View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Flinders Academic Commons

Archived at the Flinders Academic Commons:
http://dspace.flinders.edu.au/dspace/

'This is the peer reviewed version of the following article:
Taylor, D., Bright, J.-A., Kelly, H., Lin, M.-H., & Buckleton, J. (2017). A fully continuous system of DNA profile evidence evaluation that can utilise STR profile data produced under different conditions within a single analysis. Forensic Science International: Genetics, 31, 149–154. https://doi.org/10.1016/j.fsigen.2017.09.002

which has been published in final form at
https://doi.org/10.1016/j.fsigen.2017.09.002

© 2017 Elsevier B.V. This manuscript version is made available under the CC-BY-NC-ND 4.0 license:
http://creativecommons.org/licenses/by-nc-nd/4.0/

Accepted Manuscript

Title: A fully continuous system of DNA profile evidence evaluation that can utilise STR profile data produced under different conditions within a single analysis

Authors: Duncan Taylor, Jo-Anne Bright, Hannah Kelly, Meng-Han Lin, John Buckleton

PII: S1872-4973(17)30188-6
DOI: http://dx.doi.org/10.1016/j.fsigen.2017.09.002
Reference: FSIGEN 1776

To appear in: *Forensic Science International: Genetics*

Received date: 8-5-2017
Revised date: 11-8-2017
Accepted date: 6-9-2017

Please cite this article as: Duncan Taylor, Jo-Anne Bright, Hannah Kelly, Meng-Han Lin, John Buckleton, A fully continuous system of DNA profile evidence evaluation that can utilise STR profile data produced under different conditions within a single analysis, Forensic Science International: Geneticshttp://dx.doi.org/10.1016/j.fsigen.2017.09.002

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

TITLE:

A fully continuous system of DNA profile evidence evaluation that can utilise STR profile data produced under different conditions within a single analysis

AUTHOR:

Duncan Taylor[1,2,*], Jo-Anne Bright[3], Hannah Kelly[3], Meng-Han Lin[3], and John Buckleton[3,4].

1. Forensic Science SA, 21 Divett Place, Adelaide, SA 5000, Australia
2. School of Biological Sciences, Flinders University, GPO Box 2100 Adelaide SA, Australia 5001
3. Institute of Environmental Science and Research Limited, Private Bag 92021, Auckland, 1142 New Zealand
4. University of Washington, Department of Biostatistics, Seattle, WA 98195, United States

CORRESPONDING AUTHOR:

Duncan A. Taylor, PhD

Forensic Science SA

21 Divett Place

Adelaide SA

Australia 5000

Tel: +61-8 8226 7700

Fax: +61-8 8226 7777

Email: Duncan.Taylor@sa.gov.au

HIGHLIGHTS

- We explain several casework situations that can arise which require evaluation of DNA profile replicates that span a technology change
- When a likelihood ratio is required for these types of evaluations, the classic approach has been to pick one of the replicates
- We present the maths for the interpretation system STRmix™ to allow the combination of replicates produced under different conditions
- We demonstrate the performance of the proposed model on situations which may be encountered in the laboratory

ACCEPTED MANUSCRIPT

ABSTRACT

The introduction of probabilistic DNA interpretation systems has made it possible to evaluate many profiles that previously (under a manual interpretation system) were not. These probabilistic systems have been around for a number of years and it is becoming more common that their use within a laboratory has spanned at least one technology change. This may be a change in laboratory hardware, the DNA profiling kit used, or the manner in which the profile is generated. Up until this point, when replicates DNA profiles are generated, that span a technological change, the ability to utilise all the information in all replicates has been limited or non-existent. In this work we explain and derive the models required to evaluate (what we term) multi-kit analysis problems. We demonstrate the use of the multi-kit feature on a number of scenarios where such an analysis would be desired within a laboratory. Allowing the combination of profiling data that spans a technological change will further increase the amount of DNA profile information produced in a laboratory that can be evaluated.

KEYWORDS: Multiple DNA profiling kits; Continuous DNA interpretation; STRmix; modelling; DNA mixtures.

1.0 INTRODUCTION

DNA profile interpretation has substantially changed towards probabilistic genotyping. Testament to this are the number of probabilistic DNA profile interpretation systems now available [1-6], the recommendations that have been released by multiple advisory bodies on their validation [7, 8], and the vast array of published works on various models that can be used to describe different elements of DNA profile generation or behaviour [9-26].

Most of the available systems have features that carry out two standard tasks:

1) Deconvolution of a DNA profile into sets of genotypes that could describe the profile, weighted by some system of probabilistic modelling (whether this be in a semi-continuous manner that uses probabilities of drop-in and drop-out or a fully continuous system that applies probabilities for various mass parameters within the model).
2) The calculation of a likelihood ratio (*LR*) when comparing one or more references to the deconvolution and considering two competing propositions.

ACCEPTED MANUSCRIPT

These two tasks fulfil the vast majority of functions required by most forensic laboratories. However, as our knowledge and comfort with these systems increases, extensions to the above elements are starting to be explored. These have included extensions to the calculation of the *LR* to database searching [27], familial searching [28], the calculation of an *LR* comparing two mixed profiles where the propositions are whether one or more common contributors exist [29], or seeking information on potential familial relationships between donors of a mixture [30]. The deconvolution has also been extended to carry out deconvolutions looking for one or more common contributors to multiple evidence items [31].

