Characterisation of forward stutter in the AmpFlSTR® SGM Plus® PCR

Andrew J. Gibb *, Andrea-Louise Huell, Mark C. Simmons, Rosalind M. Brown

SPSA Forensic Services, Aberdeen, United Kingdom

ARTICLE INFO

Article history:
Received 5 February 2008
Received in revised form 3 May 2008
Accepted 6 May 2008

Keywords:
DNA profiles
Forward stutter
Mixtures
Guidelines
Interpretation

ABSTRACT

PCR amplification of tetrameric short tandem repeats (STRs) can lead to Taq enzyme slippage and artefact products typically one repeat unit less in size than the parent STR. These back stutter or n–4 amplification products are low-level relative to the amplification of the parent STR but are widely seen in the forensic community where tetrameric STRs are employed in the generation of DNA profiles. To aid the interpretation of DNA mixtures where minor contributor(s) might be present in comparable amounts to the back stutter products, the typical amounts of back stutter generated have been well characterised and guidelines for interpretation are in place. However, further artefacts thought to be Taq enzyme slippage leading to products with one repeat unit greater than the parent sequence (n+4 or forward stutter) or two repeats less (n–8 or double back stutter) also occur, but these have not been well characterised despite their potential influence in mixture interpretations. Here we present findings with respect to these additional artefacts from a study of 10,000 alleles and include guidelines for interpretation.

© 2008 Forensic Science Society. Published by Elsevier Ireland Ltd. All rights reserved.

1. Introduction

The use of tetrameric short tandem repeats (STRs) as forensic markers exploited by PCR amplification is well established, offering high powers of discrimination from minute amounts of DNA with short turnaround times [1–3]. PCR products are typically visualised by a combination of fluorescent-tagging and electrophoresis, creating a series of peaks the intensity of which relate to template quantity. However, one problem that can occur with PCR amplification of repetitive sequences are artefacts known as shadow bands or stutter. These artefacts are commonplace when amplifying di- and tri-nucleotide repeats even leading to ‘ladders’ of PCR products and a distinct advantage of tetrameric repeats is their relative stability [4–7]. However, although less prominent with tetrameric repeats, stutters do still occur when they are amplified (featuring ladders themselves if over-amplified) with the most common type being back stutter, which results from the loss of one repeat unit in some copies of the amplified STR [2,4]. In forensics, the peaks that result from these artefacts can complicate analysis and interpretation by appearing like true peaks from templates of less quantity such as a minor contributor in a DNA mixture [8].

In the PCR process, the limited processivity of Taq polymerases means that the generation of the extending strand employs multiple polymerases with each making a short section of the new strand before dissociating and being replaced [9–11]. Stutter artefacts are proposed to arise from a lack of fidelity at this changeover via a ‘slipped strand mispairing model’ [8]. In this model, the dissociation of the polymerase allows the template and extending strands to ‘breathe apart’. If upon re-annealing, the template strand has looped out causing mis-alignment by one repeat unit, then the extending strand will be one repeat short, as will all downstream copies created from it [8]. This is back stutter, and if this were to occur on an already shortened copy (or to mis-align by two repeat units) then it would be double back stutter. The proposed model is strengthened by the observation that utilising polymerases of lesser processivity actually increases stutter [8]. Although unproven, it has been proposed that forward stutter arises from equivalent slippage and mis-alignment in the extending strand [8].

For the SGM heptaplex, back stutter peaks were found to be <15% in area of the parent peak [12] and this figure was used as a general interpretational threshold (although this value can be exceeded, especially with low-level parent peaks [13]). More recently, locus-specific thresholds were established for the SGM Plus® hendecaplex [14], reflecting the fact that the differing sequence and repeat-arrangement properties of the loci could have a bearing on the propensity of strand mis-alignment.

