturer's instructions. Amplification fragments were resolved using the ABI PRISM® 3130xl Genetic Analyser and analysed in GeneMapper® ID-X to obtain peak height information for each profile. These mixtures are samples 22 to 31 from [10], amplified in triplicate except for sample 23 where there were only 2 replicates, leading to a total of 29 profiles. We reproduce the relevant mixture information from [10] in Table 3.

Table 3. Mixture proportions and PCR setup.

Tubes	mixture ratios for contributor C1:C2:C3:C4	Total DNA added to PCR (pg)
22	1:1:1:1	400
23		200
24		50
25		20
26		10
27	4:3:2:1	400
28		200
29		50
30		20
31		10

Profiles were analysed using ATs of 30 rfu, 50 rfu and 100 rfu. While it is possible to construct simpler mixtures that could be used in this experiment, we choose four person mixtures due to the high probability that the number of contributors can be underestimated, the higher probability that masking or dropout may occur and as an example of profiles where the use of sub-AT information could have an important impact on the interpretation. Later (in Table 4) we show how for the data sets used the number of contributors could be underestimated over half the time.

Profiles were analysed using STRmix™ V2.3 which utilises models described in [11-13] (exact software settings used are available from the corresponding author on request). In all analyses the Y-indel locus and DYS391 were ignored. A uniform probability for allelic drop-in of 0.0017 was used (up to 75 rfu) for the 30 rfu and 50 rfu AT and a drop-in probability of zero was used for the 100 rfu AT, in line with laboratory observations.

Two experiments were carried out to investigate the consequences of ignoring the sub-threshold information when determining number of contributors.

Experiment 1: Utilising sub-threshold information

Firstly the correct number of contributors was assigned to each profile during analysis and the *LR*s were calculated using the propositions:

H_p: The person of interest (POI) and 3 unknown individuals are the sources of DNA

H_d: 4 unknown individuals are the sources of DNA

The POI was varied to be each of the four known contributors and 186 randomly selected non-contributors. *LR*s were calculated using an in-house self-declared Caucasian GlobalFiler database and using the product rule. This amounts to 116 STRmix™ analyses compared to known donors and 5394 comparisons to non-donors.

Experiment 2: Ignoring sub-threshold information

In this experiment the number of contributors was chosen ignoring sub-threshold information i.e. based purely on the number of detected peaks above the varying AT. Using the chosen number of contributors, *N*, *LR*s were calculated using the propositions:

H_p: The POI and ($N-1$) unknown individuals are the sources of DNA

H_d: N unknown individuals are the sources of DNA

The POI was varied to be each of the four known contributors and 186 randomly selected non-contributors. *LR*s were calculated using an in-house self-declared Caucasian GlobalFiler database and using the product rule.