The use of multiple PCR replicates within a single interpretation has been undertaken for as long as probabilistic genotyping has been considered [32]. For semi-continuous systems this involves the accumulation of probabilities of drop-in and drop-out for each peak in each profile in the analysis. For fully continuous systems the use of replicates involves the comparison of multiple sets of observed peak heights to their expectations, built up from the models in the system.

We present here an extension to this idea of using replicates, by considering replicates that have been produced under different conditions. This would be of particular use under two broad conditions:

1) The forensic laboratory uses complementary DNA profiling kits that have the same loci, but ordered differently with respect to molecular weight i.e. as per the ESI and ESX profiling system.
2) Data has been produced in one system that has then been superseded by another before the replicate was produced. This may be due to a change in technology such as the move of a laboratory to a different profiling kit, or an update in the model of electrophoretic instrumentation, or the decision to trial an enhancement technique, such as increased PCR cycles, when standard profiling has produced insufficient data.

We term these replicates a ‘multi-kit’ analysis noting that it may be the same DNA profiling kit in use, just being used under two conditions (for example, a replicate PCR amplification after having moved from a 3130 capillary electrophoresis instrument to a 3500). The difference in the consideration of a multi-kit analysis to a standard replicate analysis is that the profiles have been produced under differing conditions, meaning that they are behaving

ACCEPTED MANUSCRIPT

differently from one another. While the same general models should apply (e.g. for stutter, peak height variability or drop-in, etc.) the parameter values (or their prior probabilities) within the models will differ between replicates. Multi-kit analyses pose additional challenges to standard continuous DNA interpretation systems as different parameters within the models may be correlated across replicates while others will not. An example of this would be to consider a mixed and degraded DNA sample that is profiled in two different kits. We would expect the relative DNA amounts of each contributor to be the same between replicates (they in fact must be as this is a property of the DNA extract, which precedes DNA profiling) where the amplification efficiencies of the loci would not be (as this is a kit related parameter).

We outline here an extension to the theory presented in [1], that considers multi-kit analyses. We then demonstrate the performance of the proposed model on a number of scenarios that laboratories could encounter, where multi-kit analysis capabilities would prove beneficial.

2.0 THEORY

Starting with the standard theory of [1, 32] we seek the probability of the observed evidence given a proposition. Nuisance parameters must be considered in order to enumerate the probability and we term these the mass parameters (***M***). We do not specifically care what the values of these parameters are to calculate the probability (hence 'nuisance'), although it may be of interest to know their values for either diagnostic purposes or for comments on higher level evidential considerations that rely on extrinsic properties of the DNA profile such as the biological source of the DNA or the activities that could have led to the presence of the DNA. We also do not know the values of these mass parameters and so integrate across their prior probability distributions, which can be achieved through the use of Markov chain Monte Carlo (MCMC) analysis. The mass parameters considered in [1] are template DNA (t), degradation (d), locus amplification efficiency (A) and replicate amplification efficiency (R). To extend the theory to multiple kits we introduce a fifth mass parameter, a kit multiplier (B). We now have:

1. A variable t_n, for each of the n contributors that may usefully be thought of as template amount,
2. A variable $d_{n,k}$, which models the decay with respect to molecular weight (m) in the template for each of the contributors to genotype set, S_j, where j is one of the J

ACCEPTED MANUSCRIPT

possible genotype sets the contributors can take. This may usefully be thought of as a measure of degradation. Note that this term is now dependent on kit, *k*, although this dependence on kit is optional and can be fixed within the analysis so that $d_{n,k} = d_{n,k'}$ under certain circumstance (explained later),

3. Locus efficiencies at each locus, A_k^l to allow for the observed amplification levels of each locus. Note that this term is now dependent on kit,
4. Replicate multipliers $R_{y,k}$. This effectively scales all peaks up or down between replicates. Note that this term is now dependent on kit, and
5. Kit multipliers, B_k. This scales all peaks in all replicates up or down between kits.

We write the mass variables, $\mathbf{D} = \{d_{n,k} : n = 1,...,N : k = 1,...,K\}$, $\mathbf{T} = \{t_n : n = 1,...,N\}$, $\mathbf{A} = \{A_k^l : l = 1,...,L : k = 1,...,K\}$, $\mathbf{R} = \{R_{y,k} : y = 1,...,Y_k : k = 1,...,K\}$ (defining Y_k as the number of PCR replicates used for kit *k*) and $\mathbf{B} = \{B_k : k = 1,...,K\}$ as **B**. The variables **D**, **A,** **R**, **T** and **B** are written collectively as **M**.

Note that the degradation term is optionally kit dependant, despite it being a property of the DNA extract and not the manner in which the profile has been generated. We make the choice to introduce an optional kit dependence on degradation as a result of the type of samples for which this analysis will commonly be used. That is, replicates that span technology (e.g. profiling kit) are likely to be separated by some time, during which degradation will have been acting on the sample. Therefore, the degradation term should be free between kits. For samples where this is not the case, the degradation values for contributors should be similar between kits and hence the term would be fixed in those situations (this property could be used as an analysis diagnostic).