These thresholds are critical when considering mixed DNA samples when the stutter peaks are of comparable size to true minor components as information can either be lost or falsely included. Even when a peak from a minor component comfortably exceeds the relevant stutter threshold allowing confidence in its calling, it should be noted that a coincident stutter product may still have influence upon the observed mixture proportion at the locus and as such needs to be taken into account. The current approach to using these established thresholds is reviewed in the national recommendations of the Technical UK DNA working group (TWG) on mixture interpretation for the NDNAD and court going purposes [15], these

* Corresponding author.
E-mail address: andrewgeorgegibb@googlemail.com (A.J. Gibb).
being in turn the UK interpretation of the recommendations of the International Society of Forensic Genetics (ISFG) [16]. The ISFG recommendation on the interpretation of mixtures with regard to stutter reads that “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 \( H_p \) [the prosecution hypothesis] should be included in the assessment.” The TWG response is to utilise the defined stutter thresholds in order to determine whether a peak in a stutter position is indistinguishable

![Fig. 1. Distribution of forward stutter observations for each locus as shown by percentage height of parent peak (primary y-axis and open circles), height in rfu of parent peak (x-axis) and height in rfu of stutter peak (secondary y-axis and closed circles). For all loci tested (those where \( n \geq 3 \)), the relationship between percentage height of parent peak and height in rfu of parent peak was found to be significant (as tested by Spearman ranked and Pearson correlation as appropriate).]
to stutter or can be safely defined as a true allele. Although methods which would allow the incorporation of information about peaks below the stutter thresholds into a likelihood ratio calculation are discussed in the ISFG recommendations [16] and elsewhere [17], many analysts still utilise a binary approach whereby the assignment of a peak as a true allele or stutter is immediately related to its position above or below the threshold respectively.

Although less frequent than back stutter, the additional artefacts of forward stutter and double back stutter when present have the same complicating influence on mixture interpretation. Despite this they remain poorly characterised. Indeed, the TWG response offers the same approach to forward stutter, but accepts that currently thresholds are not in place [15]. In this paper, we examine these artefacts in a study comprising 500 reference samples (buccal scrapes and whole blood) or 10,000 alleles, predominantly chosen from individuals of white northern European origin. We propose interpretational thresholds for forward stutter. In addition, as Taq enzyme slippage has more complicating in chance to occur in longer STR alleles [6,18], we attempted to address this lower level was utilised in this study in order to better define the forward stutter artefact by increasing the size of the dataset (although each incidence of forward stutter was only included provided the artefact was unambiguous from baseline). The ROX threshold remained at 50 rfu as none of the STRs studied here were labelled with this dye.

3. Results

500 single source profiles from reference samples were examined in this study for the presence of forward stutter (n + 4) and double back stutter (n – 8) artefacts. Due to the non-repeat unit nature of the Amelogenin alleles this locus was excluded from consideration.

3.1. Forward stutter

3.1.1. Forward stutter observations

Almost 10,000 alleles were assessed in this study. To accurately represent case work samples and to avoid marked stutter increase, over-amplified and low-level alleles (indistinct from baseline) were

### Table 1a

<table>
<thead>
<tr>
<th>Locus</th>
<th>Median</th>
<th>Maximum</th>
<th>No. observations</th>
<th>Forward stutter incidence</th>
<th>95th percentile</th>
<th>Recommended threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>D3S1358</td>
<td>1.17</td>
<td>5.45</td>
<td>27</td>
<td>4.00</td>
<td>4.30</td>
<td>4.3</td>
</tr>
<tr>
<td>vWA</td>
<td>1.21</td>
<td>2.48</td>
<td>31</td>
<td>4.00</td>
<td>2.26</td>
<td>2.3</td>
</tr>
<tr>
<td>D16S539</td>
<td>0.97</td>
<td>3.33</td>
<td>16</td>
<td>2.48</td>
<td>2.29</td>
<td>2.3</td>
</tr>
<tr>
<td>D2S1338</td>
<td>1.80</td>
<td>3.84</td>
<td>3</td>
<td>0.49</td>
<td>3.63</td>
<td>3.9</td>
</tr>
<tr>
<td>D8S1179</td>
<td>1.24</td>
<td>7.09</td>
<td>55</td>
<td>7.40</td>
<td>4.44</td>
<td>4.5</td>
</tr>
<tr>
<td>HUMD21S11</td>
<td>1.09</td>
<td>8.02</td>
<td>115</td>
<td>14.53</td>
<td>3.30</td>
<td>3.3</td>
</tr>
<tr>
<td>D18S51</td>
<td>1.70</td>
<td>8.92</td>
<td>77</td>
<td>8.80</td>
<td>3.34</td>
<td>3.4</td>
</tr>
<tr>
<td>D19S433</td>
<td>1.21</td>
<td>1.70</td>
<td>6</td>
<td>0.96</td>
<td>1.65</td>
<td>1.7</td>
</tr>
<tr>
<td>HUMTH01</td>
<td>1.83</td>
<td>1.83</td>
<td>1</td>
<td>0.10</td>
<td>–</td>
<td>1.8</td>
</tr>
<tr>
<td>HUMFUBRA</td>
<td>1.31</td>
<td>5.19</td>
<td>16</td>
<td>1.94</td>
<td>4.23</td>
<td>4.3</td>
</tr>
</tbody>
</table>

