Forensic Science International: Genetics 12 (2014) 1–11

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Analysis of allelic drop-out using the Identifiler® and PowerPlex® 16 forensic STR typing systems

Kirk E. Lohmueller [a,*], Norah Rudin [b], Keith Inman [c]

[a] Department of Integrative Biology, University of California, Berkeley, 1005 Valley Life Sciences, Building 3140, Berkeley, CA 94720-3140, United States
[b] Forensic DNA Consultant, 650 Castro Street, Suite 120-404, Mountain View, CA 94041, United States
[c] Department of Criminal Justice Administration, California State University, East Bay, 4069 Meiklejohn Hall, 25800 Carlos Bee Boulevard, Hayward, CA 94542, United States

ARTICLE INFO

Article history:
Received 29 January 2013
Received in revised form 12 March 2014
Accepted 5 April 2014

Keywords:
Low-template DNA
Allelic drop-out
Detection threshold
Statistics
Likelihood ratio

ABSTRACT

Low-template (LT) DNA profiles continue to present interpretational challenges to the forensic community. Whether the LT contribution comprises the main profile, or whether it is present as the minor component of a mixture, ambiguity arises from the possibility that alleles present in the biological sample may not be detected in the resulting DNA profile. This phenomenon is known as allelic drop-out. This ambiguity complicates both the assessment of the potential number of contributors and estimation of the weight of the DNA evidence for or against specific propositions. One solution to estimating the weight of the evidence is to use a likelihood ratio (LR) that incorporates the probability of allelic drop-out $P(D_O)$ estimated for the specific evidence sample under consideration. However, although a vast repository of data exists, few empirical studies to determine allelic drop-out probabilities have been performed to date. Here we characterized patterns of allelic drop-out in single-source samples using both universal and run-specific analytical thresholds. Not surprisingly, we found fewer instances of apparent drop-out when using a lower (run-specific) detection threshold. Also, unsurprisingly, a positive correlation exists between allele drop-out and allele length, even in good quality samples. We used logistic regression to model the fraction of alleles that dropped out of a profile as a function of the average height of the detected peaks. The equation derived from the logistic regression model allowed us to estimate the expected drop-out probability for an evidentiary sample based on the average peak height of the profile. We show that the LRs calculated using the estimated drop-out probabilities were similar to those calculated using the benchmark drop-out probabilities, suggesting that the estimates of the drop-out probability are accurate and useful. This trend holds even when using the data from the PowerPlex® 16 typing system to estimate the drop-out probability for an Identifiler® profile, and vice versa. Thus we demonstrate that use of a LR that incorporates empirically estimated allelic drop-out probabilities provides a reliable means for extracting additional information from LT forensic DNA profiles.

1. Introduction

Many biological samples typed in a forensic context contain a limited amount of DNA. Because relatively few copies of each locus may be present, the random sampling of DNA molecules during the amplification process may result in the failure to observe some alleles in the evidentiary profile that were actually present in the original biological sample. This phenomenon is known as allelic drop-out [1]. In other words, the final DNA profile may not accurately represent the profile of the person or people who deposited the sample. The possibility of allelic drop-out can severely complicate the interpretation of forensic DNA profiles, even those obtained using a standard number of PCR cycles and standard electrophoretic injection parameters [2–10]. In the instance that an evidence sample lacks an allele present in the profile from a suspected contributor, the absence of this allele can be explained in two ways. Under one proposition, the allele was not detected because the true contributor to the evidence profile lacks that particular allele; in this scenario, the true contributor and the suspected contributor are different individuals. Alternatively, the allele may have been present but not detected because it

* Corresponding author. Present address: Department of Ecology and Evolutionary Biology, University of California, Los Angeles, 621 Charles E. Young Dr. South, Los Angeles, CA 90095-1606, United States. Tel.: +1 310 825 7636; fax: +1 310 206 048.
E-mail address: klohmueller@ucla.edu (K.E. Lohmueller).

http://dx.doi.org/10.1016/j.fsigen.2014.04.003
1872-4973/

had dropped out of the evidence profile; under this scenario, an explanation exists for the difference and we must quantify the probability of seeing the observed profile if the proposed explanation is true.

Due to these complications, workers have proposed various statistical approaches that account for the possibility of allelic drop-out when analyzing low-template (LT) DNA evidence [2–5,9,11–18]. One approach uses a stochastic threshold. The stochastic threshold is the peak height below which drop-out is considered to be possible [13,19]. Such loci are then typically excluded from any statistical calculations to weight the DNA evidence [13,19]. Drop-out is considered to be unlikely for loci where all peaks are above the stochastic threshold. Thus, these loci are analyzed using the standard calculations employed for DNA evidence present in sufficient amounts. One potential drawback to the stochastic threshold approach is that many loci may be discarded from the calculation, especially in mixtures with low-level minor contributors, resulting in the loss of important information. While this often serves simply to underweight the evidence with regard to some particular individual, ignoring these loci may also result in discarding potentially exclusionary information as well. Further, underweighting the evidence against a possible contributor may under certain circumstances be prejudicial to a defendant if, for example, that contributor is an alternate suspect [20]. Finally, we note that it is never permissible to exclude loci from statistical calculations in which, after comparison between the evidence profile and the profile of the suspected contributor, alleles that are present in the profile of the suspected contributor are not detected in the evidence profile. Such an approach assumes that drop-out has occurred and ignores the alternative explanation for the missing allele: the true contributor to the evidence was a homozygote and no drop-out is required to explain the evidence. Further, this approach was shown to frequently over-state the strength of the evidence [21]. Rather, if loci are to be excluded from statistical calculations at all (not an approach we recommend), this must be done prior to the comparison to possible contributors.

Another statistical tool that can be applied to estimate the weight of LT DNA evidence is a likelihood ratio (LR) that can incorporate the probability of allelic drop-out into a LR [2–4,11,16–18,22,23]. Freely available computer programs facilitate the calculation of LRs with a model of drop-out for profiles containing components in the stochastic range, where allelic drop-out is possible [4,24,25]. One advantage to this approach is that no need exists to restrict the statistical calculation to only those loci where all peaks exceed a stochastic threshold, allowing all loci to be used in assessing the strength of the evidence. A probabilistic framework also relieves the analyst from deciding, *a priori*, whether the suspected contributor should be included or excluded. Rather, a hypothesis and an alternate hypothesis are stated, and a LR calculated for the comparison; the direction and distance from one then informs us as to the support, or lack thereof, for each of the named pair of propositions.

LR approaches that are used to assess the weight of LT DNA evidence can include a parameter that describes the probability of allelic drop-out $P(D_O)$. Thus, $P(D_O)$ must be estimated for a particular evidence profile in order to compute the LR. Several different approaches have been implemented to obtain an estimate of $P(D_O)$. Each of the various approaches possesses different advantages and limitations. It is not yet clear whether any one approach performs better than another. One type of method estimates $P(D_O)$ from the evidence profile directly, exclusive of external data. This is accomplished by finding the value of $P(D_O)$ that maximizes the probability of the evidence under each of the hypotheses being compared in the LR. While promising, this approach is computationally intensive [24]. Additionally, it will require empirical validation using ground truth samples. A second category of approaches estimates $P(D_O)$ using ground truth samples for which the specific drop-out events are known. Specifically, the relationship between drop-out and either peak heights or the quantity of DNA is determined from the validation data. This relationship is used to estimate $P(D_O)$ for a particular evidence profile. Recently Tvedebrink et al. made a significant advance in this area by using a logistic regression framework to model the drop-out probability based on the average peak height in a profile [26–28]. The logistic regression model parameters are estimated from validation data in which the relationship between the average peak height and the proportion of alleles that have dropped out is known. The curve generated by the logistic model can then be used to predict the drop-out probability for an evidentiary sample based on the average RFU values of the peaks detected in the profile. However, this group did not assess the quality of their estimates of the drop-out probability by comparing the estimated LRs, to ground-truth or whether the estimated drop-out probabilities were kit or lab specific.

Here we assess allelic drop-out in single source LT DNA profiles from the Identifiler® and PowerPlex® 16 typing systems. We provide a formula to estimate drop-out probabilities as a function of average peaks height using logistic regression (similar to what was done by Tvedebrink et al. [26–28]). Importantly, none of the previous studies using peak heights [20,26–28] have evaluated the accuracy of the final LRs calculated using the estimated drop-out probabilities. To address these potential concerns, we performed experiments to assess the accuracy of LRs calculated using empirically derived drop-out probability estimates. Simulations show that LRs calculated using an estimated $P(D_O)$ are comparable to those calculated using the benchmark $P(D_O)$. This work also suggests that, despite the existence of statistically significant correlations between drop-out and allele length, using logistic regression to estimate an average drop-out probability still produces a final LR that accurately weights the strength of the LT DNA evidence. This work provides new insight into allelic drop-out and contributes to expanding the use of probabilistic approaches using LRs to interpret LT DNA profiles.