Knowing the values of the parameters in **M** allows the calculation of Total Allelic Product (*T*), the total amount of fluorescence expected resulting from an allele in a DNA extract, which will ultimately get broken into components of fluorescence in an allelic position and its stutter positions on the electropherogram (EPG). Calculation of *T*, for a combination of contributor, kit, locus, replicate and allele, is achieved formulaically by:

$$T_{a,n,y,k}^l = t_n \times A_k^l \times R_{y,k} \times B_k \times X_n^l \times e^{-d_{n,k}\left(m_{a,k}^l - offset_k\right)} \quad (1)$$

ACCEPTED MANUSCRIPT

The X_n^l term in equation 1 represents a 'dose' and takes values of 1 or 2. The dose considers that if contributor *n* is homozygous for allele *a* at locus *l* ($X_n^l = 2$), then the expected value for *T* will be twice as high than if allele *a* was one in a heterozygous pair. The offset marks the molecular weight at which degradation starts to be applied, i.e. at the offset (and technically before it), degradation is not acting to reduce fluorescence. This offset is usually set to be the lowest molecular weight peak observed in one or more electropherograms (or some value below it), and so only depends on the profiling kit being used. As the PCR occurs, some of the fluorescence that was destined for the allele will shift to stutter positions on the EPG. There are a number of stutter types that can occur (back stutter, forward stutter, half stutter, double stutter, etc.) and we will define the number of types of stutter as *I*, the stutter ratio of stutter type *i*, for locus *l* in kit *k* for allele *a* as $\pi_{a,k}^{l,i}$ and the position of stutter type *i* relative to the parent peak, *a*, as Δ^i. We can now split the total allelic product into components of, respectively, allele and stutter by:

$$E_{a,n,y,k}^{l,\bar{i}} = \frac{T_{a,n,y,k}^l}{1+\sum_i \pi_{a,k}^{l,i}} \quad (2)$$

and

$$E_{a+\Delta^i,n,y,k}^{l,i} = \frac{T_{a,n,y,k}^l \pi_{a,k}^{l,i}}{1+\sum_i \pi_{a,k}^{l,i}} \quad (3)$$

where $\bar{i}$, indicates 'not *i*' and hence the allelic component. The total expected height of a peak at a locus, replicate and kit combination is then the sum of the stutter and allelic components of all individuals that fall on that allelic position:

$$E_{a',y,k}^l = \sum_n E_{a',n,y,k}^{l,\bar{i}} + \sum_n \sum_i E_{a+\Delta^i,n,y,k}^{l,i} \quad \text{where } a \text{ is chosen for each } i \text{ so that } a+\Delta^i = a' \quad (4)$$

Doing this for all contributors, alleles, loci, replicates and kits results in $\sum_k Y_k$ expected profiles, each of which has an observed counterpart, for which each peak height can be compared. Differences between observed and expected peak heights, $\Pr\left(O_{a,y,k}^l \mid E_{a,y,k}^l\right)$, are modelled by:

$$\log_{10}\left(\frac{O_{a,y,k}^l}{E_{a,y,k}^l}\right) \sim N\left[0, \left(1-\sum_i P^i\right)\frac{\left(c_k^{\bar{i}}\right)^2}{E_{a,y,k}^l} + \sum_i \frac{P^i\left(c_k^i\right)^2}{O_{a+\Delta^i,y,k}^l}\right] \quad (5)$$

ACCEPTED MANUSCRIPT

Where P^i is the proportion of peak *a* that is stutter type *i*. The right-hand side of equation 5 signifies modelling using a normal distribution, in the form N[mean,variance]. Note that the c^2 parameters in the variance term also depend on kit. Each has a prior gamma distribution modelled by:

$$\left(c_k^i\right)^2 \sim \Gamma\left(\alpha_k^i, \beta_k^i\right)$$

allowing each peak height variability term to adjust according to the observed data.

The functioning of the maths can be difficult to glean from formulae alone and so we show (in Figure 1) a flow diagram of the process for one iteration of an MCMC analysis. The top row of Figure 1 shows the parameters in the model. Some of these are used to build up expected DNA profile(s), of which some are mass parameters and others are fixed parameters (e.g. stutter ratios). This can be seen occurring as one moves down through Figure 1 from the top to approximately the centre of the diagram. Other parameters in the top row are used to calculate the likelihood of the observed data given the mass parameters, which can be seen moving vertically down through Figure 1 from approximately the centre to the bottom of the diagram.

In Figure 1 all values that are used in the model are shown in the top layer as vectors of values (this is indicated by bolding of the font type). Parameters in squares are fixed (i.e. their values are set prior to the analysis) whereas values in circle are variables in the MCMC. As one moves through the layers of resolution from analysis to kit, replicate, locus, allele, stutter type and contributor each parameter within the calculation is broken into single values or sometimes smaller vectors of values if the initial vector was multi-dimensional. Where the values are split in Figure 1 demonstrates where dependencies exist between that parameters and the layer of resolution. At the finest level of resolution we end up with expected peak heights for each kit, replicate, locus, allele, contributor and stutter type combination as shown in equations 2 and 3. These are summed together across stutter type and contributor to obtain the expected heights of alleles per kit, replicate, locus and allele as per Equation 4 (and procedurally shown in Figure 1 in the central zone). At this stage the expected peak heights can be compared with their corresponding observed peak heights and (with the use of the variance terms that are specific to both kit and stutter type) likelihoods for each allele, locus, replicate and kit combination can be calculated. The product of these across each of the layers

ACCEPTED MANUSCRIPT

of resolution ultimately ends up with the likelihood of the data given all the mass parameter values in the model for that MCMC iteration. This likelihood is then used in an acceptance rejection algorithm to either accept or reject the proposed **M**.