The median, maximum, 95th percentile and the recommended threshold are listed as percentage height of the parent peak. The recommended thresholds are based on the 95th percentiles for all datasets when n > 10. When this is not the case, the maximum observation is used in preference in order to ensure a conservative value. In addition, all final values are rounded-up to the nearest 0.1%.

### Table 1b

<table>
<thead>
<tr>
<th>Locus</th>
<th>Median</th>
<th>Maximum</th>
<th>No. observations</th>
<th>Forward stutter incidence</th>
<th>95th percentile</th>
<th>Recommended threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>D3S1358</td>
<td>1.39</td>
<td>6.84</td>
<td>27</td>
<td>4.00</td>
<td>4.72</td>
<td>4.8</td>
</tr>
<tr>
<td>vWA</td>
<td>1.13</td>
<td>2.34</td>
<td>31</td>
<td>4.00</td>
<td>2.20</td>
<td>2.2</td>
</tr>
<tr>
<td>D16S539</td>
<td>0.84</td>
<td>3.04</td>
<td>16</td>
<td>2.48</td>
<td>2.13</td>
<td>2.2</td>
</tr>
<tr>
<td>D2S1338</td>
<td>0.86</td>
<td>3.87</td>
<td>3</td>
<td>0.49</td>
<td>3.57</td>
<td>3.9</td>
</tr>
<tr>
<td>D8S1179</td>
<td>1.02</td>
<td>6.31</td>
<td>55</td>
<td>7.40</td>
<td>3.89</td>
<td>3.9</td>
</tr>
<tr>
<td>HUMD21S11</td>
<td>1.04</td>
<td>7.32</td>
<td>115</td>
<td>14.53</td>
<td>3.08</td>
<td>3.1</td>
</tr>
<tr>
<td>D18S51</td>
<td>1.49</td>
<td>8.68</td>
<td>77</td>
<td>8.80</td>
<td>3.33</td>
<td>3.4</td>
</tr>
<tr>
<td>D19S433</td>
<td>1.10</td>
<td>1.34</td>
<td>6</td>
<td>0.96</td>
<td>1.31</td>
<td>1.4</td>
</tr>
<tr>
<td>HUMTH01</td>
<td>2.20</td>
<td>2.20</td>
<td>1</td>
<td>0.10</td>
<td>–</td>
<td>2.2</td>
</tr>
<tr>
<td>HUMFUBRA</td>
<td>1.34</td>
<td>5.19</td>
<td>16</td>
<td>1.94</td>
<td>4.23</td>
<td>4.3</td>
</tr>
</tbody>
</table>