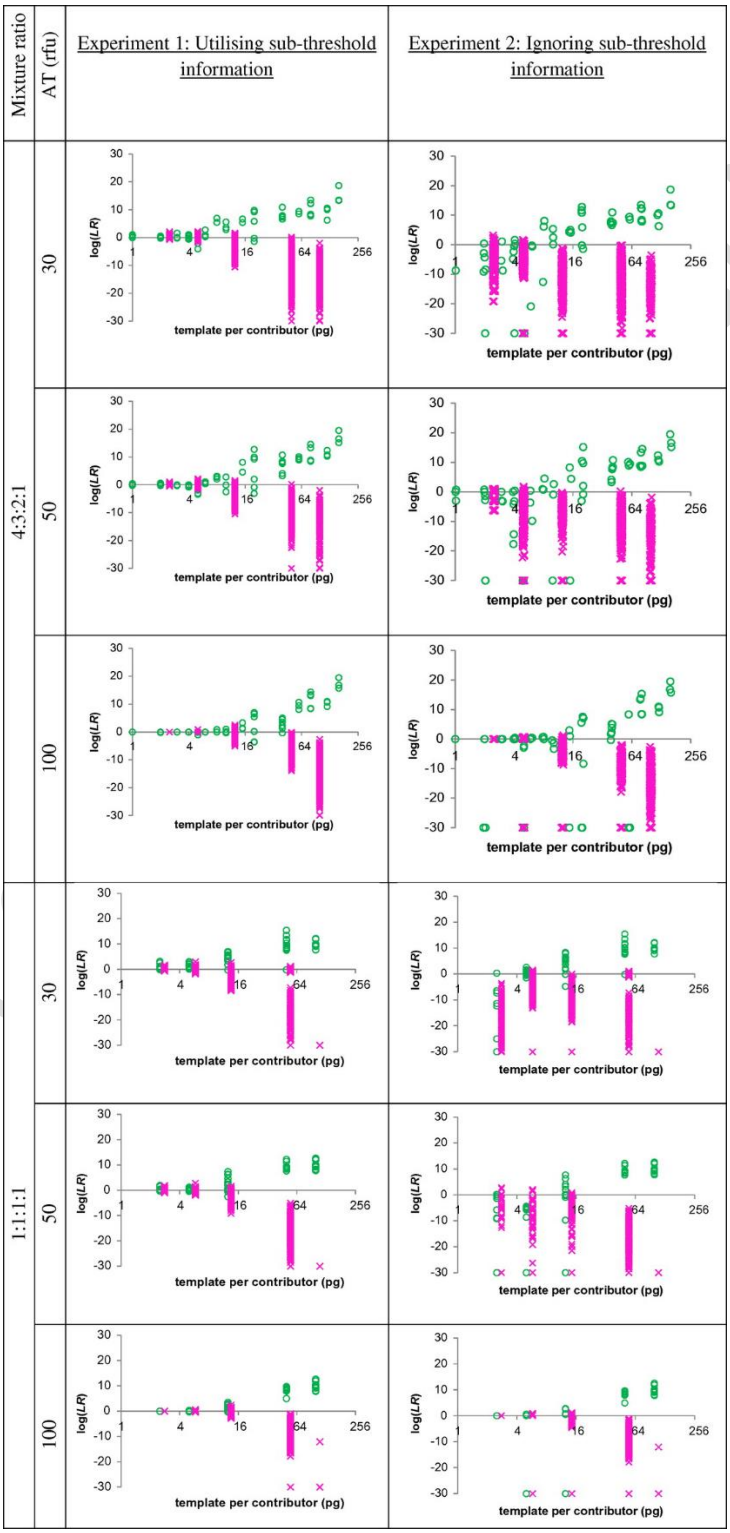
Figure 1 shows the $\log_{10}(LR)$ produced for these comparisons. The *LRs* produced from comparisons to known contributors are signified by a green point and those produced from comparisons to known non-contributors are signified by a pink cross. A minimum value for $\log_{10}(LR)$ of -30 was used, and any *LR*s obtained that fell below this were given the value of -30. The amount of DNA contributed by each known contributor was known from the experimental design. When comparing to non-contributors, the choice of input DNA (for Figure 1) was not known as the non-contributor could align with any of the contributors' input DNA amounts. For known non-contributors the amount of input DNA was assigned as the total amount of DNA added to the PCR divided by the number of contributors. Due to the amount of information present in these graphs we also provide (as supplementary material) the same information but displayed by plotting the $\log_{10}(LR)$ value when considering or ignoring sub-threshold information against each other.

Figure 1 shows that underestimating the number of contributors can cause a $\log_{10}(LR)$ to become less than 0 (sometimes to minimum cap of the graphs) of a true trace contributor in some cases (note the scattered green circles at low $\log(LR)$ for low template). This is the expected outcome for underestimation [4, 5]. We have chosen profiles that are most difficult to interpret due to complexity and high levels of dropout. In addition a detailed examination of peak heights will be of some but limited use since the donor in dispute is trace and at the limits of the AT. In theory there should be a greater ability to exclude using fewer contributors and this is visible in the results (note the generally lower values for the red crosses in the right hand set of graphs in Figure 1).

This experiment looks at consequences of underestimation of *N* and shows that utilising sub-threshold information can partially mitigate the issue. However, use of sub-threshold peaks should be tempered by the relative strength and amount of the putative additional contributor. When assigning a number of contributors based on sub-threshold information there is a risk that an overestimation can occur if any artefacts are considered allelic. It should therefore be balanced by reference to the previously published work [5, 14] which showed that an increase in

N beyond that required, can alter the *LR* for a true trace contributor and mildly increase the risk of low grade *LR* greater than one.

Figure 1. $Log_{10}(LR)$ versus template per contributor (pg) using sub-threshold information (experiment 1) or ignoring sub-threshold information (experiment 2) for a range of four person profiles.