Note that Figure 1 is only a flow diagram and does not show the various prior probability distributions for each parameter in the model.

3.0 VALIDATION METHOD

Profiles were interpreted covering the following situations where laboratories might use multi-kit analyses:

1. Single source and two person mixtures amplified using Applied Biosystems Identifiler™ profiles analysed using both a 3130 and 3500 CE after an instrument upgrade, 18 examples shown.
2. Single source DNA amplified using both the Applied Biosystems Identifiler™ and MiniFiler™ multiplexes in order to improve profile quality for degraded and inhibited profiles, 15 examples shown.
3. Single source DNA amplified using both the PowerPlex® ESI 17 Pro and Applied Biosystems NGM SElect™ multiplexes, DNA profiling kits with the same loci, most loci ordered differently with respect to molecular weight, 11 examples shown.
4. *In silico* profiles generated using the Identifiler™ and GlobalFiler™ multiplexes, mimicking a two person mixture where the person of interest (POI) is partial and dropping out. The GlobalFiler™ profile was deliberately degraded, modelling a time delay between amplifications, 11 examples shown.

All profiles were interpreted using STRmix™ version 2.5.08 both as a single replicate using one kit and combined using multiple kits (i.e. three interpretations per sample set). Likelihood ratios were calculated using the African American allele frequencies from the FBI expanded CODIS core set [33] and a theta (F_{ST}) of 0.01. The propositions considered were:

H_p: the DNA originated from the person of interest and N-1 unknown contributors

H_d: the DNA originated from N unknown contributors

Where N was the number of contributors.

ACCEPTED MANUSCRIPT

4.0 RESULTS

A plot of the log(*LR*) for the single kit analysis versus the increase in the log(*LR*) after the multi-kit analysis for each dataset is provided in Figure 2. The increase is calculated by the difference between the *LR* obtained when comparing the reference to the deconvolution carried out using information from a single kit, and the *LR* obtained when comparing the reference to the deconvolution carried out using information from both kits. Note that an increase in the *LR* is not necessarily expected from all multi-kit analyses, just as an increase in the *LR* will not always be the results of adding replicates in a single kit analysis. Whether an increase or a decrease is obtained will depend on the additional information being provided, and how well the reference being compared can explain the observed data. In the experiments carried out for the mutli-kit analysis, all additions did result in an increase in the *LR*, as seen in Figure 2.

We also demonstrate an *in-silico* multi-kit example that illustrates a degraded profile typed in two kits with complementary locus sizes. A toy kit example has been fabricated to demonstrate the effect. Figure 3 shows the two profiles.

The profiles were set up so that the low molecular weight genotypes are unambiguous, the central molecular weight locus is ambiguous (i.e. dropout could have occurred) and there is very limited, or no, high molecular weight information. The genotype set weights resulting from the deconvolution are shown in Table 1.

The missing information in one kit (due to degradation) is present in the other allowing a complete resolution of the donor's genotype when deconvolution is carried out using both kits together. This is not a situation that could be rectified by additional replicate amplifications of the profile in a single kit only as it is likely the longer fragments would never amplify above the analytical threshold.

5.0 DISCUSSION:

Inspection of Figure 2 shows that, as expected, the increase in log(*LR*) after multi kit analysis depends on the amount of profiling data within the two profiles. For example, the combined interpretation of Identifiler and MiniFiler profiles has little improvement in *LR* when compared to the single Identifiler profile interpretation but significant (up to nine orders of

ACCEPTED MANUSCRIPT

magnitude) improvement compared to the single MiniFiler profile interpretation. This is intuitive given that the eight loci within the MiniFiler kit are a subset of the Identifiler kit loci. The MiniFiler profiles contained little additional profiling information compared to their Identifiler replicates despite targeting degraded DNA samples.

The position that PCR fragments elute from capillary electrophoresis instruments can be deliberately affected by mobility modifiers [34] (we term this the observed fragment sizes), which have been added during PCR, rather than being indicative of the fragment size as targeted by the primers (we term this, actual fragment size). The effect that these mobility modifiers have on the profile is that degradation will not have its most extreme effect (of reducing fluorescence) on the fragments that eluted last (i.e. the largest observed fragments). This is not an issue that is new to multi-kit analysis, as the presence of extreme size shifts by mobility modifiers in single kit analyses could be at odds with the degradation model employed (or indeed any model that expects a general downward trend in peak heights with observed fragment size). From the author's experiences with many profiling systems, degradation does appear to act in the manner that would be expected if observed fragment size and actual fragment size were roughly aligned, suggesting that mobility modifiers may not be used for extreme size shifts, but rather finer scale adjustments. For these finer scale adjustments, the amplification efficiency multipliers that are used in our model account for slight deviations in expected and observed peak heights that would be caused by differences in observed and actual molecular weights from mobility modifiers.