2. Methods

2.1. Description of the data

We analyzed data from single-source DNA profiles generated by John Butler's group at the National Institute of Standards and Technology (NIST) [29]. Briefly, DNA from two individuals (MT97150 and PT88411) was adjusted variously to amplify 100 pg, 30 pg or 10 pg. Each of the six samples (three concentrations of each of the two individuals) was amplified 10 times with both the Identifiler® and PowerPlex® 16 kits using the standard number of PCR cycles suggested by the manufacturer (28 cycles for Identifiler and 30 cycles for PowerPlex). A total of 60 profiles were thus produced for each kit. Further information regarding the experimental conditions is given in [29]. In our study, data from the two different kits, and for each condition, were analyzed separately unless otherwise noted. The data from each run was analyzed using GeneMapper® software with the default stutter filters activated, providing both allele calls and RFU values for the peaks detected by the software. Pull-up, spikes, and obvious kit artifacts were eliminated by visual inspection of each electropherogram. The electropherograms were further inspected to identify peaks that were above the analytical threshold, but were foreign to the known input profiles, for example peaks in the minus or plus stutter positions that rose above both the analytical threshold and the default stutter filter percentages. For the PowerPlex® 16 profiles, no peaks remained that met these criteria. For the

Identifiler® profiles, several peaks in plus or minus stutter positions rose above the default stutter threshold, and one possible drop-in allele was detected. These observations will be addressed in Section 4. Obvious artifacts and peaks in the stutter positions were removed from the electropherograms prior to statistical analyses. LRs were computed both including and excluding the possible drop-in allele and results were nearly identical (see Section 4).

2.2. Statistical analyses

We used two different methods to set an allelic detection threshold for each LT profile. One method was the standard approach used by most forensic laboratories. This typically involves picking the lowest analytical threshold at which most detected peaks result from DNA signal rather than instrument noise. Most labs currently use thresholds between 50 RFUs and 150 RFUs. We elected to apply a detection threshold of 50 RFUs. We also used a limit of detection/limit of quantitation (LOD/LOQ) approach to determine a data-driven allelic detection threshold for these samples [19,30–36]. This threshold was determined by using the average instrument noise plus 10 standard deviations of the noise in the reagent blank samples [19,30–35]. For these data, that threshold was determined to be approximately 30 RFUs.

We compiled several summary statistics from each profile. First, we tabulated the average height (in RFUs) of all peaks present in the profile. Let $H_{ir.}$ be the average height of all detected peaks in the rth replicate from individual i. We chose to use the average peak height across all peaks in the profile, rather than calculate individual statistics for each locus, because less sampling variation exists in estimating an average compared to a single data point. Further, Tvedebrink showed that average peak height was a useful statistic [26]. Second, we determined the proportion of alleles that dropped out of a given profile. For the purposes of this paper, we define drop-out to be the situation where a particular peak known to exist in the sample does not rise above the allelic detection threshold, and, as a result, is not detected in the profile. Define d_{irl} to be the number of alleles that dropped out of the lth locus in the rth replicate profile from individual i. Then, $d_{ir.}$, or the total number of alleles that dropped out of replicate r from the ith individual is simply

$$d_{ir.} = \sum_{i=1}^{15} d_{irl}.$$

Then, let $P_{ir.}$ be the proportion of alleles that dropped out of the rth replicate for individual i. Then

$$P_{ir.} = \frac{d_{ir.}}{30}$$

because both individuals analyzed were heterozygous at all loci.

We used a logistic regression framework to model the proportion of alleles that dropped out of each profile as a function of the average height (in RFUs) of all peaks present in the profile using the following equation,

$$In\left[\frac{P_{ir.}}{1-P_{ir.}}\right] = \beta_0 + \beta_1 H_{ir.}$$

where $P_{ir.}$ is the proportion of alleles that dropped out of the rth replicate for individual i, and where $H_{ir.}$ is the average height of all detected peaks in the rth replicate from individual i. In this model β_0 and β_1 are the intercept and slope, respectively, of the logistic function. We estimated β_0 and β_1 from the Identifiler® and PowerPlex® 16 profiles using the **glm** function in the R statistical package [37]. The logistic regression equation was then used to obtain an estimate of the drop-out probability for a particular profile, given the peak heights of that profile. This estimated drop-out probability is denoted as $P(\hat{D}_O)$ and is used for all loci within the profile of interest.

We also tested whether the amplicon size affected the proportion of alleles that dropped out at a particular locus. We took the average length of the two alleles (in units of bp) at each locus for each of the two individuals. This quantity, $L_{i.l}$, is the average length of the alleles at locus l in individual i. The length of each allele was taken from STRbase [38]. We next tabulated the proportion of alleles that dropped out from all profiles for a particular individual at a particular locus. Let $d_{i.l}$, be the number of alleles that dropped out at locus l across all 30 replicate profiles for individual i. It is calculated as

$$d_{i.l} = \sum_{r=1}^{30} d_{irl}.$$

We then used linear regression to model the relationship between these two variables. Specifically, we used the model,

$$d_{i.l} = \alpha_0 + \alpha_1 L_{i.l},$$

where α_0 and α_1 are the intercept and slope, respectively, of the linear model. We also tested if d and L were correlated using Pearson's and Spearman's correlation coefficients.

2.3. Evaluating the performance of $P(\hat{D}_O)$

We evaluated the performance of our estimates of the drop-out probability, $P(\hat{D}_O)$ that were obtained from the logistic regression approach described above. Briefly, each of the LT DNA profiles, simulating an evidentiary profile, was compared with the profile of a suspected contributor. For each of these comparisons, we computed two different LRs, one using the estimated drop-out probability, $P(\hat{D}_O)$, and another using the benchmark drop-out probability (defined as $P(\hat{D}_O)$). The benchmark drop-out probability is the maximum likelihood estimate of the drop-out probability from a binomial distribution if one could directly count the number of alleles that had dropped out. From ground-truth samples, such a count was possible. This would not be possible in casework analyses. The benchmark drop-out probability for each locus within each profile was calculated as follows. Let $P(D_{irl})$ be the benchmark drop-out probability for replicate r from the ith individual at locus l, then:

$$P(D_{irl}) = \begin{cases} 0, & d_{irl} = 0 \\ 0.5, & d_{irl} = 1 \\ 1, & d_{irl} = 2 \end{cases}$$

If the estimated drop-out probabilities are accurate, then the LRs calculated using them should be similar to those calculated using the benchmark drop-out probability. This approach is similar to that of the Tippett plots previously applied to forensic DNA profiles [23,39–41].

To assess how well the drop-out estimator functions, we calculated LRs comparing the LT profiles (in the role of evidentiary samples) to both true contributors and to known non-contributors. This type of comparison has become a standard tool in investigating statistical models applied to complex forensic DNA profiles (see Refs. [18,23], for example). Specifically, to model a true contributor, we compared the hypothesized contributor (one of the LT profiles) to the full profile of the known true contributor. In contrast, to model a known non-contributor, a randomly generated individual simulated from a US population database with European ancestry was compared to the hypothesized

contributor. This comparison was performed using the LT samples derived from both MT97150 and PT84411.

For the known non-contributors, we generated two populations of 5000 random individuals by drawing alleles from a multinomial distribution with parameters 2 and **p**, where **p** is the vector of allele frequencies at the particular locus, using the allele frequencies published in Butler et al. [42] for the Identifiler® profiles and from Levadokou et al. [43] for the PowerPlex® 16 profiles. The model used to simulate the genotypes assumes Hardy–Weinberg equilibrium and linkage equilibrium.

LRs were calculated for each of the LT DNA profiles generated by Butler against each of the randomly generated known non-contributors using the R-code developed by David Balding, based on Balding and Buckleton [4]. The code was modified to run in an automated manner within a bash shell script. We set the drop-in probability to 0.01 and the population structure adjustment to $\theta = 0.01$. The homozygous drop-out probability was computed as $\alpha P(\hat{D}_O)^2$, where $\alpha = 0.5$, as suggested by Balding and Buckleton [4]. If the LT profile simulating the evidence profile originated from the first individual (MT97150), then the logistic equation from the 30 profiles from the other individual (PT88411) was used to calculate $P(\hat{D}_O)$. Similarly, if the LT evidence profile originated from the second individual, the logistic equation from the first individual was used to calculate $P(\hat{D}_O)$. This procedure avoids using data employed to estimate the parameters of the model to test the adequacy of the model.

3. Results

3.1. Overall patterns of drop-out

The number of alleles that dropped out of a particular LT DNA profile ($d_{ir.}$) was correlated to the average height of the peaks across all loci ($H_{ir.}$; Fig. 1, Table 1). Table 1 shows the number of alleles that dropped out of each of the LT profiles compared with average peak height. As one might predict, drop-out generally occurred more frequently in profiles with lower average peak heights for both of the typing systems. Drop-out tended to be rare for profiles with $H_{ir.} > 200$ RFUs.

Apparent drop-out was, again unsurprisingly, also influenced by the detection threshold imposed upon the profile. For example, significantly more alleles dropped out from each Identifiler® profile when using the 50 RFU detection threshold (average of 11.8 alleles) than when applying the 30 RFU detection threshold (average of 9.2 alleles, $P < 10^{-6}$, paired t-test). Table 1 shows that this trend holds for both typing systems and across the entire range of average peak heights. The number of profiles in which no alleles were detected was, as expected, greater in the data using a 50 RFU threshold (9) than from the data generated using a 30 RFU threshold (4). These findings indicate that many alleles which would otherwise be discarded at a policy-driven detection threshold can be recovered if a data-driven threshold is employed.