The median, maximum, 95th percentile and the recommended threshold are listed as percentage area of the parent peak. Forward stutter incidence is listed as a percentage of total number of alleles observed (adjusting for homozygote alleles).
excluded from the dataset. From the remaining 9576 alleles, 347 incidences of forward stutter were observed from across all SGM Plus™ loci giving an incidence rate (adjusting for incidences arising from homozygote alleles) of 4.51%. Their total distribution per locus is shown in Fig. 1. Over all loci an increase in percentage height of parent peak (and area, data not shown) was seen with forward stutter arising from parent peaks below 1000 rfu height. Tables 1a and 1b list the medians, number of observations, forward stutter incidence rates, upper 95th percentiles and recommended thresholds for each forward stutter locus dataset by height and area respectively. For datasets where \(n > 10\), the recommendations are based on the 95th percentile but when \(n < 10\) the maximum value observed has been used instead in order to ensure a conservative threshold value. In addition, all final values are rounded-up to the nearest 0.1%. For each locus with \(n > 3\) forward stutter incidences, the relationship between percentage height of parent peak and height in rfu of the parent peak was tested for correlation and each was found to be significant (Spearman ranked correlations: D3S1358, \(r_s = -0.924\), \(p = 0.000\); vWA, \(r_s = -0.856\), \(p = 0.000\); D16S539, \(r_s = -0.942\), \(p = 0.005\); D8S1179, \(r_s = -0.887\), \(p = 0.000\); HUMD21S11, \(r_s = -0.800\), \(p = 0.000\); D18S51, \(r_s = -0.897\), \(p = 0.000\); HUMFIBRA, \(r_s = -0.959\), \(p = 0.005\). Pearson correlation: D19S433, \(r = -0.867\), \(p = 0.025\)). These correlations remained significant when the data with parent peaks below 1000 rfu height were excluded (data not shown).

The loci most prone to forward stutter were HUMD21S11 and D18S51 with a total of 115 and 77 incidences respectively. Conversely, D16S539, D2S1338 and HUMFIBRA showed only one incidence each, with HUMTH01 showing none. In all cases, our own laboratory’s threshold of 50% of the back stutter peak’s area would have been adequate. No relationship was seen between parent peak size and the prevalence of double back stutter (data not shown). Further analyses were not undertaken due to the limited dataset.

4. Discussion

4.1. The relative propensities of back and forward stutter

Walsh et al. [8] proposed that if during enzyme changeover instead of the template strand looping out leading to the extending strand being completed with one repeat unit less than its template, the extending strand looped out then an ‘extra’ repeat unit could be incorporated. This would present as the forward stutter observed in this study. Walsh et al. speculated that the reason forward stutter is the less frequent occurrence is that an association between the Taq polymerase and the 3’-end of the extending strand might inhibit loop out formation in the extending strand [8]. An alternative explanation arises when we consider the relative energetics of expansions and contractions; if there is no breathing apart of the strands, but instead only a single shift of repeat units occurs then Shinde et al. [7] discussed that a contraction would be energetically more favourable. For a contraction to occur only one repeat unit’s bonds from the

![Fig. 2.](image1.png) For vWA and D18S51, the proportion (y-axis) of each allele (x-axis) giving rise to forward stutter showing that for these loci the relationship is significant (Spearman ranked and Pearson correlations respectively).

![Fig. 3.](image2.png) An illustration of how forward stutter can complicate mixture analysis and the need for threshold definition. The diagram represents a two person mixture of major component cd. There are two peaks (a and b) that could be attributable to the minor component. However, peak b is in a forward stutter position. This means that the minor component is either ab or af (i.e. a and any other allele) depending on the relative magnitude of peak b to peak c and so to the forward stutter threshold (of course aa is also a possible call for the minor component should the magnitude of peak a be above the drop-out threshold, but this is unlikely to be the case when considering peaks that are of comparable size to forward stutter). In the diagram, peak b is comfortably above the threshold and so there is sufficient confidence to call the peak a true allele and assign the minor component the genotype ab.
The results in this paper, from a review of 10,000 alleles in single-person profiles, showed that forward stutter occurs at significant enough levels to warrant consideration when interpreting AmpFLSTR® SGM Plus™ multiplex PCRs resulting from mixtures of template DNA. The work showed that the 6% by area guideline for forward stutter previously used at this laboratory was an adequate but conservative guideline for all the SGM Plus™ loci. In line with the recent adoption of ABI-published loci-specific back stutter \( n \neq 4 \) thresholds [14], we recommend similar guidelines based on this dataset for the interpretation of forward stutter peaks in DNA mixture analyses. Fig. 3 illustrates how these thresholds would typically be applied in the binary approach to mixture interpretation. If a peak in a forward stutter position is in excess of the recommended threshold then it can be concluded that at least part of its size is due the presence of a true allele peak. If below the threshold, even if arising from a genuine allele, there is not enough confidence that this is the case and as such the peak must be discarded from consideration. However, due to the elevated forward stutter proportions observed from parent peaks below 1000 rfu height, we recommend caution is applied when considering whether associated minor peaks in stutter positions are indeed genuine alleles.