The inclusion of profiling information generated using a 3500 CE instrument to that previously generated using a 3130 CE instrument did not generally result in an increase in *LR*. This was because the laboratory validated 3500 AT was relatively high (260 rfu all loci) compared to the 3130 AT (50 rfu all loci) which resulted in (on average) more profiling information being captured in the 3130 profiles. Having said this there were instances for both technologies where the addition of the replicate produced in the other technology provided additional information to the system and higher *LR*s for known contributors. This is expected, if for nothing else, that the addition of any replicates (produced under the same or different conditions) in an analysis is likely to lead to better discrimination power [26].

Significant (up to 12 orders of magnitude) improvement was observed in the combined interpretation of ESI 17 and NGM SElect profiles compared to the single ESI 17 profile

ACCEPTED MANUSCRIPT

interpretation. Both kits have laboratory validated analytical thresholds of 50 rfu (all loci). Additional profile information was available when combining the kits, despite both kits containing the same loci and this was due to the locus order being different so that degradation differently on the same locus between kits.

The *in silico* profiles demonstrated the most significant increase in *LR*. These profiles were manufactured deliberately to have an effect. The person of interest was low template, exhibiting drop out at many loci. The corresponding GlobalFiler profiles, despite being artificially degraded, had many additional alleles contributing to the increased *LR*.

The effect was more significant (up to nine orders of magnitude) when the original single kit *LR* was low. This was the general trend for all data sets.

6.0 CONCLUSION

We present the maths for deconvoluting a DNA profile, considering replicates that have been produced under different conditions. These may include being different profiling systems, or the same system under different laboratory protocols, or using different hardware. Conceptually, the combination of such replicates is similar to the combination of replicates produced under the same conditions, in that genotypic weights become a combination of how well the observed data is described in all replicates by the chosen genotype set and mass parameters. The extension that allows the analysis to span replicates produced under different conditions is simply that some of the mass parameters (particularly those that are related to the functioning of the PCR) are allowed to vary between kits (or amplification conditions). We demonstrate the performance of the maths on some examples that align with scenarios likely to be encountered in casework and obtain the expected results. Again the results are similar to validations that consider the effect of replication (e.g. [26]) and follow intuitively sensible trends when additional loci are considered within the analysis.

The maths presented here should allow an even greater level of information to be utilised within forensic laboratories and under an expanded set of circumstances.

ACCEPTED MANUSCRIPT

ACKNOWLEDGEMENTS

This work was supported in part by grant 2011-DN-BX-K541 from the US National Institute of Justice. Points of view in this document are those of the authors and do not necessarily represent the official position or policies of their organisations.

ACCEPTED MANUSCRIPT

REFERENCES

[1] Taylor D, Bright J-A, Buckleton J. The interpretation of single source and mixed DNA profiles. Forensic Science International: Genetics. 2013;7:516-28.

[2] Perlin MW, Legler MM, Spencer CE, Smith JL, Allan WP, Belrose JL, et al. Validating TrueAllele® DNA mixture interpretation. Journal of Forensic Sciences. 2011;56:1430-47.

[3] Lohmueller K, Rudin N. Calculating the weight of evidence in low-template forensic DNA casework. Journal of Forensic Sciences. 2013;58:s234-59.

[4] Gill P, Kirkham A, Curran J. LoComatioN: A software tool for the analysis of low copy number DNA profiles. Forensic Science International. 2007;166:128-38.

[5] Balding DJ, Steele CD. likeLTD v6.0: an illustrative analysis, explanation of the model, results of validation tests and version history

2015. [Accessed] 07/08/2017. https://sites.google.com/site/baldingstatisticalgenetics/software/likeltd-r-forensic-dna-r-code

[6] Balding DJ, Buckleton J. Interpreting low template DNA profiles. Forensic Science International: Genetics. 2009;4:1-10.

[7] Scientific Working Group on DNA Analysis Methods (SWGDAM). Guidelines for the Validation of Probabilistic Genotyping Systems. 2015. [Accessed] 07/08/2017. http://media.wix.com/ugd/4344b0_22776006b67c4a32a5ffc04fe3b56515.pdf

[8] Coble MD, Buckleton J, Butler JM, Egeland T, Fimmers R, Gill P, et al. DNA Commission of the International Society for Forensic Genetics: Recommendations on the validation of software programs performing biostatistical calculations for forensic genetics applications. Forensic Science International: Genetics. 2016;25:191-7.

[9] Bright J-A, Taylor D, McGovern C, Cooper S, Russell L, Abarno D, et al. Developmental validation of STRmix™, expert software for the interpretation of forensic DNA profiles. Forensic Science International: Genetics. 2016;23:226-39.

[10] Cooper S, McGovern C, Bright J-A, Taylor D, Buckleton J. Investigating a common approach to DNA profile interpretation using probabilistic software. Forensic Science International: Genetics. 2015;16:121-31.

[11] Taylor D, Bright J-A, McGoven C, Hefford C, Kalafut T, Buckleton J. Validating multiplexes for use in conjunction with modern interpretation strategies. Forensic Science International: Genetics. 2015;20:6-19.

[12] Bright J-A, Curran JM, Buckleton JS. Investigation into the performance of different models for predicting stutter. Forensic Science International: Genetics. 2013;7:422-7.

ACCEPTED MANUSCRIPT

[13] Bright J-A, Taylor D, Curran JM, Buckleton JS. Developing allelic and stutter peak height models for a continuous method of DNA interpretation. Forensic Science International: Genetics. 2013;7:296-304.