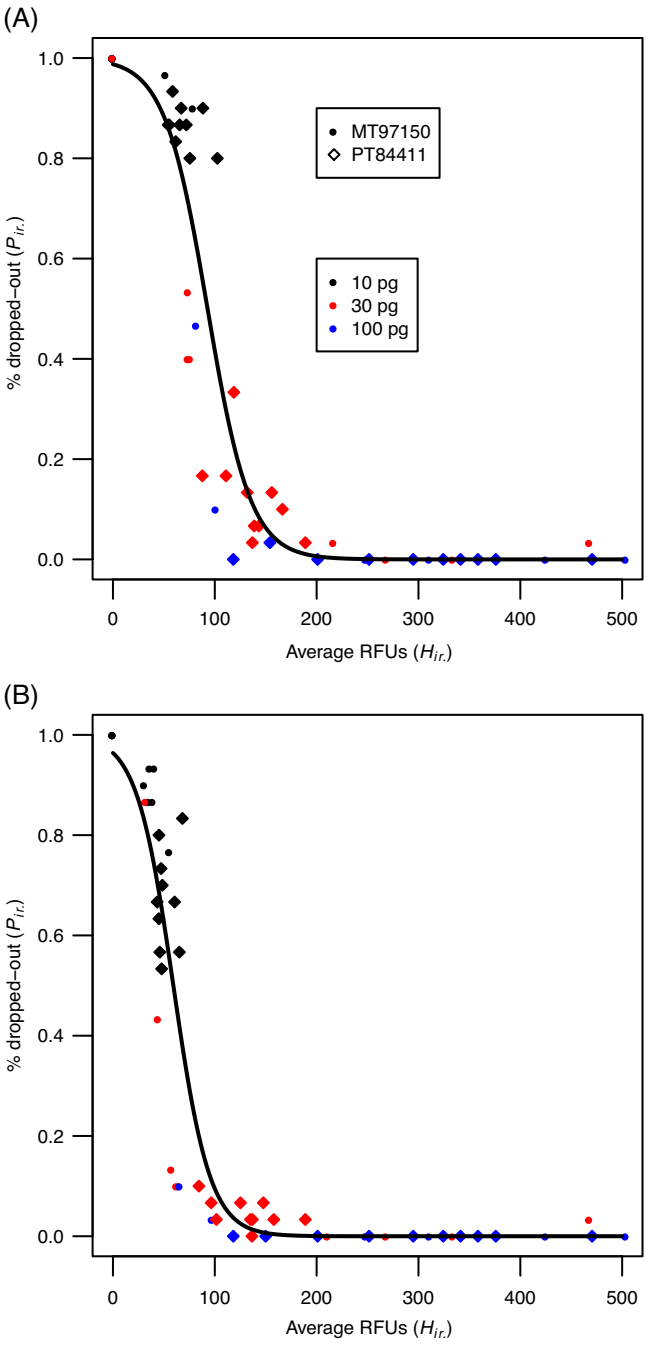

Fig. 1. Relationship between the proportion of alleles that dropped out of each Identifiler® profile ($P_{ir.}$) and the average RFUs of all detected peaks in the profile ($H_{ir.}$). (A) 50 RFU detection threshold, and (B) 30 RFU detection threshold. The logistic regression curves were fit to all 30 profiles from individual MT97150 and 30 profiles from individual PT8441. The Identifiler® data are depicted. *Note*: the plotting area was truncated at 500 RFUs. Several points fell beyond this region.

Table 1
Average (SD) number of alleles that dropped out of each profile, grouped by average height of the detected peaks across all loci.

Typing system	Detection threshold (RFU)	$H_{ir.} < 100$ RFUs[a]	$100 < H_{ir.} < 200$ RFUs[a]	$H_{ir.} > 200$ RFUs[a]
Identifiler®	50	24.8 (6.9)	4.6 (6.4)	0.1 (0.3)
	30	18.5 (10.1)	0.9 (0.7)	0.1 (0.2)
PowerPlex® 16	50	17.7 (3.6)	1.8 (3.9)	0.1 (0.3)
	30	10.7 (3.7)	0.46 (1.0)	0.1 (0.2)

[a] $H_{ir.}$ denotes the average height of all detected peaks within the profile.

3.2. Estimation of drop-out probabilities as a function of peak height

The correlation of the proportion of allele drop-out from a profile ($P_{ir.}$) with the average peaks heights within a particular profile ($H_{ir.}$) supports using logistic regression to model the relationship between these two variables (Fig. 1 and Table 2). Based on the clustering of the points along the logistic curve, the model visually appears to fit the data quite well for both the 50 RFU and the 30 RFU detection thresholds (Fig. 1). Estimates of the parameters for the logistic model applied to different subsets of the

Table 2
Parameter estimates from the logistic regression model.

Detection threshold (RFU)	Typing system	Intercept MT sample	Slope MT sample	Intercept PT sample	Slope PT sample	Intercept both samples	Slope both samples
50	Identifiler	4.24	−0.051	5.10	−0.052	4.39	−0.047
	PowerPlex	3.49	−0.038	4.03	−0.042	3.76	−0.040
30	Identifiler	4.52	−0.086	3.07	−0.048	3.29	−0.055
	PowerPlex	2.07	−0.039	1.73	−0.033	1.86	−0.036

data are presented in Table 2. The satisfactory fit of the logistic model was also observed for the PowerPlex® 16 data (Fig. S1).

As expected, the 100 pg samples (blue points in Fig. 1) tend to have higher average peak heights (>150 RFUs) and less drop-out than samples containing smaller amounts of DNA. However, several exceptions to this pattern were observed. For example, nearly half of the alleles dropped out of one 100 pg profile with $H_{ir.} < 100$ RFUs (Fig. 1A).

3.3. Comparison of logistic model parameters across different experimental conditions

One goal of our study was to evaluate the differences in patterns of drop-out across different samples, detection thresholds, and typing systems. Fig. 2 shows the logistic regression curves applied to different sub-sets of the data. Overall, the curves are roughly similar to each other, suggesting the same general patterns of drop-out for all of the data in our study. However several statistically significant differences in model parameters are evident (Tables S1–S3). The greatest difference was seen when comparing drop-out probability curves from the 30 RFU detection threshold (red curves in Fig. 2) to those from the 50 RFU detection threshold (black curves in Fig. 2; Table S1). The curves from the 30 RFU detection threshold were shifted to the left relative to those from the 50 RFU detection threshold. This is expected based on our earlier observations that more drop-out is displayed by the 50 RFU data. The curves from the MT97150 and PT84411 Identifiler® data appear different from each other, especially at the 30 RFU threshold (Table S2). Interestingly, the curves from the Identifiler® and the PowerPlex® 16 data are very similar to each other, with few statistically significant differences (Table S3).

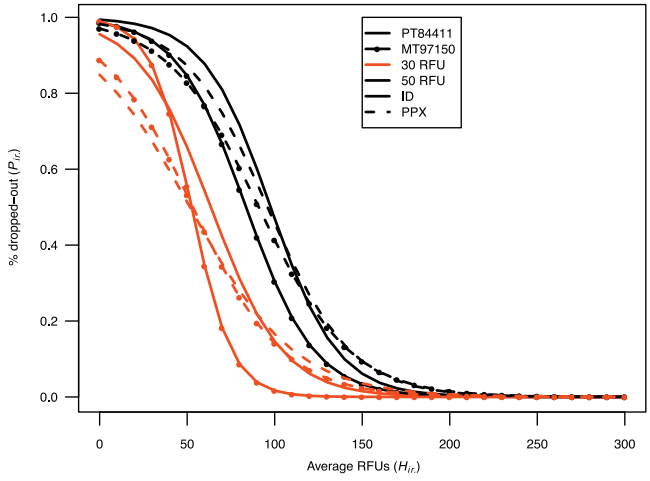

Fig. 2. Logistic curves showing the relationship between the probability of drop-out and average RFUs for different detection thresholds (30 RFUs and 50 RFUs), individuals (PT84411 and MT97150), and typing systems (Identifiler® and PowerPlex® 16).