1.2 Change the method by which data is generated (either by lowering the AT or carrying out replicate PCRs)

To investigate the extent to which generating additional data can assist in interpretation we considered two possible strategies, firstly a lowering of the AT and secondly generating additional PCR replicates. It has already been shown [10] that providing additional, relevant information into the analysis of DNA profile data, increases the ability to distinguish a true from a false proposition. We also recognise that due to reasons of practicality there is going to be a limit to which laboratories are willing to lower their AT, and as stated in the introduction, no matter where this level is, there will always be data that appears just below it. We show the effect of lowering the AT as a means to assist laboratories in their choice of AT, when they will inevitably have to weigh up throughput considerations again data generation.

We analyse the 29 mixed DNA profiles outlined in Table 3 using 4 different AT (10, 30, 50 and 100 rfu) and considering each of the three PCR replicates individually or in combination in order to determine the number of contributors.

Table 4 shows the effect that lowering AT, using sub-threshold information or carrying out replicates has on the ability to determine the number of contributors for the data used in this study. For example, inspection of the 1:1:1:1 mixture results at 20 pg individual DNA from Table 4 shows that at AT=50 rfu each of the three individual profiles (1 PCR) appeared to have originated from only one contributor based on allele count. When the AT was reduced to 30 rfu the profiles appeared to have originated from two contributors with more unmasked alleles observed for each contributor. At 10 rfu, when all three replicates are analysed together (3 PCR), the correct assignment of four contributors is made.

1.2.1 Replication

Replication led to some improvement particularly at the fringes when significant portions of the data are dropping out. This can be seen in Table 4 in the 50 pg samples using an AT of 30 rfu, all 6 of these samples individually detected information that could be described by three individuals, but were clearly four when taking multiple replicates into account. The results in Table 4 also show that amplification can only assist so much. Sticking with an AT of 30 rfu, any samples that were amplified with 10 pg or 20 pg of DNA remained describable by fewer than four individuals even with three replicates. For these samples there is a need to consider

what the correct answer is. For example if the peaks above AT come from three of the four contributors the "correct" answer is probably nearer to three rather than four.

There is a resource cost associated with routine repeat amplifications that will need to be considered in forensic laboratories.

Table 4. Assigned number of contributors (based on peak count) are given showing the effect that lowering AT or carrying out replicates has on the ability to determine the number of contributors.

Template (pg)	ratio	replicate	AT = 10 rfu		AT = 30 rfu		AT = 50 rfu		AT = 100 rfu	
			1 PCR	3 PCR	1 PCR	3 PCR	1 PCR	3 PCR	1 PCR	3 PCR
400	1:1:1:1			4		4		4		4
		1	4		4		4		4	
		2	4		4		4		4	
		3	4		4		4		4	
	4:3:2:1	1	4	4	4	4	4	4	4	4
		2	4		4		4		4	
		3	4		4		4		4	
200	1:1:1:1	1	4	4	4	4	4	4	4	4
		2	4		4		4		4	
	4:3:2:1	1	4	4	4	4	4	4	3	3
		2	4		4		4		3	
		3	4		4		4		3	
50	1:1:1:1	1	3	4	3	4	3	3	1	2
		2	4		3		2		1	
		3	4		3		3		2	
	4:3:2:1	1	3	4	3	4	2	3	2	2
		2	4		3		3		2	
		3	3		3		3		2	
20	1:1:1:1	1	3	4	2	2	1	1	0	1
		2	3		2		1		0	
		3	3		2		1		1	
	4:3:2:1	1	2	3	2	2	1	2	0	1
		2	3		1		1		1	
		3	3		2		2		1	
10	1:1:1:1	1	2	3	1	1	1	1	0	0
		2	2		1		1		0	
		3	2		1		1		0	
	4:3:2:1	1	2	3	1	2	1	1	0	0
		2	2		1		0		0	
		3	3		2		1		0	

1.2.2 Lowering the AT

Comparing graphs vertically in Figure 1 shows very little noticeable improvement in the ability to discriminate true from false donors. However comparing rows horizontally in Table 4 suggests that lowering the AT or using sub-threshold information leads to improved ability to assign the number of contributors. There is a cost in expert time in using very low thresholds. Although no evidence is presented here we assume that at very low thresholds even the most skilled experts will let through artefacts occasionally.