### 4.3. Correlations between forward stutter and parent peak size

The significant correlations between forward stutter percentage height of parent peak and parent peak height were unexpected. The higher relative magnitudes of forward stutter for the smallest parent peaks (those below 1000 rfu by height) can be explained by the influence of stochastic effects as well as a greater contribution of baseline noise to the recorded stutter peak signal (in line with a similar finding for back stutter by Leclair et al. [13]). However, these effects do not alone offer an explanation for the significant correlations as the significance remained even after the removal the low parent alleles data (data not shown). The observation that the relative magnitudes of forward stutter peaks continue to decrease as parent peak heights increase suggests the involvement of a process-based effect. Despite all the template DNA in this study being from reference samples it could be that the greater peak heights represent PCRs with less inhibitors present and thus less susceptibility to stutter. Alternatively, if there were inaccuracies in the quantification of the template DNA or in its dilution to the optimal value, then this could lead to differing starting quantities of template DNA which could in turn have differing effects on PCR kinetics and so stutter formation. Whatever the cause of these effects, our recommended thresholds include all the observed data in order to provide a single reference point for each locus and a conservative set of guidelines.

### 4.4. Differences in forward stutter between and within loci

The loci varied both in exhibited proportions of forward stutter as well as in the number of observations made. Possible reasons for these differences are discussed below. Most of the loci showed stutter thresholds between 2 and 4% by height, with only D3S1358, D8S1179 and HUMFIBRA exceeding this range, and only D19S433 and HUMTH01 having a value below. The value of 1.8% for HUMTH01 should be used with caution as it is a single value. This low occurrence of forward stutter at HUMTH01 is expected given the observed low occurrence of back stutter at this locus [14].

#### 4.4.1. Allele size

As the slipped strand mispairing model proposes that longer alleles will exhibit greater proportions of stutter due to more changeovers of polymerases during PCR extension [8], we also investigated the relationship between allele size and the relative frequency of forward stutter for each locus. Although most of the loci appeared to show a positive relationship thus supporting the model, only VWA and D18S51 showed significant correlations. The lack of significant correlations could, in part, be due to the small datasets employed (where total forward stutter datasets considered=16 to 115 depending on locus) although differing compositions of some alleles will also have had an effect; Walsh et al. [8] demonstrated with sequencing data that a vWA allele with particularly low back stutter had a core repeat that was interrupted by a related but different repeat. The group presumed that this arrangement inhibited incorrect re-alignment following breathing apart. It is likely that other non-sequential alleles in this dataset will also have had an impact.

#### 4.4.2. Polymorphism

As some suggested mechanisms of STR instability in vivo include replication-mediated expansion/contractions that work through composition-dependent stable single-strand DNA intermediates [24,25], we speculated that, by analogy of the PCR process to replication, more polymorphic STR loci (i.e. those having least copy number fidelity) could be the most prone to stutter effects by way of having more stable intermediates. To investigate whether loci with higher mutation rates gave more forward stutter, the number of forward stutter observations at each locus were was plotted against polymorphism, but no significant correlation was found (data not shown). It could be that differences in the influence of the constituent factors of polymorphism such as repeat-unit arrangements and GC content [26] could affect the possibility of seeing such a straightforward correlation.

### 4.5. Conclusions for forward stutter characterisations

In conclusion, the results presented here demonstrate that forward stutter has the potential to interfere with the interpretation of mixed profiles for all SGM Plus™ loci, but like back stutter it can also be characterised allowing the definition of guidelines. The results confirm that an all-loci encompassing guideline for the consideration of forward stutter in DNA mixture analyses does not reflect the nature of this artefact; the marked variation observed between loci for forward stutter is expected from studies examining back stutter [8,12,14]. Indeed, as different alleles within the same locus do not necessarily differ in a sequential manner and with other factors such as GC content affecting the propensity to breathe apart in the first instance, if all other factors could be held constant, it is likely that allele-specific stutter guidelines would be viable and justified. However, as we have also observed that parent peak height has an influence on the relative magnitude of forward stutter and that this could be compounded by baseline influence and stochastic effects, it is more appropriate for a general guideline per locus to be applied.