[14] Bright J-A, Taylor D, J.M. C, Buckleton JS. Degradation of forensic DNA profiles. Australian Journal of Forensic Sciences. 2013;45:445-9.

[15] Bright J-A, McManus K, Harbison S, Gill P, Buckleton J. A comparison of stochastic variation in mixed and unmixed casework and synthetic samples. Forensic Science International: Genetics. 2012;6:180-4.

[16] Brookes C, Bright J-A, Harbison S, Buckleton J. Characterising stutter in forensic STR multiplexes. Forensic Science International: Genetics. 2012;6:58-63.

[17] Ballantyne J, Hanson EK, Perlin MW. DNA mixture genotyping by probabilistic computer interpretation of binomially-sampled laser captured cell populations: Combining quantitative data for greater identification information. Science & Justice. 2012.

[18] Perlin MW, Sinelnikov A. An information gap in DNA evidence interpretation. PLoS ONE. 2009;4:e8327.

[19] Gill P, Haned H, Bleka O, Hansson O, Dorum G, Egeland T. Genotyping and Interpretation of STR-DNA: Low-template,Mixtures and Database Matches - twenty years of research anddevelopment. Forensic Science International: Genetics. 2015;18:100-17.

[20] Prieto L, Haned H, Mosquera A, Crespillo M, Alemañ M, Aler M, et al. Euroforgen-NoE collaborative exercise on LRmix to demonstrate standardization of the interpretation of complex DNA profiles. Forensic Science International: Genetics. 2014;9:47-54.

[21] Gill P, Haned H. A new methodological framework to interpret complex DNA profiles using likelihood ratios. Forensic Science International: Genetics. 2013;7:251-63.

[22] Haned H, Dorum G, Egeland E, Gill P. On the meaning of the likelihood ratio: is a large number always an indication of strength of evidence? 25th Congress of the International Society for Forensic Genetics. Melbourne, Australia2013.

[23] Benschop CCG, Haned H, de Blaeij TJP, Meulenbroek AJ, Sijen T. Assessment of mock cases involving complex low template DNA mixtures: A descriptive study. Forensic Science International: Genetics. 2012;6:697-707.

[24] Haned H, Slooten K, Gill P. Exploratory data analysis for the interpretation of low template DNA mixtures. Forensic Science International: Genetics. 2012;6:762-74.

[25] Haned H, Egeland T, Pontier D, Pène L, Gill P. Estimating drop-out probabilities in forensic DNA samples: A simulation approach to evaluate different models. Forensic Science International: Genetics. 2011;5:525-31.

ACCEPTED MANUSCRIPT

[26] Taylor D. Using continuous DNA interpretation methods to revisit likelihood ratio behaviour. Forensic Science International: Genetics. 2014;11:144-53.
[27] Bright J-A, Taylor D, Curran J, Buckleton J. Searching mixed DNA profiles directly against profile databases. Forensic Science International: Genetics. 2014;9:102-10.
[28] Taylor D, Bright J-A, Buckleton J. Considering relatives when assessing the evidential strength of mixed DNA profiles. Forensic Science International: Genetics. 2014;13:259-63.
[29] Slooten K. Identifying common donors in DNA mixtures, with applications to database searches. Forensic Science International: Genetics. 2017;26:40-7.
[30] Green P, Mortera J. Paternity testing and other inference about relationships from DNA mixtures. Forensic Science International: Genetics. 2017;28:128-37.
[31] Perlin M. Combining DNA evidence for greater match information. Forensic Science International: Genetics Supplement Series. 2011;3:e510-e1.
[32] Gill P, Whitaker JP, Flaxman C, Brown N, Buckleton JS. An investigation of the rigor of interpretation rules for STR's derived from less that 100 pg of DNA. Forensic Science International. 2000;112:17-40.
[33] Moretti T, Moreno L, Smerick J, Pignone M, Hizon R, Buckleton J, et al. Population data on the expanded CODIS core STR loci for eleven populations of significance for forensic DNA analyses in the United States. Forensic Science International: Genetics. 2016;25:175-81.
[34] Hill CR, Kline MC, Duewer DL, Butler JM. Strategies for concordance testing. Profiles in DNA (Promega). 2010;13.