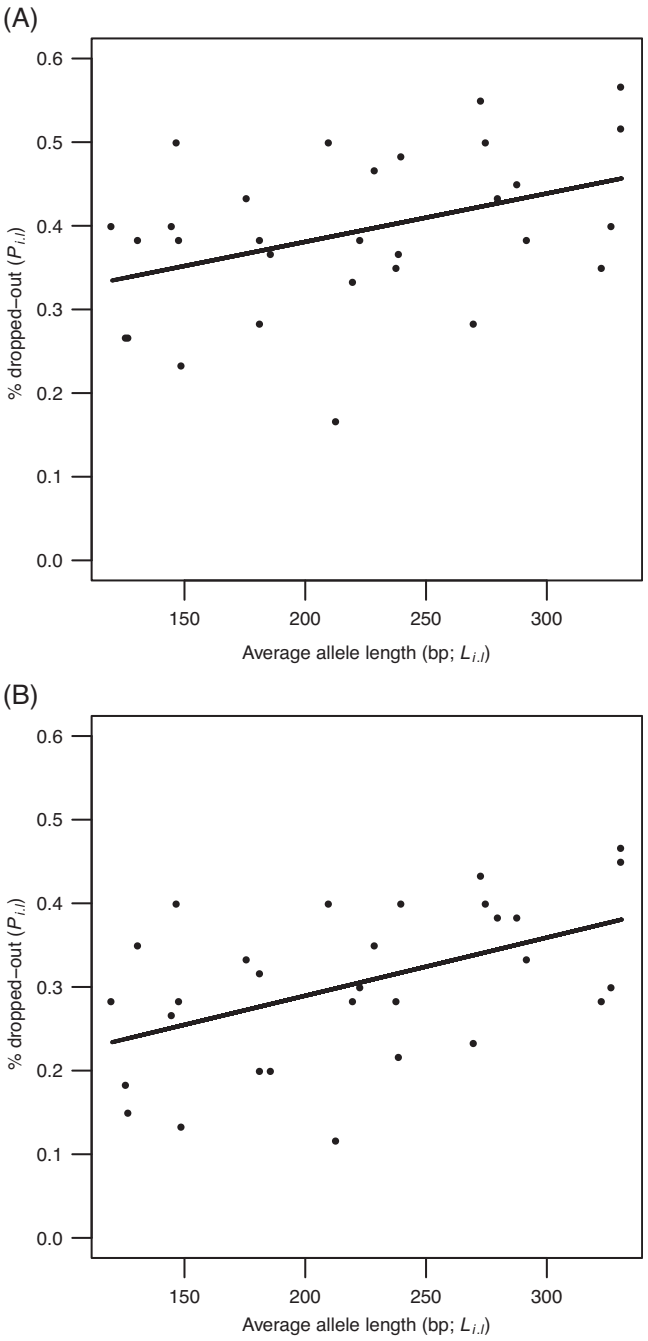

Fig. 3. Relationship between the proportion of alleles that dropped out at a particular locus in a particular profile ($P_{i.l}$, y-axis) and the average length of the alleles in the particular individual at that particular locus ($L_{i.l}$, x-axis). Data from the 60 Identifiler profiles are depicted using (A) a 50 RFU detection threshold and (B) a 30 RFU detection threshold. The line is the best-fit linear regression line fit to the data.

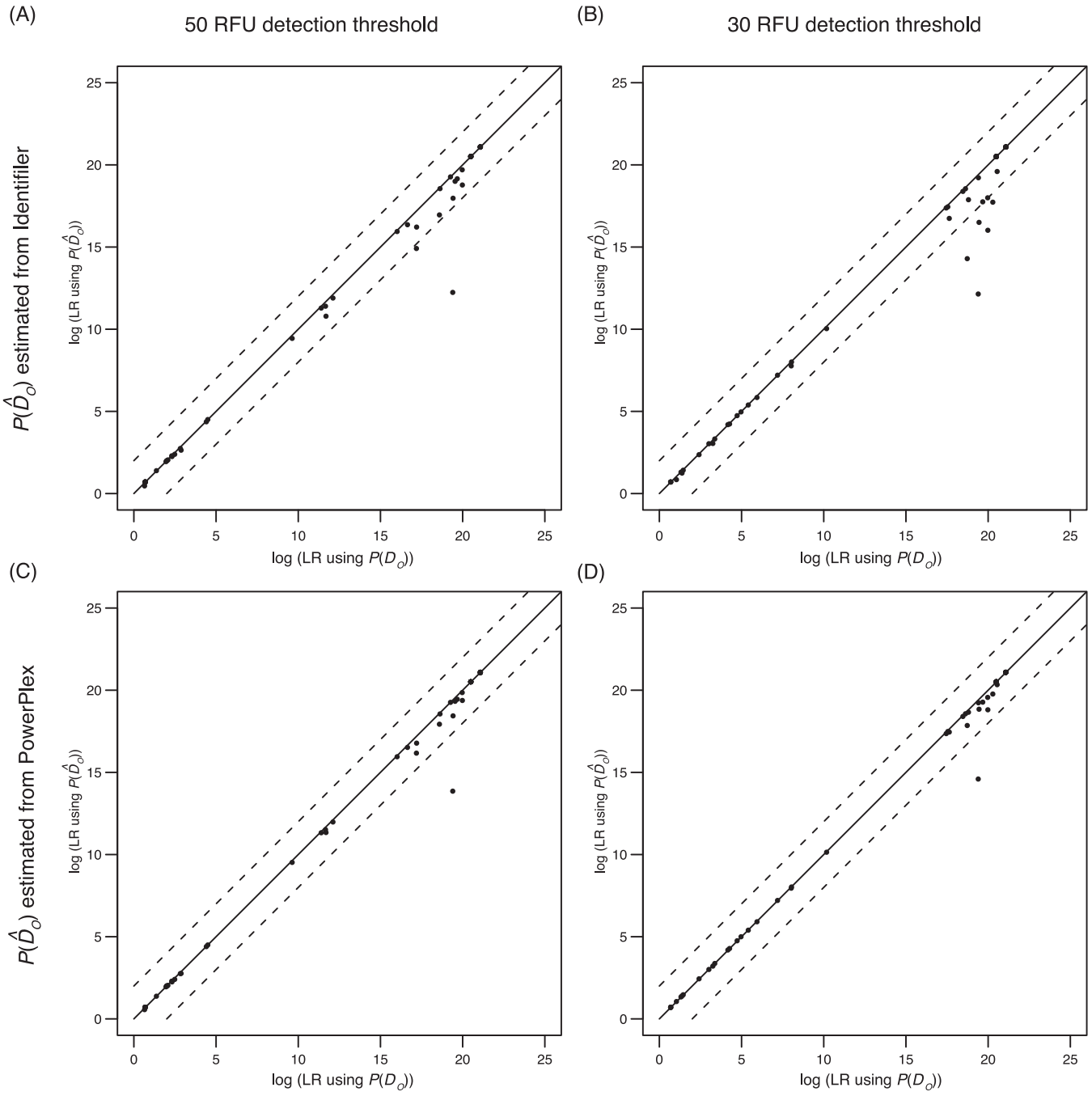

Fig. 4. LRs calculated from the Balding and Buckleton program when the hypothesized contributor was the true donor of the Identifiler LT DNA profile. Each dot represents the log(LR) for a particular LT DNA profile calculated using the estimated drop-out probability ($P(\hat{D_O})$, *y*-axis) vs. the benchmark drop-out probability ($P(D_O)$, *x*-axis). The solid line is the diagonal and the dashed lines denote 2 orders of magnitude in either direction around the diagonal. (A) 50 RFU detection threshold, where $P(\hat{D_O})$ was estimated from the other individual's Identifiler data. (B) 30 RFU detection threshold, where $P(\hat{D_O})$ was estimated from the other individual's Identifiler®. (C) 50 RFU detection threshold, where $P(\hat{D_O})$ was estimated from the other individual's PowerPlex® 16 data. (D) 30 RFU detection threshold, where $P(\hat{D_O})$ was estimated from the other individual's PowerPlex data® 16.

3.4. Drop-out and allele length

We next examined whether allele length influenced drop-out in a locus-specific manner. Fig. 3 shows the proportion of alleles that dropped out at each locus in each of the two individuals ($P_{i,l}$) vs. the average length of the alleles in individual i at locus l ($L_{i,l}$). We find a strong positive correlation between these two variables (50 RFU detection threshold: Spearman's $\rho = 0.39$, $P = 0.034$, Pearson's $\rho = 0.41$, $P = 0.025$; 30 RFU detection threshold: Spearman's $\rho = 0.49$, $P = 0.006$, Pearson's $\rho = 0.50$, $P = 0.005$). Thus, those loci with longer average amplicons tended to show higher levels of drop-out.

3.5. Comparison of estimated and benchmark drop-out probabilities when used in LRs

To evaluate the efficacy of our estimates of drop-out on the results of LRs in which they are used, we first evaluated LRs in which the numerator represents the true donor and the denominator represents a random individual. For each of the 60 Identifiler® profiles, we calculated the LR using both the estimated drop-out probability ($P(\hat{D_O})$) and the benchmark drop-out probability ($P(D_O)$, see Section 2). Across all four different sets of comparisons, the LRs calculated using $P(\hat{D_O})$ typically fall within

two orders of magnitude of those calculated using $P(D_O)$ (Fig. 4). When the LRs differ by more than two orders of magnitude, the LRs using $P(\hat{D}_O)$ are lower than those obtained from $P(D_O)$. This pattern holds for both the 50 RFU and the 30 RFU detection thresholds, as well as when using the equations from the PowerPlex® 16 data to estimate $P(\hat{D}_O)$ for the Identifiler® profiles. A similar high degree of concordance between the LRs calculated using the estimated drop-out probabilities and the benchmark drop-out probabilities was also observed for the PowerPlex® 16 data (Fig. S2). The one outlier data point in all four panels of Fig. 4 was for a profile from which one allele had dropped out, despite the high average RFUs ($H_{ir.}$ = 468 RFUs, when using either detection threshold). The high average peak heights led to a low $P(\hat{D}_O)$ ($<10^{-8}$) which then produced a lower LR because one instance of drop-out was required under the hypothesis that this individual was the contributor.

