



**Part I:**

# **Estimation of STRmix™ Parameters**

**District of Columbia  
Department of Forensic Science  
Forensic Science Laboratory  
Forensic Biology Unit**

The Institute of Environmental Science & Research (ESR) has taken all reasonable measures to ensure that the information and data presented in this document is accurate and current. However, ESR makes no express or implied warranty regarding such information or data, and hereby expressly disclaims all legal liability and responsibility to persons or entities that use or access this document and its content.

© 2015 Institute of Environmental Science and Research Limited (ESR)

## STRmix™ Implementation

This document describes the estimation of the STRmix™ parameters for Identifiler® Plus DNA profiling data from the Department of Forensic Science, Washington D.C. (DFS) for use in STRmix™ V2.3.

## STRmix™ Parameters

There are a number of parameters which are not optimised by the MCMC in a STRmix™ analysis. These parameters must be set by the user and are either determined by analysis of empirical data or modelled within STRmix™ using Model Maker. The laboratory specific parameters that are determined prior to use of STRmix™ are:

- Analytical threshold (detection threshold)
- Stutter ratios
- Drop-in parameters
- Saturation
- Allelic and stutter peak height variance
- The hyper-parameter for the variance of locus specific amplification effects (LSAE).

These parameters need to be defined for each STR kit, each protocol (e.g. cycle number variation), and CE platform (e.g. 3130 or 3500), and potentially each time there is a significant change of platform (e.g. a camera or laser change). Stutter ratios and saturation were determined for DFS 28 cycle Identifiler® Plus data analysed on a 3130 capillary electrophoresis instrument. Peak height variance and locus specific amplification efficiencies are calculated using Model Maker within STRmix™ from analysis of empirical profile data. The results of these analyses are described within this report.

## Analytical Thresholds

The assignment of a signal as allelic product as opposed to baseline or noise is important in DNA profile analysis. This differentiation is usually undertaken using a set threshold above which peaks are deemed to be allelic if they also meet certain morphological requirements, and below which they are ignored, regardless of morphology. The issue is to assign an analytical threshold (AT) to minimise the detection of artefacts whilst maximising the detection of allelic peaks.

The original validation of Identifiler® Plus was approved in 2012 and established an overall analytical threshold of 70 rfu for all samples [1]. This value was set using the practices, guidelines and literature which were available at the time the studies were conducted. In 2014, the laboratory received criticism regarding the mixture interpretation and statistical calculation methods used in its casework. As a result, recommendations were made by the Department of Forensic Sciences Scientific Advisory Board [2] to update its procedures including the parameters established in the original validation. Several methods of calculating and applying analytical thresholds were tested using a range of high and low-level input DNA [3]. These methods were evaluated based on the quantity and quality of the resulting profiles. The following method for analysis was recommended:

Samples with at least one allele >1000rfu

Analytical Threshold (rfu)

Blue:	60
Green:	50
Yellow:	70
Red:	60
Orange:	70

Samples with all alleles <1000 rfu

Analytical Threshold (rfu)

Blue:	20
Green:	50
Yellow:	40
Red:	40
Orange:	70

The analytical thresholds were based on the equation  $2 * (\text{Maximum noise peak} - \text{minimum noise peak})$  which is suggested in the validation guidelines published by the Scientific Working Group on DNA Analysis Methods [4]. The noise peaks, or baseline, from samples amplified at ideal targets were used to set the analytical thresholds for samples with at least one allele greater than 1000 rfu. The noise peaks, or baseline, from negative amplification controls were used to set the analytical thresholds for samples with all alleles less than 1000 rfu.

In February 2015, a site-visit was conducted which included a review of the newly established analytical thresholds. As a result of the site-visit, a report was issued criticizing the method used to establish the new analytical thresholds [5]. It stated that using negative controls to set analytical thresholds was an inappropriate method and should be re-calculated and evaluated using low level samples. This criticism was based on their assumption that the negative controls used to set the lower analytical thresholds were not amplified. However, the laboratory later clarified that the negative controls used were negative amplification controls, samples which were considered acceptable. In addition, the laboratory outsourced a review of the validation studies to Sorenson Forensics. The detailed report issued in July 2015 contained recommendations for improvement of other studies, but no criticism of the process used to set the analytical thresholds.