ACCEPTED MANUSCRIPT

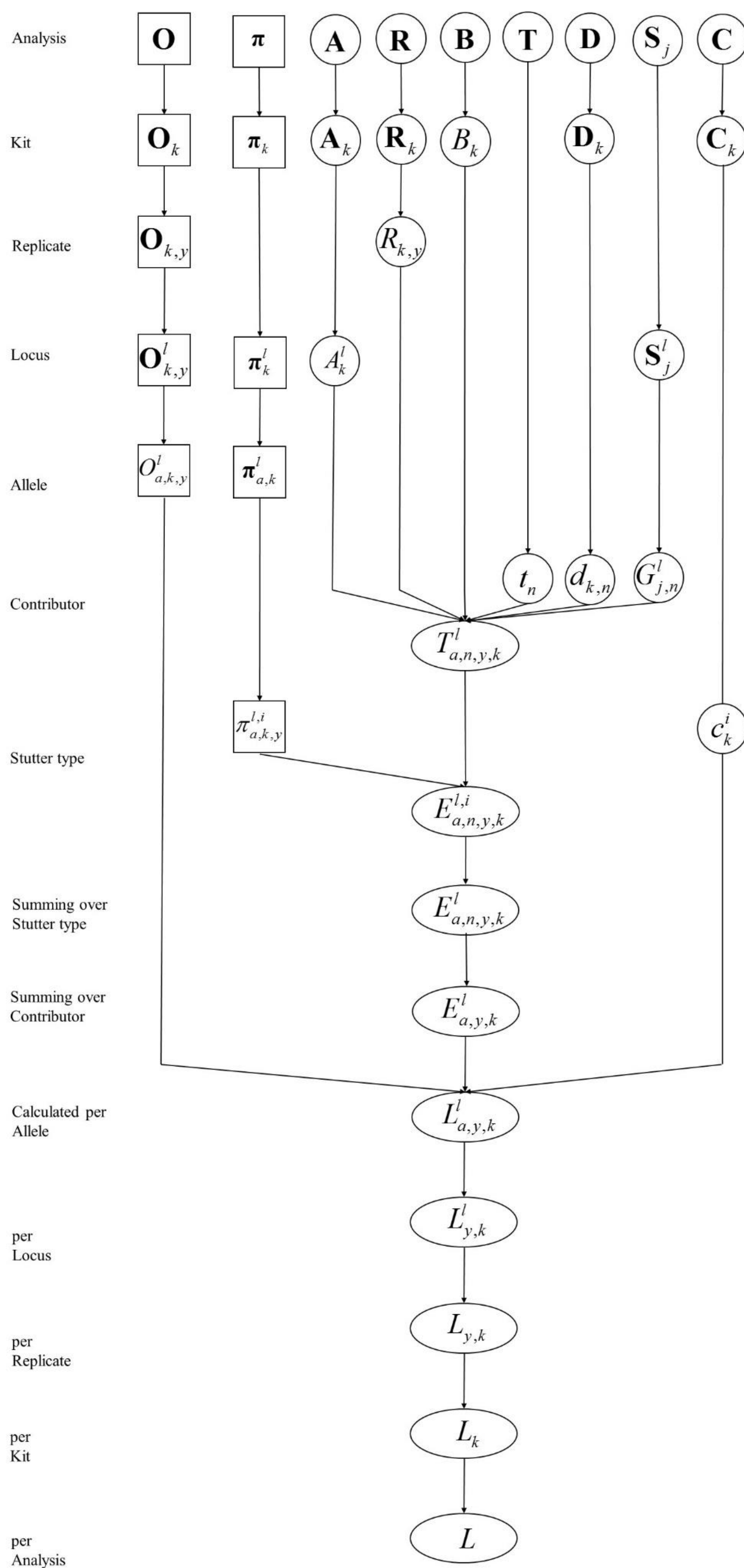

Figure 1: Diagrammatic representation of multi-kit profile building and evaluation