Swaminathan *et al.*[15] created a continuous method for contributor number assignment (called NOC*It*) and compared this to maximum allele count and maximum likelihood methods. When carrying out the maximum allele count method they found that allowing the AT to shift to the point of baseline noise (19 to 52 rfu) performed worse at estimating number of contributors than having it fixed at a higher level above baseline noise (50 rfu). While the text does not specifically comment on the reasons for this finding, it may be due to low level artefacts, or stutters appearing above the ratio threshold used being counted as allelic.

1.3 Use informed priors on mixture proportion in a probabilistic system

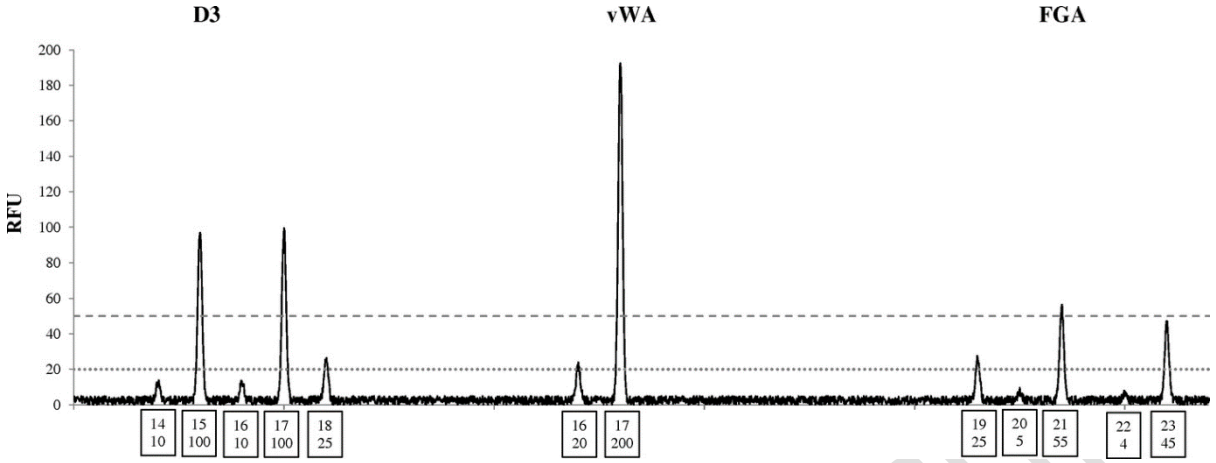
It is possible to provide the analytical system with information that a low level sub-threshold contributor is believed to exist. Consider the mixed DNA profile shown in Figure 2. The known sources of DNA are:

Contributor 1: D3:[15,17], vWA:[17,17], FGA:[21,23]

Contributor 2: D3:[17,18], vWA:[16,18], FGA:[19,19]

In this instance considering the AT as 50 rfu there appears to be a sub-threshold contributor present, however the detected information present in the profile can be described by a single contributor. Peaks detected at 50 rfu are too weak to be paired with complete certainty at D3 or designated as a homozygote at vWA (using only a single replicate), although their pairing would be the most supported combination. There is therefore likely to be a mild impact of the presence of the sub-threshold peaks on the detected peaks, i.e. the presence of the sub-threshold D3:18 means we would accept a [15,18] or [17,18] pairing for the 'major' some proportion of the time with the 17 or 15 peaks (respectively) coming from a second contributor. The analyst may choose to use the presence of the sub-threshold peaks to consider the profile as originating from two individuals.

Figure 2. Three loci of a mixed DNA profile with AT shown as a dashed line for 50 rfu and dotted line for 20 rfu. Boxes show peak designation and height.



We demonstrate the power that providing information, even seemingly minor, can have on the ability of continuous systems to interpret DNA profile data. Before carrying out the experiment there are several predictions that can be made from theory. Consider two *LR*s that could be calculated from these data:

Proposition pair 1

H_{p1}: Contributor 1 and an unknown individual are the sources of DNA

H_{d1}: Two unknown individuals are the sources of DNA

Proposition pair 2

H_{p2}: Contributor 2 and an unknown individual are the sources of DNA

H_{d2}: Two unknown individuals are the sources of DNA

If the profile is analysed as a two person mixture with no guiding information from the analyst even with no significant imbalances in the observed peaks then the analysis will likely split the profile into two roughly equal contributors. Proposition pair 1 will yield an *LR* that favours H_{p1} as most of Contributor 1's peaks are detected, but it will be low as the genotype probability will be spread approximately evenly across a number of genotypes. Proposition pair 2 will yield an *LR* that will likely provide some support for H_{d2} to the profile. The reason for this is that Contributor 2's peaks are not detected and so their presence would have to be explained with multiple dropouts. If the system is supplied with some guiding information that there are two unevenly contributing individuals then we would expect that more weight would be placed on pairing the observed peaks for the major, which we would expect to translate to an *LR* that provides more support for H_{p1} in proposition pair 1. For contributor 2 to be the minor contributor their peaks have still dropped out, however now the system is expecting a low

template contributor and will be more tolerant of dropout. We therefore would expect the *LR* obtained from proposition pair 2 to be closer to one. Finally, when reading to AT of 20 rfu then more information is given to the system. Informed priors for mixture proportion are no longer required as the information being used to interpret the profile is all being used in the analysis. We would expect a divergence of mixture proportion to be obtained naturally from the data provided and that the *LR* produced from either proposition pair will support the corresponding prosecution proposition.

We now turn to results obtained in practice. The DNA profile in Figure 2 was analysed using STRmix™ V2.3.06 firstly using an AT of 50 rfu and providing the system with no information beyond that it has originated from two individuals. Due to the low peak heights under these circumstances the mixture proportions obtained were 47%:53%.

Secondly the same analysis was carried out in STRmix™ but supplying mild prior distributions for mixture proportions of $N(0.75, 0.25)$ for contributor 1 and $N(0.25, 0.25)$ for contributor 2. We use priors on mixture proportion, however realise that it is in fact template DNA amount that these priors will be acting on. Priors for mixture proportions are displayed for the ease of the user because doing so does not need them to consider how other effects within the DNA profile such as degradation and locus specific amplification efficiencies interact with template to generate peak heights. Mixture proportions will automatically scale with peak intensity and so the user does not need to scale their priors for each similarly proportioned mixture. We also recognise that Gaussian distributions extend beyond the interval [0,1] but only apply them within this range.

The mean of the posterior for mixture proportions from the analysis were 85%:15%. The third analysis was for data using AT of 20 rfu, and not providing informed priors for mixture proportion. This time the mean of the posterior for mixture proportions from the analysis were 79%:21%. The *LR*s when comparing contributors to the three analyses can be seen in Table 5. The trend of *LR*s fits what is expected by theory and demonstrates the point that even just supplying the information that the analyst has a prior belief in the mixture proportions based on sub-threshold data (without supplying that specific data to the analysis system) aids in the analysis and produces a result that is more intuitively aligned with the human assessment.

Table 5. LRs produced for comparison to contributors to epg shown in Figure 2.

		Uniform priors AT = 50 rfu	Using informed priors AT = 50 rfu	Uniform priors AT = 20 rfu
Contributor 1	*LR*	63	108	310
Contributor 2		0.097	0.24	6

1.4 Do not interpret the DNA profile

At the laboratory at Forensic Science South Australia an audit of samples received over a one month period revealed that 54% of samples fell into what is classically called transfer or contact DNA and 34% of samples yielded a total DNA concentration of less than 10 pg/µL. There would be many more that would possess less than this level for individual contributors to mixed samples. These profiles are likely to suffer from significant allelic dropout and be within the range where sub-threshold information will be present.

A simple solution to the problems of interpreting epgs with sub-threshold peaks might be to deem all such profiles as too complex; however, given the portion of profiles that this group would represent it is unlikely this would be a sustainable practice. We do not mean this to be an excuse to interpret poor quality data, quite the contrary, instead we mean this statement to highlight the need to determine what data can be interpreted (which we hope we have started in this work).

The question must be asked whether certain profiles *should* be analysed. This is a different question to whether a profile *can* be analysed. Taking a position of theoretical purity, all data can be analysed as long as models exist to describe it. As the information content of the data decreases, or the uncertainty surrounding the interpreted profile increases, there will be an inevitable drop in the discriminating power the model will provide using the data. This is the desired behaviour and correctly represents the strength of the data. There is no limit to which this thinking can be applied. For example the models already exist that an analyst could obtain an epg that exhibits a single weak peak of putative artefactual status and choose to analyse it, considering it may originate from anywhere between one and five individuals. After what is likely to be several hours of processing and analysis, utilising highly complex statistical, mathematical and biological theory and being provided with many pages of detailed output the interpretation system would no doubt inform the analyst of what they already knew, there is no information in the datum to discriminate true from false propositions.