3.6. Effect of analytical threshold on LRs using estimated and benchmark drop-out probabilities

As reported earlier in this paper, lower allelic detection threshold results in fewer instances of apparent drop-out. To evaluate what effect this has on the LRs, we compared the same LRs as previously calculated using a 50 RFU detection threshold to those using a 30 RFU detection threshold. As expected, when the suspected contributor was the true contributor, a 30 RFU detection threshold produces LRs that are higher than those obtained using a 50 RFU detection threshold (Fig. 5). This trend is reflected by the excess of points above the diagonal in Fig. 5. The large differences were particularly apparent for the comparisons with LRs $< 10^{15}$ using the 50 RFU detection threshold, where drop-out tended to occur more frequently. As such, lowering the allelic detection threshold results in detecting more (real) alleles, thereby providing additional information, and thus strengthening the weight of evidence in favor of the true contributor.

3.7. Effect of known non-contributors using estimated and benchmark drop-out probabilities

We next computed LRs where each of the LT DNA profiles was compared to profiles from random individuals whose genotypes were simulated in proportion to the US Caucasian allele frequencies. Specifically, in these LRs, the hypothesis in the numerator was that the simulated reference profile individual was the true contributor. The hypothesis in the denominator was that an unknown individual was the true contributor. These LRs were computed using the estimated drop-out probability ($P(\hat{D}_O)$) and the benchmark drop-out probability ($P(D_O)$). Because none of these simulated individuals were the actual donors of the LT DNA profiles, the majority of the LRs should be less than one. This analysis allows us to evaluate whether estimating the drop-out probability using this logistic regression approach leads to unexpectedly large LRs which would falsely support individuals who were not the true donors.

For the majority of the comparisons (>80%), the LRs were <1, regardless of detection threshold, typing system, or how $P(\hat{D}_O)$ was estimated (Tables 3 and 4). These LRs correctly reflected support for the hypothesis that the known non-contributor was not the true donor. However, the 30 RFU detection threshold led to a greater number (about 8%) of LRs < 1 than those computed using the 50 RFU detection threshold. The 30 RFU detection threshold allows the detection of additional alleles, some of which were exclusionary, supporting the hypothesis that an unknown individual was the contributor. In nearly all of the comparisons (>99%), the LRs were <2, indicating that, in most instances, the LR approach correctly rejects the proposition of random individuals as true donors of the evidence. Further, using our estimates of $P(\hat{D}_O)$ from the logistic regression calculation results in fewer instances of large LRs compared to using the benchmark drop-out probability ($P(D_O)$). For example, for the 50 RFU detection threshold in the Identifiler® data, about 10 times fewer LRs > 1000 were obtained using the estimated drop-out probability compared with the benchmark drop-out probability. Also, the maximum LRs obtained using $P(\hat{D}_O)$ were lower than those obtained using $P(D_O)$. Finally, the distribution of the LRs for the Identifiler® profiles calculated using $P(\hat{D}_O)$ estimated from the Identifiler® data was nearly identical to the distribution of the LRs for the Identifiler® profiles using $P(\hat{D}_O)$ estimated from the PowerPlex® 16 data (Table 3) and

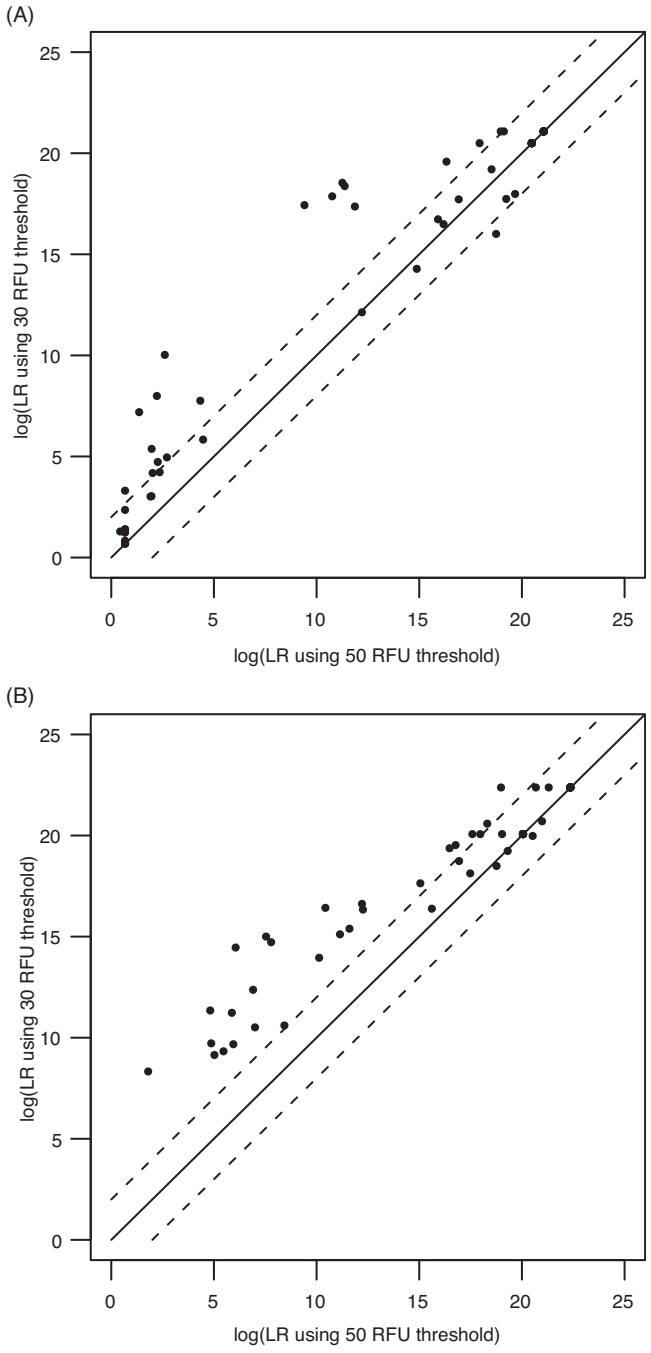

Fig. 5. Comparison between the LRs calculated using the 30 RFU detection threshold vs. the 50 RFU detection threshold when the hypothesized contributor was the true donor of the LT DNA profile. The LRs were calculated using the estimated drop-out probability $P(\hat{D}_O)$ for each sample. (A) Identifiler profiles®. (B) PowerPlex® 16 profiles.

Table 3
Distribution of the LRs calculated comparing 60 LT Identifiler® profiles to 5000 simulated European individuals.[a]

Detection threshold (RFU)	Drop-out probability[b]	% LRs <1	% LRs <2	% LRs >100	%LRs >1000	99.9% of LRs are < than...	Max LR
50	Benchmark $P(D_O)$	84.2	99.4	0.053	0.010	34	6.6×10^4
	Estimated using Identifiler	84.3	99.5	0.031	0.001	27	2.4×10^3
	Estimated using PPX	84.2	99.5	0.030	0.001	27	2.4×10^3
30	Benchmark $P(D_O)$	92.9	99.6	0.041	0.005	31	1.5×10^6
	Estimated using Identifiler	92.8	99.6	0.012	0.001	17	3.2×10^5
	Estimated using PPX	92.8	99.6	0.012	0.001	18	3.4×10^5

[a] LT profiles containing 0 alleles were left in this analysis, but the LRs were set to 1.
[b] Denotes the approach used to estimate the drop-out probability. The benchmark $P(D_O)$ is the best drop-out probability for each locus at each profile. The benchmark $P(D_O)$ was 0, if 2 alleles were detected at a locus. The benchmark $P(D_O)$ was 0.5 if a single allele was detected at a locus. The benchmark $P(D_O)$ was 1 if 0 alleles were detected at a locus. Estimated using Identifiler® refers to drop-out probabilities estimates using the logistic regression equations fit to the Identifiler data from the 30 samples from the other individual. Estimated using PowerPlex® 16 refers to drop-out probabilities estimates using the logistic regression equations fit to the PowerPlex® 16 data from the 30 samples from the other individual.

vice versa (Table 4). This finding suggests that the differences between the estimates of $P(\hat{D}_O)$ obtained from these two typing systems are negligible with regard to their effect on the final LR.