### 4.6. Double back stutter

The results in this paper also show that double-back stutter has the potential to interfere with interpretation of minor profiles within mixtures but to a lesser degree, with only 44 incidences and each being less than 50% of the parent back stutter peak. Due to the low number of observations in this dataset, we do not offer loci-specific
thresholds for this artefact, but instead recommend that our general guideline of 50% of the back stutter peak is retained.

4.7. External factors affecting all stutter types

As reaction conditions are critical to the integrity of PCR it follows that differences in the performance of individual PCR machines would lead to differences in the propensity to form stutter. Increasing the injection time of a fragment analysis machine would also be expected to alter stutter ratios by increasing the sampling of peaks arising from stochastic effects. Furthermore, the performance of both PCR and fragment analysers can be affected by external factors such as temperature/location. As all of our PCRs were performed on either one of two PCR machines located adjacent to each other and the product processed on the one fragment analyser with the same injection time we feel we have addressed these factors as far as possible. However, these effects should be kept in mind particularly if higher values of forward stutter than those presented here are being experienced.

5. Recommendations

5.1. When to consider forward stutter

The situations where the artefact of forward stutter would be problematic are those involving mixtures of DNA, particularly two person mixtures where there is a strong major component associated with a weak minor component, for example a rape case with trace evidence from a male perpetrator on an intimate sample from the victim. Due to the often incomplete nature of the trace profile, alleles coincident with stutter positions will be critical, both to inclusion and strength of the statistic and to exclusion and thus elimination of potential suspects. For mixtures of DNA from more than two contributors, the greater chance of coincident peaks, both stutter-allele and allele-allele combinations, means that interpretation is in turn more complicated. However, the characterisation of forward stutter presented here will still aid the expert analyst when deciding which peaks to include when evaluating its significance must of course be followed, thus the same rules must apply whether the an investigation is a match or would eliminate.

5.2. Reproducibility

As forward stutter is a process artefact, it will vary more than the true amplification of a low-level template, and as such, reproducibility of any low-level peak in a forward stutter position will add considerable confidence to its calling as a true allele. Whenever possible, this should be the preferred approach with critical samples.

5.3. Use of intelligence vs. evidence

It should be noted that once generated, even for a two person mixture, the use of these thresholds is not as straightforward as the binary approach suggests. Imagine considering a peak near to the threshold value. The danger with discarding the peak is that it would unknowingly allow an apparent but false match to an accused or suspect. On the other hand, if the peak represents a stutter product but is instead treated as an allele peak, it will more likely lead to false exclusion by chance, potentially de-railing an investigation by eliminating the true culprit. This is because at any one locus the chance that any stutter peak would match by coincidence will always be less than the chance that it could eliminate, as only one or two alleles can ever match while the others would all eliminate. Although we reiterate that we would recommend a repeat PCR wherever possible as the best way of determining the validity of any peak in question, we argue, despite the possibility of intelligence becoming evidence, that in any single instance being conservative and discarding the peak is the better choice to make as a false inclusion is much less likely and a lesser evil than a false exclusion. Indeed, profiles in which such a problem would arise are likely to be low and/or partial with a poor statistic, and their main use would be in helping the police to eliminate suspects, therefore the potential for a false exclusion must be kept to a minimum.

No matter the approach undertaken, in the interpretation of all DNA profiles the principle of deciding the validity of a peak before evaluating its significance must of course be followed, thus the same rules must apply whether the peak in question is a match or would eliminate.