Whether something should be analysed will depend on a number of factors, many of which will not directly relate to the epg in question. Ultimately it will be a decision made by the analyst that the potential discriminating power that epg could provide, in context of the case and laboratory environment, is worth the interpretation and analysis time.

2. Interpretation of putative stutter peaks

When interpreting a DNA profile that has a major component and one or more minor components that are in the same peak height range as stutter of the major then some assessment of the nature of small peaks in stutter positions will need to be made by the analyst.

It is worth discussing the 2006 ISFG [16] recommendation 6 that states:

> 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.

It is the authors' experience this statement is sometimes taken as meaning "*all peaks in stutter positions must be treated as allelic*" as it has been used as such for interpretational attack in court. We suggest that this is not the intent of the authors of [16] when making this recommendation. In the same publication, the preceding sentence gives an example of when the recommendation would have an effect, and states that under those circumstances "*…the probability of stutter must be considered…"*. Probabilistic systems take into account the ambiguous nature of peaks by calculating the probability of that peak if it is purely stutter as opposed to it being partially allelic (given a number of parameter dealing with intrinsic properties of the DNA profile such as DNA amounts, degradation, genotype sets, etc). Sometimes the choice of number of contributors will mean that the certain peaks within the profile will be considered unambiguously as entirely stutter, however this is a perfectly acceptable outcome. To consider all peaks in stutter positions as allelic would see an overestimation of the number of contributors in a large proportion of samples and would be against the ethos that each party is allowed its best explanation of the evidence.

This leaves the analyst with the task of making an assessment of the nature of peaks in stutter positions as to their status. There is a risk here of either overestimating or underestimating the number of contributors to the profile and we point the reader to [4, 5] for the outcomes of either

of these eventualities when using a continuous system including examples of ambiguous stutter peaks. Our intention in this paper is not to trial or recommend methods for dealing with ambiguous peak in stutter positions and we do not do so. All we suggest is that the method used should take into account known stutter values for alleles/loci and the profile should be considered holistically, which may include an assessment of the presence of peaks below the AT.

3. Conclusion:

Continuous systems (at least STRmix™ as trialled here) can overcome the issues of missing low-level data with minimal effects on the outcome of the analysis. The effects of overestimation of the number of contributors may not be too severe as long as the system has been reliably validated for this policy. This situation should not be used to enable a reduction of valid quality practices such as replication and careful expert inspection of profiles and cannot be assumed to be conservative. However, any system, even one possessing the soundest theoretical basis, that cannot withstand the rigours of practical use, is destined to remain nothing more than a nice idea. We have discussed strategies to mitigate the effect of uncertainty in the number of trace contributors present when sub-threshold information is present in a DNA profile. We support replication and lowering the AT whenever practical. The use of sub-threshold data without lowering the AT may be useful in some cases. The effects of miss-assignment of *N* in either direction are relatively mild and restricted to *LR*s less than one when comparing known contributors and low *LR*s greater than one when comparing known non-contributors.

We believe that treating the number of contributors as an unknown nuisance variable is the best long term solution. An even better solution would be to combine the treatment of number of contributors as a nuisance variable with an expert system that utilises fluorescent signal directly and has models for different known artefacts. In such a system all data would be treated probabilistically and the tyranny of thresholds would be completely abolished. We are not aware of any system that can perform at this level and so can provide no examples of how it would perform.

Last, we suggest that some profiles are simply too complex and should not be interpreted. Ultimately it is the role of the scientist to assess each profile on its own merits and the case context in order to determine if and how analysis will proceed.

Acknowledgements

We thank Lisa Melia and Johanna Veth and two anonymous reviewers for their valuable comments that greatly improved this manuscript.