We next focused on those comparisons between random individuals and the LT DNA profiles that gave LRs > 1 using $P(\hat{D}_O)$. We wanted to determine whether these relatively high LRs resulted from using an estimated drop-out probability, or were instead inherent features of the profiles and the LR model. Fig. 6 shows the LRs calculated using $P(D_O)$ vs. the LRs calculated using $P(\hat{D}_O)$. Points below the diagonal represent those comparisons for which the LRs calculated using $P(\hat{D}_O)$ were larger than those calculated using $P(D_O)$. Conversely, points above the diagonal represent those comparisons for which the LRs calculated using $P(\hat{D}_O)$ were smaller than those using $P(D_O)$. Additionally, points that fall below the dotted line represent those comparisons where the LRs using $P(D_O)$ are <1, but the LRs calculated using the $P(\hat{D}_O)$ were >1. We first note that few points fall below the dotted line and those that do are <2 (using a $\log_{10}$ scale) on the *x*-axis. This means that when the LRs calculated using $P(D_O)$ are <1, they are always <100 using $P(\hat{D}_O)$. Similarly, for those profiles for which using $P(\hat{D}_O)$ gives a higher LR than using $P(D_O)$, the LR using $P(\hat{D}_O)$ is almost always <100. Those points that fall above the diagonal are scattered throughout the entire range. These results indicate that in the relatively few instances for which $P(\hat{D}_O)$ gives an estimate of the LR that overstates the strength of the evidence, the absolute LR will still be small (<100). Finally, LRs computed using the 30 RFU threshold resulted in fewer points below the dotted line than the 50 RFU data. This finding indicates that, for the lower detection threshold, fewer instances were seen for which the LR calculated using $P(D_O)$ was <1 and the LR calculated using $P(\hat{D}_O)$ was >1 than when using the higher (50 RFU) threshold. Similar patterns hold for the PowerPlex® 16 typing system (Fig. S3). This result supports using data-driven thresholds that rescue low-level, but informative, data.

4. Discussion

Here we have shown that the relationship between the average peak height and the proportion of allelic drop-out in LT DNA samples can be modeled using logistic regression. We have presented the parameters of the logistic model that can be used to estimate drop-out probabilities as a function of average peak height for 30 RFU and 50 RFU detection thresholds.

We have evaluated the difference between using a standard, policy-driven allelic detection threshold of 50 RFUs and using a data-derived (or, as previously used throughout the manuscript, data-driven) threshold of 30 RFUs [19,30,31,33–35]. Not unexpectedly, we found fewer instances of apparent drop-out when using the lower detection threshold. This finding indicates that useful genetic information, both inclusionary and exclusionary, is discarded by ignoring alleles that fall below a conventional 50 RFU threshold.

One obvious question is whether lowering the detection threshold to 30 RFUs resulted in the detection of additional noise not found with a higher threshold? After removing obvious artifacts, a number of peaks that were not represented in the known input profiles were detected. All of these peaks, except for one, were in plus or minus stutter positions. In one instance, a D13 11 peak in the minus stutter position that was detected at both the 30 RFU and 50 RFU analytical thresholds was called because, at 9% of the parent peak, it just exceeded the manufacturer's default stutter filter of 8%. In another instance, a D16 12 peak that was also detected at both the 30 RFU and 50 RFU analytical thresholds was located between two parent peaks such that it comprised contributions from both plus and minus stutter; thus the standard D16 minus stutter threshold of 10.4% would not necessarily filter the peak appropriately. Several additional instances of peaks in a plus stutter position rising above the analytical threshold were noted; one instance was detected at both the 50 RFU and 30 RFU

Table 4
Distribution of the LRs calculated comparing 60 LT PowerPlex® 16 profiles to 5000 simulated European individuals.[a]

Detection threshold (RFU)	Drop-out probability[b]	% LRs < 1	% LRs < 2	% LRs > 100	% LRs > 1000	99.9% of LRs are < than...	Max LR
50	Benchmark $P(D_O)$	99.98	99.98	0.005	0.001	0.001	2.7×10^3
	Estimated using PPX	99.98	99.99	0.004	0.0	0.038	30
	Estimated using Identifiler	99.98	99.99	0.004	0.0	0.029	490
30	Benchmark $P(D_O)$	100	100	0	0	0	6.6×10^{-7}
	Estimated using PPX	100	100	0	0	2.8×10^{-13}	1.4×10^{-2}
	Estimated using Identifiler	100	100	0	0	4.9×10^{-13}	7.3×10^{-2}

[a] All 60 LT DNA profiles had at least one detected allele.
[b] Denotes the approach used to estimate the drop-out probability. The benchmark $P(D_O)$ is the best drop-out probability for each locus at each profile. The benchmark $P(D_O)$ was 0, if 2 alleles were detected at a locus. The benchmark $P(D_O)$ was 0.5 if a single allele was detected at a locus. The benchmark $P(D_O)$ was 1 if 0 alleles were detected at a locus. Estimated using PowerPlex® 16 refers to drop-out probabilities estimates using the logistic regression equations fit to the PowerPlex® 16 data from the 30 samples from the other individual. Estimated using Identifiler® refers to drop-out probabilities estimates using the logistic regression equations fit to the Identifiler® data from the 30 samples from the other individual.

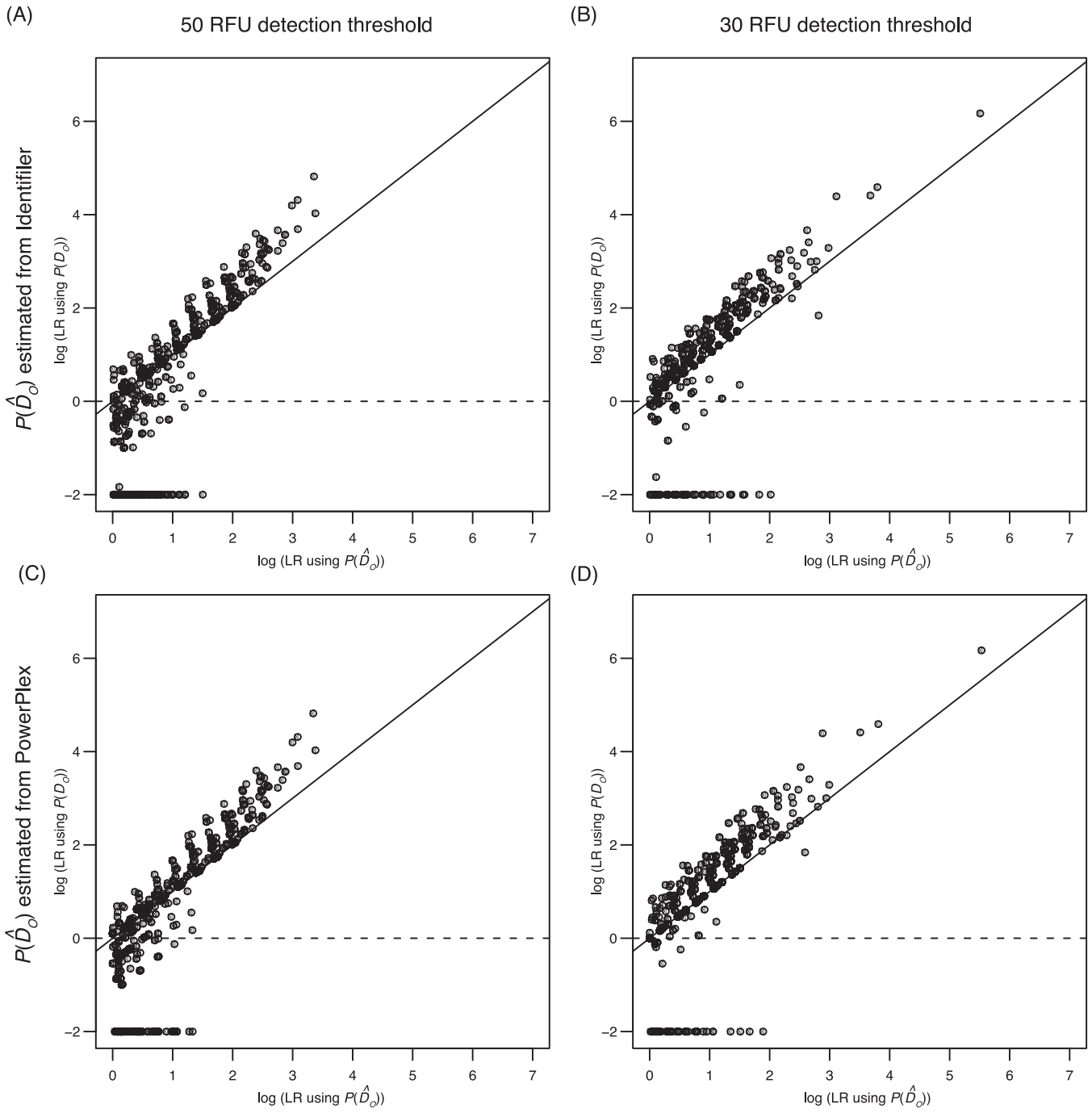

Fig. 6. LRs calculated from the Balding and Buckleton program when the hypothesized contributor was a random individual who was not the true donor of the Identifiler® LT DNA profile. Each dot represents the log(LR) for a particular LT DNA profile calculated using the benchmark drop-out probability ($P(D_O)$, *y*-axis) vs. the estimated drop-out probability ($P(\hat{D}_O)$, *x*-axis). Only comparisons where LR > 1 using the estimated drop-out probability, $P(\hat{D}_O)$, are shown. The solid line is the diagonal. *Note*: for plotting purposes, LRs calculated using $P(D_O)$ that were <0.01 were set to 0.01, without changing the LR calculated using $P(\hat{D}_O)$. (A) 50 RFU detection threshold, where $P(\hat{D}_O)$ was estimated from the other individual's Identifiler® data. (B) 30 RFU detection threshold, where $P(\hat{D}_O)$ was estimated from the other individual's Identifiler data. (C) 50 RFU detection threshold, where $P(\hat{D}_O)$ was estimated from the other individual's PowerPlex® 16 data. (D) 30 RFU detection threshold, where $P(\hat{D}_O)$ was estimated from the other individual's PowerPlex® 16 data.

analytical threshold, and three instances, two involving the same allele, at only the 30 RFU threshold. We did not designate these peaks as drop-in. Although some workers choose to combine these peaks in stutter positions with true drop-in [18], we advocate that the better approach is to reinvestigate both peak height ratios (PHRs), and plus and minus stutter percentages specifically for LT samples in each specific system. Application of empirical PHRs and stutter thresholds should minimize the risk of overestimating the incidence of true drop-in for low-template samples.