ACCEPTED MANUSCRIPT

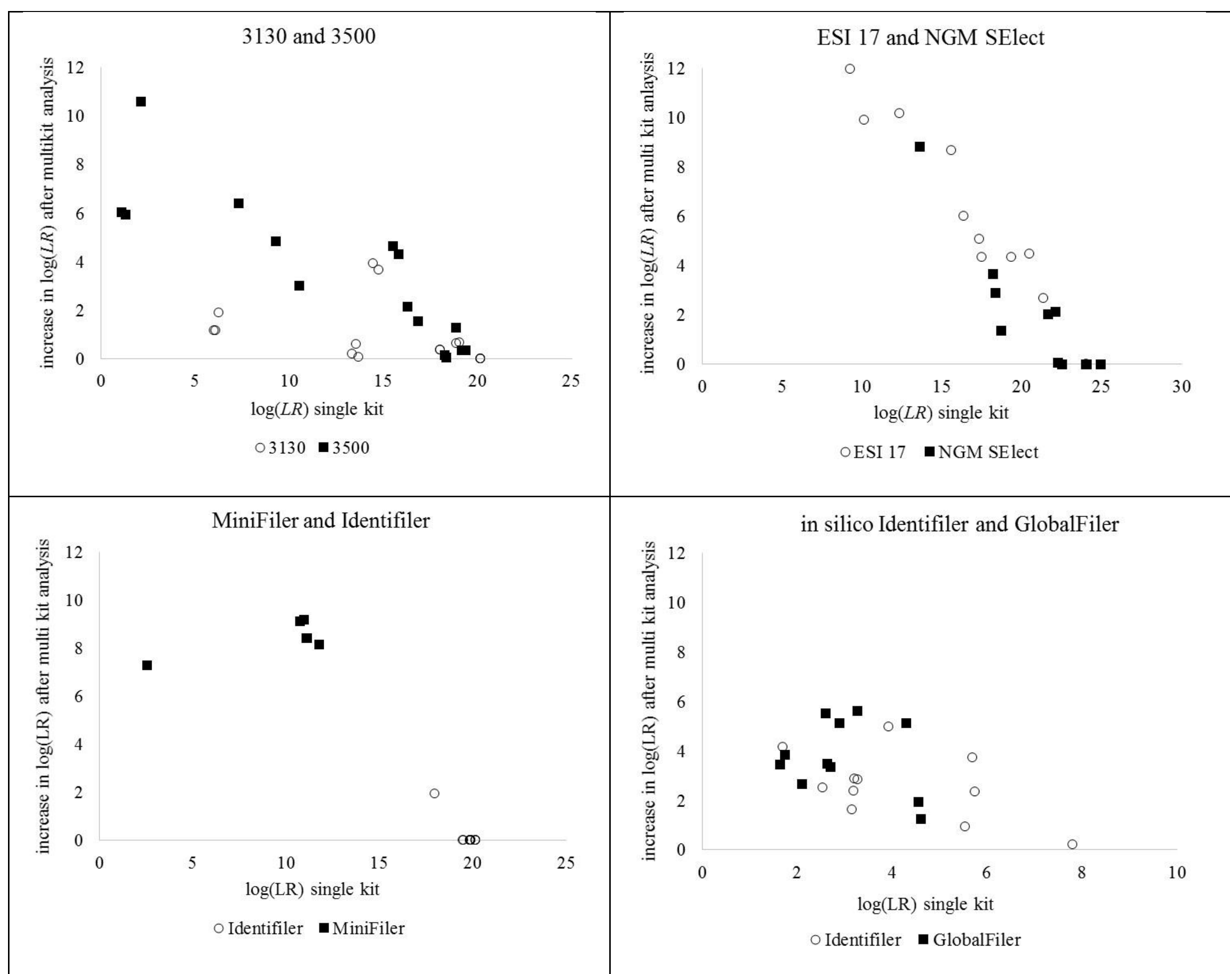

Figure 2: Log(LR) for the single kit analysis versus the increase in the log(LR) after the multi-kit analysis for each dataset

ACCEPTED MANUSCRIPT

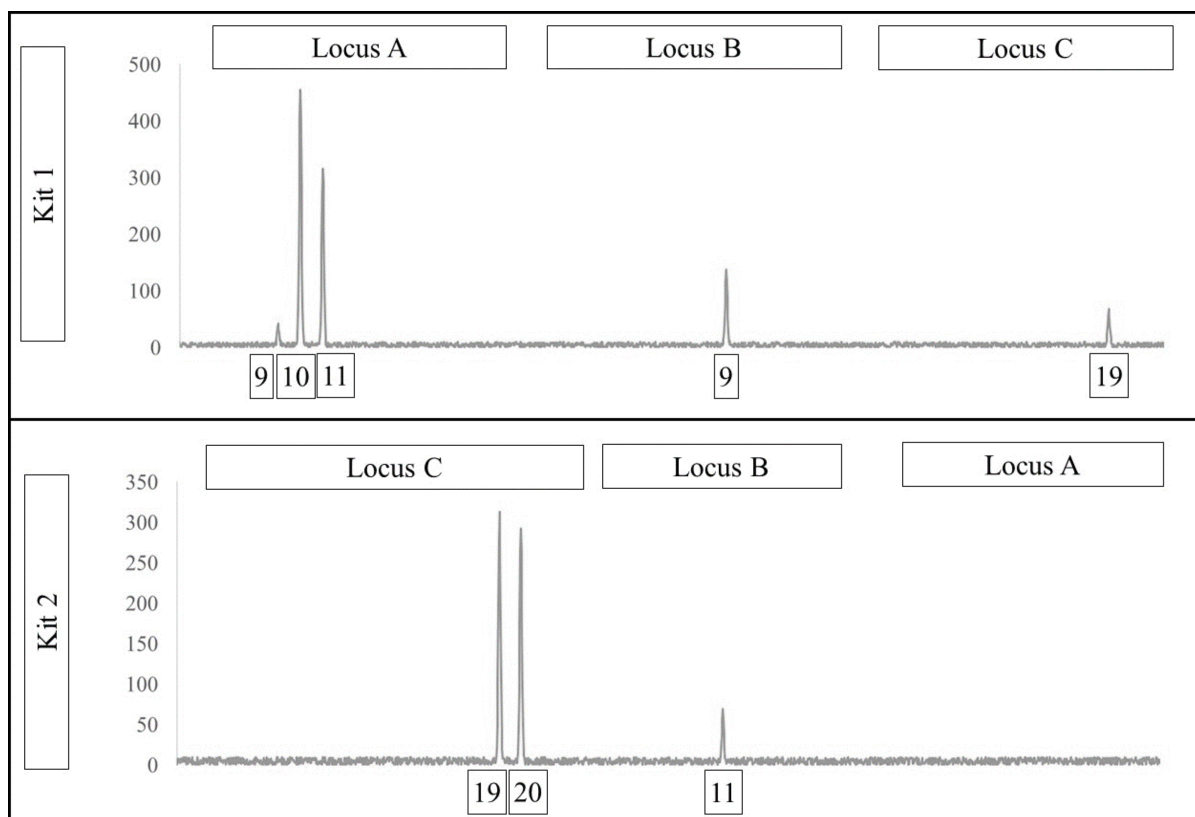

Figure 3: Example profiles from two profiling systems of a degraded sample

ACCEPTED MANUSCRIPT

Table 1: Weights produced from deconvolution of DNA profiles seen in Figure 3. Q is a generic letter used to signify an allele that has dropped out.

Locus	Kit 1 only	Kit 2 only	Kit 1 and Kit 2 together
A	[10,11] – 100%	-	[10,11] – 100%
B	[9,9] – 98% [9,Q] – 2%	[11,11] – 76% [11,Q] – 24%	[9,11] – 100%
C	[19,19] – 80% [19,Q] – 20%	[19,20] – 100%	[19,20] – 100%