Funding

This work was supported in part by grant 2014-DN-BX-K028 from the US National Institute of Justice. Points of view in this document are those of the authors and do not necessarily represent the official position or policies of the US Department of Justice or Commerce. Certain commercial equipment, instruments, or materials (or suppliers, or software) are identified in this paper to foster understanding. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

References

1. Bregu J, Conklin D, Coronado E, Terrill M, Cotton R, Grgicak C. Analytical Thresholds and Sensitivity: Establishing RFU Thresholds for Forensci DNA Analysis, Journal of Forensic Sciences. 2012; 58(1): 120-29.
2. Taylor D, Bright J-A, McGoven C, Hefford C, Kalafut T, Buckleton J. Validating multiplexes for use in conjunction with modern interpretation strategies, Forensic Science International: Genetics. 2015; http://dx.doi.org/10.1016/j.fsigen.2015.09.011.
3. Gilder JR, Doom TE, Inman K, Crim M, Krane DE. Run-Specific Limits of Detection and Quantitation for STR-Based DNA Typing. , Journal of the Forensic Sciences. 2006; 52(1): 97-101.
4. Bright J-A, Curran J, Buckleton J. The effect of the uncertainty in the number of contributors to mixed DNA profiles on profile interpretation, Forensic Science International: Genetics. 2014; 12 208-14.
5. Bright J-A, Taylor D, Curran J, Buckleton J. Searching mixed DNA profiles directly against profile databases, Forensic Science International: Genetics. 2014; 9 102-10.
6. Benschop C, Haned H, Jeurissen L, Gill P, Sijen T. The effect of varying the number of contributors on likleihood ratios for complex DNA mixtures, Forensic Science International: Genetics. 2015; 19 92-99.
7. Scientific Working Group on DNA Analysis Methods (SWGDAM). Guidelines for the Validation of Probabilistic Genotyping Systems, 2015.
8. Taylor D, Bright J-A, Buckleton J. Interpreting forensic DNA profiling evidence without specifying the number of contributors, Forensic Science International: Genetics. 2014; 13(0): 269-80.
9. Taylor D, Buckleton J. Do low template DNA profiles have useful quantitative data?, Forensic Science International: Genetics. 2015; 16(0): 13-16.
10. Taylor D. Using continuous DNA interpretation methods to revisit likelihood ratio behaviour, Forensic Science International: Genetics. 2014; 11 144-53.
11. Taylor D, Bright J-A, Buckleton J. The interpretation of single source and mixed DNA profiles, Forensic Science International: Genetics. 2013; 7(5): 516-28.
12. Bright J-A, Taylor D, Curran JM, Buckleton JS. Developing allelic and stutter peak height models for a continuous method of DNA interpretation, Forensic Science International: Genetics. 2013; 7(2): 296-304.
13. Bright J-A, Taylor D, J.M. C, Buckleton JS. Degradation of forensic DNA profiles, Australian Journal of Forensic Sciences. 2013; 45(4): 445-49.
14. Bright J-A, Curran JM, Buckleton JS. The effect of the uncertainty in the number of contributors to mixed DNA profiles on profile interpretation, Forensic Science International: Genetics. 2014; 12(0): 208-14.
15. Swaminathan H, Grgicak CM, Medard M, Lun DS. NOCIt: A computational method to infer the number of contributors to DNA samples analyzed by STR genotyping, Forensic Science International: Genetics. 2015; 16 172-80.
16. Gill P, Brenner CH, Buckleton JS, Carracedo A, Krawczak M, Mayr WR, et al. DNA commission of the International Society of Forensic Genetics: Recommendations on the interpretation of mixtures. , Forensic Science International. 2006; 160 90-101.

Appendix 1

Peaks for each of the two contributors were simulated from a lognormal distribution with mean μ and variance $\frac{4}{\mu}$. With probability 0.2 a peak was masked. Masking can be thought of as happening because a major contributor is present or because the two traces mask each other. The number of peaks per locus was counted and any profile that had only 0-2 peaks per locus was checked to see that it did have contributions from each contributor. This is the number of profiles out of the 1,000 simulations appearing in table 1.

Appendix 2

Let

S: be the event that the peaks above AT come from a single source

T: be the event that the peaks above AT come from two sources

AS: be the event that the peaks above AT appear to come from a single source by simple allele count.

Values for the mean μ were drawn from either U[10,50 rfu] and U[10,100 rfu] for each of the two contributors. Pr(*AS*|*S*) and Pr(*AS*|*T*) were calculated using the simulation described in appendix 1 (1,000 simulations were used). Masking was set at 0.2 and 0.5.

The desired probability was obtained as:

$$\Pr(S \mid AS) = \frac{\Pr(AS \mid S)\Pr(S)}{\Pr(AS \mid S)\Pr(S) + \Pr(AS \mid T)\Pr(T)}$$

and assuming Pr(*S*)=Pr(*T*). These values appear in table 2.