At D3, we detected a 17 when using the 30 RFU threshold (31 RFUs) which could not conclusively be determined to be artifactual. We compared this profile to the known contributors and known non-contributors both including and excluding this peak. Because two alleles were detected at every locus in this profile, the extra peak made little difference in the final LRs (Table S4).

We also corroborated the findings previously observed for degraded DNA samples [27,44,45] that amplicon length was positively correlated with the number of alleles that had dropped

out at a particular locus. Because the LT samples analyzed here were of good quality, the preferential drop-out of longer alleles must be explained by some mechanism other than degradation. Nevertheless, this work demonstrates that, while special circumstances might require the use of locus-specific drop-out probabilities, in general, LRs calculated using averaged drop-out probabilities accurately assess the weight of the evidence.

Next, we have shown that, when estimated drop-out probabilities were used to compute LRs, these LRs were very similar to the LRs calculated using the benchmark drop-out probability. The concordance between LRs calculated from estimated drop-out probabilities and the LRs calculated using the benchmark drop-out probabilities may be surprising in light of several statistically significant differences (Tables S1–S3) in the logistic regression curves shown in Fig. 2. However, these two findings are not incompatible with each other. Our simulations show that the magnitude of the differences did not substantively change the ability to compute accurate estimates of the drop-out probability and, consequently, accurate LRs.

Similarly, the drop-out probabilities estimated for Identifiler® profiles using the logistic regression equation fit to the PowerPlex® 16 data gave LRs very similar to those obtained using the benchmark drop-out probability. The same pattern was observed for the opposite situation. These results are somewhat surprising given that the two different kits employ different fluorescent dyes, a different number of PCR cycles, different PCR primers, and different reaction conditions [46,47]. However, it is also comforting to confirm that the *in vitro* mechanics apparently proceed in a very similar fashion for both systems. The fact that the estimates of drop-out probabilities could be used interchangeably across the typing systems does not mean that no differences exist in the characteristics of drop-out between the two systems. Rather, it means that any differences between the two systems are relatively minor compared to the overall variability inherent in LT DNA evidence. Further, the fact that the logistic regression equations could be used interchangeably across systems speaks to the robustness of the estimates. An estimate of the drop-out probability for a particular sample is sufficiently accurate, even if the conditions used for the analysis of that sample are not exactly the same as those used to estimate the parameters of the logistic regression model.

While the interchangeability between the PowerPlex® 16 and Identifiler® data is encouraging, it is important not to over-interpret this result. Other sources of variability exist in actual forensic casework that were not included in our study. For example, all of the samples in this study were typed in the same laboratory and were analyzed on the same ABI 3130 genetic analyzer. The range of inter-laboratory and inter-instrumental variability was not tested here.

Another consideration is how accurately drop-out probabilities estimated from LT single-source samples model the drop-out probabilities of LT DNA mixtures. One line of existing evidence suggests that drop-out probabilities estimated from single-source samples may be applicable to mixtures. Bright et al. [48] found little difference between peak height ratios computed from single-source samples and in mixtures with known mixture proportions. We have observed a similar trend in reviewing laboratory validation data received in connection with casework (unpublished data). If this pattern holds for allelic drop-out, the estimation of drop-out probabilities would be simplified.

Another potential concern is whether the logistic regression curve modeling the relationship between drop-out and average RFUs would hold for samples with different concentrations of DNA than those used in this study (i.e. 10 pg, 30 pg, and 100 pg.) However, the relationship between the quantity of DNA in the sample and drop-out are highly correlated with the average peak-heights [26]. Given this high correlation, the concentration should be represented in the average peak heights. However, if this correlation is not perfect, then sampling a wider number of DNA concentrations may give more accurate models.

Although we demonstrated that allele length was positively correlated with the number of alleles that had dropped-out at a particular locus, we found that using an average drop-out probability for all loci nonetheless yielded accurate estimates of the LRs. Thus, for the data analyzed here, the locus-specific differences in drop-out were subtle and did not appreciably affect the outcome when calculating LRs that included drop-out probabilities. While locus-specific drop-out probabilities were not necessary for this data set, they might be required for degraded samples, as suggested by Tvedebrink et al. [27], because degradation preferentially affects the larger loci. However, Bright et al. [49] found little difference between heterozygote peak height ratios in pristine validation samples and forensic casework samples, suggesting that for some attributes of the DNA profiles, pristine samples are an adequate proxy for actual casework. An additional line of evidence suggests caution when applying parameter rich models, such as those including locus-specific estimates of the logistic regression parameters to model drop-out in forensic applications. Specifically, a recent study [50] suggests that locus-specific logistic regression parameters may differ across different datasets, decreasing their utility in casework situations. However, a better strategy could be to account for locus-specific effects by adjusting the peak heights as a function of the molecular weight, as suggested by Buckleton et al. [49].

Finally, we have evaluated the difference between using a standard, policy-driven allelic detection threshold of 50 RFUs and using a data-derived LOD/LOQ threshold of 30 RFUs [19,30–35]. The alleles that were included using a 30 RFU detection threshold, but were discarded using the 50 RFU detection threshold, hold substantial value. When comparing the true contributor to the LT profile, including such alleles can lead to LRs which are orders of magnitude higher than when excluding true alleles between 30 RFUs and 50 RFUs. Similarly, for our data here, the lower detection threshold resulted in LRs < 1 for an additional 10% of random individuals. This demonstrates the potential risk of ignoring data based on a policy-driven detection threshold. However, even for the rare instance in which a LR fell above 1 for a true non-contributor, the LR using the Identifiler® data tended to be higher for the 30 RFU detection threshold than for the 50 RFU detection threshold. The reason for this is that incorporation of the additional alleles more strongly supports either the hypothesis that the individual was not a contributor (pushing the LR < 1) or, conversely, that the individual was a contributor (pushing the LR > 1). This increase in accuracy, in particular, is critically important, and would strongly refute the prevailing misunderstanding that ignoring data is always "conservative for the defendant" [20,34,35].

Our results indicate that it is possible to accurately compute LRs using drop-out probabilities estimated from empirical data. This contributes to the wider application of probabilistic approaches to forensic casework using LRs that incorporate an estimate of the drop-out probability. Combining the appropriate detection threshold with LRs that incorporate empirically estimated drop-out probabilities should provide a sensitive and statistically rigorous framework to reliably interpret LT DNA evidence.

Acknowledgements

We thank John Butler and his lab for generating the data used in this study and for providing it to us. KEL was supported by funds from the Miller Research Institute at UC Berkeley and start-up funds from UCLA.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.fsigen.2014.04.003.