5.4. ISFG interpretation recommendations — alternatives to exclusion

As well as discussing the binary approach to the consideration of stutter in DNA mixtures, the ISFG recommendations in Appendix A [16] suggest a likelihood ratio (LR) approach to stutter peaks that have values below the thresholds that goes beyond simple exclusion. From Fig. 3, suppose that the victim is of genotype cd and that the accused is of genotype aa and that the prosecution hypothesis, \(H_p = V + S\) (i.e. that the mixture comprises victim + suspect). In this situation the defence hypothesis, \(H_d = V + U\) (i.e. that the mixture comprises victim + unknown, which is to say any possible genotypes for the minor component, i.e. aa, ab, ac and ad assuming no drop-out). The standard LR calculation can be written as below with \(p_i\) being the allele probability of the ith allele:

\[
LR = \frac{Pr(E|H_p)}{Pr(E|H_d)} = \frac{P_i^2}{P_i^2 + 2p_ip_s + 2p_ip_v + 2p_ip_d}.
\]

When the probability of stutter, \(Pr(St)\), is close to 0, the LR calculation would favour the defence hypothesis because peak b must then be a true allele and thus aa is excluded as a possible genotype for the minor component (assuming a two person mixture). In the binary approach the possibility of aa would be discarded in the same way because \(Pr(St)=0\) would equate to a value above the 95th percentile-based stutter threshold.

The methods differ however in their treatment of peaks in stutter positions that fall below the thresholds. The advantage of the LR method is that these peaks can still be considered as being true alleles as long as their probability of being stutter is incorporated into the relevant calculation. This means that for a peak in a stutter position, instead of \(Pr(St)\) changing effectively from 0 to 1 as its magnitude drops below the threshold, the probability value lies somewhere between 0 and 1 with the exact value depending on its magnitude relative to the threshold and to the baseline. Thus, in the example, \(H_p\) is not favoured as much as it would be if peak b is discarded as soon as it was below the threshold.

In the example, the numerator of the LR calculation is the probability of the minor component genotype necessary for the prosecution hypothesis to be true (aa). If peak a was the only minor peak observed then the probability of the evidence given the prosecution hypothesis is 1. However, if the peak in the stutter position is a true allele then the aa genotype cannot be represented. Thus, the numerator becomes \(Pr(St)\) as the presence of this genotype requires the occurrence of forward stutter in order to explain peak b. The denominator on the other hand is the probability of the genotypes possible under the defence hypothesis. These include genotype combinations that can occur alongside forward stutter (aa, ac, ad as
well as ab provided peak b is sufficiently large) and these are multiplied by Pr(St), as well as combinations that only arise without forward stutter taking place (ab if peak b is insufficient) and these are multiplied by Pr(St), (the probability of stutter not having taken place (i.e. 1−Pr(St))). With these adaptations, the LR calculation becomes:

\[
LR = \frac{Pr(St)}{p_a^2 + 2p_ap_b + 2p_b^2 + 2p_ap_c + Pr(St)} + \frac{p_c^2}{Pr(St)}
\]

which can be simplified to:

\[
LR = \frac{1}{p_a^2 + 2p_ap_c + 2p_b^2 + (2p_ap_b)Pr(St)}
\]

The smaller the relative magnitude of peak b the larger Pr(St) becomes which in turn increases the LR calculation favouring \(H_p\). Conversely, with bigger relative magnitudes of peak b Pr(St) approaches 0 adding more favour to \(H_d\) until we would conclude that the peak is a true allele and \(H_p\) is rejected. It is argued that the values of Pr(St) can then be derived from experimental observations, such as the ones in this study, and fed into a series of locus-specific calculations. There will of course be a limit on how accurate the Pr(St) values can be. For example from this study we do not consider the 95th percentile as the true equivalent of Pr(St)=0.05. This is because the 95th percentile values are derived from datasets that include the elevated stutter proportions observed when the parent peak heights are <1000 rfu. As is expected for most loci stutter values above the 95th percentile threshold are clustered below 1000 rfu height parent peak. Therefore for peaks where the parent peak height is >1000 rfu the 95th percentile is more akin to Pr(St)≈0. If employed accurately Pr(St) values, the empirical data fed into the LR method would need to take into account the dynamic nature of the artefact as evidenced in Fig. 1, possibly including treating peaks in stutter positions differently depending on the nature of their (potential) parent peak. Even given the correlations between parent peak height and the relative magnitude of forward stutter peaks, the relationships are not exact and small differences in Pr(St) could make big differences to a LR calculation. Nevertheless, this calculation would generally give a lower Pr(St|H_p) than would result from immediately discounting a peak that would favour the defence hypothesis once it dropped below the stutter threshold (or conversely including one as soon as it was above the threshold, i.e. if S=ab) although we do have some concern from some test calculations that we performed whether the method remains conservative at low values of Pr(St) i.e. on the cusp of exclusion/inclusion of a potential stutter peak.