References

[1] P. Taberlet, S. Griffin, B. Goossens, S. Questiau, V. Manceau, N. Escaravage, et al., Reliable genotyping of samples with very low DNA quantities using PCR, Nucleic Acids Res. 24 (1996) 3189–3194.
[2] J. Buckleton, P. Gill, Low copy number, in: J. Buckleton, C.M. Triggs, S.J. Walsh (Eds.), Forensic DNA Evidence Interpretation, CRC, Press, Boa Raton, FL, 2005, pp. 275–297.
[3] P. Gill, C.H. Brenner, J.S. Buckleton, A. Carracedo, M. Krawczak, W.R. Mayr, et al., DNA commission of the International Society of Forensic Genetics: recommendations on the interpretation of mixtures, Forensic Sci. Int. 160 (2006) 90–101.
[4] D.J. Balding, J. Buckleton, Interpreting low template DNA profiles, Forensic Sci. Int. Genet. 4 (2009) 1–10.
[5] J. Buckleton, Validation issues around DNA typing of low level DNA, Forensic Sci. Int. Genet. 3 (2009) 255–260.
[6] B. Budowle, A.J. Eisenberg, A. van Daal, Validity of low copy number typing and applications to forensic science, Croat. Med. J. 50 (2009) 207–217.
[7] N. Gilbert, Science in court: DNA's identity crisis, Nature 464 (2010) 347–348.
[8] P. Gill, J. Buckleton, Commentary on: Budowle B, Onorato AJ, Callaghan TF, Della Manna A, Gross AM, Guerrieri RA, Luttman JC, McClure DL. Mixture interpretation: defining the relevant features for guidelines for the assessment of mixed DNA profiles in forensic casework. J Forensic Sci 2009;54(4):810–21, J. Forensic Sci. 55 (2010) 265–268, author reply 269–72.
[9] B. Budowle, A. van Daal, Comment on "A universal strategy to interpret DNA profiles that does not require a definition of low copy number" by Peter Gill and John Buckleton, 2010, Forensic Sci. Int. Genetics 4, 221–227, Forensic Sci. Int. Genet. 5 (2011) 15.
[10] M. Kayser, P. de Knijff, Improving human forensics through advances in genetics, genomics and molecular biology, Nat. Rev. Genet. 12 (2011) 179–192.
[11] 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.
[12] P. Gill, Application of low copy number DNA profiling, Croat. Med. J. 42 (2001) 229–232.
[13] B. Budowle, A.J. Onorato, T.F. Callaghan, A. Della Manna, A.M. Gross, R.A. Guerrieri, et al., Mixture interpretation: defining the relevant features for guidelines for the assessment of mixed DNA profiles in forensic casework, J. Forensic Sci. 54 (2009) 810–821.
[14] P. Gill, J. Buckleton, A universal strategy to interpret DNA profiles that does not require a definition of low-copy-number, Forensic Sci. Int. Genet. 4 (2010) 221–227.
[15] J. Buckleton, P. Gill, Further comment on "Low copy number typing has yet to achieve general acceptance" by Budowle, B. et al., 2009. Forensic Sci. Int. Genetics: Supplement Series 2, 551–552, Forensic Sci. Int. Genet. 5 (2011) 7–11, discussion 12–14.
[16] P. Gill, L. Gusmao, H. Haned, W.R. Mayr, N. Morling, W. Parson, et al., DNA commission of the International Society of Forensic Genetics: recommendations on the evaluation of STR typing results that may include drop-out and/or drop-in using probabilistic methods, Forensic Sci. Int. Genet. 6 (2012) 679–688.
[17] H. Haned, K. Slooten, P. Gill, Exploratory data analysis for the interpretation of low template DNA mixtures, Forensic Sci. Int. Genet. 6 (2012) 762–774.
[18] A.A. Mitchell, J. Tamariz, K. O'Connell, N. Ducasse, Z. Budimlija, M. Prinz, et al., Validation of a DNA mixture statistics tool incorporating allelic drop-out and drop-in, Forensic Sci. Int. Genet. 6 (2012) 749–761.
[19] Scientific Working Group on DNA Analysis Methods, SWGDAM Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories, 2010 Available from: http://www.fbi.gov/about-us/lab/codis/swgdam.pdf.
[20] K.E. Lohmueller, N. Rudin, Calculating the weight of evidence in low-template forensic DNA casework, J. Forensic Sci. 58 (2013) S243–S249.
[21] J.M. Curran, J. Buckleton, Inclusion probabilities and dropout, J. Forensic Sci. 55 (2010) 1171–1173.
[22] P. Gill, A. Kirkham, J. Curran, LoComatioN: a software tool for the analysis of low copy number DNA profiles, Forensic Sci. Int. 166 (2007) 128–138.
[23] P. Gill, J. Curran, C. Neumann, A. Kirkham, T. Clayton, J. Whitaker, et al., Interpretation of complex DNA profiles using empirical models and a method to measure their robustness, Forensic Sci. Int. Genet. 2 (2008) 91–103.
[24] D.J. Balding, Evaluation of mixed-source, low-template DNA profiles in forensic science, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 12241–12246.
[25] K. Cheng, L. Inman-Semerau, N. Rudin, K. Inman, K.E. Lohmueller, Lab Retriever, 2013 www.scieg.org/Lab_retriever.html.
[26] T. Tvedebrink, P.S. Eriksen, H.S. Mogensen, N. Morling, Estimating the probability of allelic drop-out of STR alleles in forensic genetics, Forensic Sci. Int. Genet. 3 (2009) 222–226.
[27] T. Tvedebrink, P.S. Eriksen, H.S. Mogensen, N. Morling, Statistical model for degraded DNA samples and adjusted probabilities for allelic drop-out, Forensic Sci. Int. Genet. 6 (1) (2012) 97–101.
[28] T. Tvedebrink, P.S. Eriksen, M. Asplund, H.S. Mogensen, N. Morling, Allelic drop-out probabilities estimated by logistic regression – further considerations and practical implementation, Forensic Sci. Int. Genet. 6 (2012) 263–267.
[29] J.M. Butler, C.R. Hill, Scientific issues with analysis of low amounts of DNA, Profiles in DNA, 13, 2010.
[30] D.J. Anderson, Determination of the lower limit of detection, Clin. Chem. 35 (1989) 2152–2153.
[31] D.A. Armbruster, M.D. Tillman, L.M. Hubbs, Limit of detection (LQD)/limit of quantitation (LOQ): comparison of the empirical and the statistical methods exemplified with GC–MS assays of abused drugs, Clin. Chem. 40 (1994) 1233–1238.
[32] K.A. Rubinson, J.F. Rubinson, Sample Size and Major, Minor, Trace, and Ultratrace Components, Contemporary Instrumental Analysis, Prentice Hall, Upper Saddle River, NJ, 2000, pp. 150–158.
[33] V. Thomsen, D. Schatzlein, D. Mercuro, Limits of detection in spectroscopy, Spectroscopy 18 (2003) 112–114.
[34] J.R. Gilder, T.E. Doom, K. Inman, D.E. Krane, Run-specific limits of detection and quantitation for STR-based DNA testing, J. Forensic Sci. 52 (2007) 97–101.
[35] C.A. Rakay, J. Bregu, C.M. Grgicak, Maximizing allele detection: effects of analytical threshold and DNA levels on rates of allele and locus drop-out, Forensic Sci. Int. Genet. 6 (2012) 723–728.
[36] J. Bregu, D. Conklin, E. Coronado, M. Terrill, R.W. Cotton, C.M. Grgicak, Analytical thresholds and sensitivity: establishing RFU thresholds for forensic DNA analysis, J. Forensic Sci. 58 (2013) 120–129.
[37] R Development Core Team, R: A Language and Environment for Statistical Computing, 2010.
[38] J.M. Butler, STRbase, 2012, http://www.cstl.nist.gov/strbase/multiplx.htm.
[39] I.W. Evett, J. Scranage, R. Pinchin, An illustration of the advantages of efficient statistical methods for RFLP analysis in forensic science, Am. J. Hum. Genet. 52 (1993) 498–505.
[40] I.W. Evett, P.D. Gill, J.K. Scrange, B.S. Weir, Establishing the robustness of short-tandem-repeat statistics for forensic applications, Am. J. Hum. Genet. 58 (1996) 398–407.
[41] I.W. Evett, B.S. Weir, Interpreting DNA Evidence: Statistical Genetics for Forensic Scientists, Sunderland, MA, Sinauer Associates, 1998.
[42] J.M. Butler, R. Schoske, P.M. Vallone, J.W. Redman, M.C. Kline, Allele frequencies for 15 autosomal STR loci on US Caucasian, African American, and Hispanic populations, J. Forensic Sci. 48 (2003) 908–911.
[43] E.N. Levadokou, D.A. Freeman, M.J. Budzynski, B.E. Early, K.C. McElfresh, J.W. Schumm, et al., Allele frequencies for fourteen STR loci of the PowerPlex 1.1 and 2.1 multiplex systems and Penta D locus in Caucasians, African-Americans, Hispanics, and other populations of the United States of America and Brazil, J. Forensic Sci. 46 (2001) 736–761.
[44] P.M. Schneider, K. Bender, W.R. Mayr, W. Parson, B. Hoste, R. Decorte, et al., STR analysis of artificially degraded DNA – results of a collaborative European exercise, Forensic Sci. Int. 139 (2004) 123–134.
[45] L.A. Dixon, A.E. Dobbins, H.K. Pulker, J.M. Butler, P.M. Vallone, M.D. Coble, et al., Analysis of artificially degraded DNA using STRs and SNPs – results of a collaborative European (EDNAP) exercise, Forensic Sci. Int. 164 (2006) 33–44.
[46] B.E. Krenke, A. Tereba, S.J. Anderson, E. Buel, S. Culhane, C.J. Finis, et al., Validation of a 16-locus fluorescent multiplex system, J. Forensic Sci. 47 (2002) 773–785.
[47] P.J. Collins, L.K. Hennessy, C.S. Leibelt, R.K. Roby, D.J. Reeder, P.A. Foxall, Developmental validation of a single-tube amplification of the 13 CODIS STR loci, D2S1338, D19S433, and amelogenin: the AmpFlSTR Identifiler PCR Amplification Kit, J. Forensic Sci. 49 (2004) 1265–1277.
[48] J.A. Bright, J. Turkington, J. Buckleton, Examination of the variability in mixed DNA profile parameters for the Identifiler multiplex, Forensic Sci. Int. Genet. 4 (2010) 111–114.
[49] J.A. Bright, K. McManus, S. Harbison, P. Gill, J. Buckleton, A comparison of stochastic variation in mixed and unmixed casework and synthetic samples, Forensic Sci. Int. Genet. 6 (2012) 180–184.
[50] J. Buckleton, J. Kelly, J.A. Bright, D. Taylor, T. Tvedebrink, J.M. Curran, Utilising allelic dropout probabilities estimated by logistic regression in casework, Forensic Sci. Int. Genet. 9 (2014) 9–11.