It is worth noting that given the relative low level of forward stutter peaks, it is extremely unlikely that in the given example peak a would ever be assigned a homozygote call as its height would most likely be insufficient to exclude drop-off. Therefore, we suggest that for practical use the only issue to the competing hypotheses in such a circumstance would be whether the call was a|F (allowing aa and therefore \(H_p\)) or ab (excluding aa and therefore not allowing \(H_p\)). The inclusion into the LR calculation of drop-off likelihood is also discussed in the ISFG recommendations and TWG interpretation [16,17].

Although the inclusion of Pr(St) (and by extension Pr(D), the probability of drop-off) into the LR method offers a significant advantage in that all the available data is included in the assessment its added complex nature means that its use is not widespread. Indeed not all laboratories including this one utilise likelihood ratios preferring instead to delineate the competing hypotheses. Nonetheless we look forward to further development and demonstration of this model.

5.5. Sole plank evidence not used in the UK

It should always be kept in mind that other factors than DNA will usually be available in any active investigation to help eliminate any names forwarded as potential matches. In the UK a conviction cannot be based solely on even the best of DNA evidence (the “sole plank” principle) and therefore should such a profile eventually be used as evidence in court the poor statistic would speak for itself.

6. Further work

6.1. Extraction techniques

It is possible that the type of extraction method utilised could affect the degree of stutter artefacts observed. Chelex is a well-utilised extraction method for forensic purposes due to its economic advantages and simplicity of use with rapid processing [19]. Chelex works by absorbing metal ions from the substrate which would otherwise inhibit downstream PCR [19]. However, use of Chelex with degraded samples has suggested that there is an upper limit to its chelating ability and for some samples it may not be sufficient [27,28]. Such criticism has led some in the forensic community to prefer silica based kit extraction procedures [24,29], which are argued to be more effective at purifying the DNA. If free divalent metal ions are able to contaminate the PCR environment, then fewer polymerases will be available for extension. It follows that even mild inhibition could lead to a longer lag time between one enzyme dissociating from the extending strand and another joining on, which in turn would allow more chance of the template and extending strand breathing apart giving rise to looping and therefore stutter. A comparative study of extraction types directly examining stutter effects has yet to be undertaken, but would be of significant interest for the field.

6.2. Spectral effects

Another possible reason for variation between loci that should be considered relates to spectral effects. Partial analysis windows are employed in the ABI GeneMapper™ software [23]. These are designed to not include the primer flare as such high intensity fluorescence could be compensated for by software algorithms resulting in raised baselines and a subsequent drop in sensitivity or ‘burying’ of peaks. It could be that baseline artefacts sometimes seen at the low molecular weight end of the partial window represent the very end of the primer flare. We speculate that if these are compensated for by the software algorithms raising the baseline then they could have the knock on effect that some low-level peaks such as forward stutter are not as readily observed in some dyes as in others. Such possible effects have so far not been addressed.

Acknowledgements

The authors would like to thank SPSS Forensic Services — Edinburgh, Glasgow and Dundee for helpful discussions with respect to forward stutter. We also thank Strathclyde University as Andrea-Louise Huell’s MSc thesis was the basis of this work. We would also like to thank Chris Gannicliffe of this laboratory for helpful comments on the manuscript and Sónia Mendes for guidance with the statistics employed. We also thank our anonymous referees for their constructive guidance. Funding for this project came from the laboratory budget which at the time of the work was provided by Grampian Police and Northern Constabulary.

References


学霸图书馆

www.xuebalib.com

本文献由“学霸图书馆-文献云下载”收集自网络，仅供学习交流使用。

学霸图书馆（www.xuebalib.com）是一个“整合众多图书馆数据库资源，提供一站式文献检索和下载服务”的24小时在线不限IP图书馆。

图书馆致力于便利、促进学习与科研，提供最强文献下载服务。

图书馆导航：

图书馆首页 文献云下载 图书馆入口 外文数据库大全 疑难文献辅助工具