Despite the misunderstanding from the April 2015 report, the laboratory decided to demonstrate that the criticism was taken seriously and perform a third evaluation of the analytical thresholds. 21 low level samples (rfu less than 800) from past performance checks and amplification kit quality control checks were used to calculate a new set of analytical thresholds. The calculated results are the following:

**Low Level Samples (0~800 rfu)**

SWGAM Analytical Threshold = 2(Maximum peak height - minimum peak height)

	max	min	Analytical Threshold
blue	14	1	26
green	11	1	20
yellow	19	1	36
red	23	1	44
overall	23	1	44

While Limit of Detection (LOD) and Limit of Quantitation (LOQ) were calculated for these samples, no comparisons were made. In the Re-evaluation of Identifiler Plus validation study, it was determined that LOD and LOQ were not practical methods of calculating analytical threshold.

These values only show slight variation in rfu when compared to the thresholds calculated from the negative amplification controls (listed below), with the exception of the green dye channel.

**Negative Samples**

SWGAM Analytical Threshold = 2(Maximum peak height - minimum peak height)

	Max	min	Analytical Threshold
blue	13	1	24
green	31	1	60
yellow	22	1	42
red	23	1	44
overall	31	1	60

Further investigation of the noise peaks observed in the green dye channel indicated that the undefined 31 rfu peak in the negative amplification control may be an outlier. The next highest peaks observed between both the negative and low level samples was at 13 rfu.

A new set of analytical thresholds based on the low level samples tested above and the positive samples from the re-evaluation were then tested using similar sample sets as the Methods sections of the Identifiler Plus re-evaluation studies. All values were left unrounded.

Samples with at least one allele >1000 rfu

Analytical Threshold (rfu)

Blue: 58  
Green: 46  
Yellow: 72  
Red: 56  
Orange: 70

Samples with all alleles <1000rfu

Analytical Threshold (rfu)

Blue:	26
Green:	20
Yellow:	36
Red:	44
Orange:	70

The sample set analysed included 265 samples of varying peak heights. Each discordant peak was evaluated to ensure no drop-in was observed. All 162 discordant peaks were easy to identify artifacts such as pull up, background, minus A, and stutter. Of the 8449 concordant alleles, 996 would not have been detected with the original 70 rfu, non-color specific, analytical threshold.

Based on the above listed study, the possible outlier in the green dye channel from the re-evaluation and the previous criticism to set the analytical thresholds using the types of samples to which they will be applied, it was proposed to use the newly calculated analytical thresholds derived from low level samples for profiles where all allelic peaks are less than 1000rfu. Therefore, the optimum AT values which have been determined by DFS for all the Identifiler® Plus loci are dye-specific. Where any peak within a profile > 1000 rfu, the AT values are B=58 rfu, G=46 rfu, Y=72 rfu, and R=56 rfu (hereafter high AT) and where all peaks within a profile < 1000 rfu the AT values are B=26 rfu, G=20 rfu, Y=36 rfu, R=44 rfu (hereafter low AT). These values were used for all data analysis within this report.

#### Stutter

The laboratory recently completed an in-house stutter study as a part of the Re-evaluation of Identifiler® Plus referred to in the previous section. This study was conducted to compare laboratory stutter ratios to manufacturer stutter ratios and to establish a laboratory-specific stutter filter in the GeneMapper® ID-X software (GMID-X). One of the limitations of GMID-X is that the stutter filter can only be applied per locus and research has shown that modelling stutter on a per allele basis is a better fit of expected stutter ratios. Therefore, a new stutter study was conducted for the implementation of STRmix because it is able to model stutter on a per allele basis.

There are three parameters within STRmix™ that calculate expected stutter rates and therefore require optimisation. The first is the maximum allowable stutter ratio. The maximum allowable stutter ratio reduces run time by only permitting peaks in a stutter position below a certain percentage to be considered stutter. This parameter has been set at 0.3 (30%) based on inspection of laboratory stutter ratio data.

The second parameter is a file used to model the expected heights of the stutter peaks based on their partner allele designation. The values used to determine expected stutter heights are 'per allele'. Per allele stutter ratios are calculated using a linear equation and regressing stutter ratio against allele. Within STRmix™, stutter is estimated using the model  $SR = m \times Allele + c$  where the intercept ( $c$ ) and slope ( $m$ ) are determined using regression. Values for  $m$  and  $c$  were calculated in R (data analysis software). A plot of  $SR$  versus Allele for each locus is provided in Appendix 1. A summary of the STRmix™ allelic stutter files for the DFS data is given in Table 1.

Table 1: DFS per allele Identifier® Plus stutter values for STRmix™

Locus	Intercept	Slope
D8S1179	0.00916	0.00403
D21S11	-0.0767	0.00488
D7S820	-0.0492	0.00931
CSF1PO	-0.04	0.00837
D3S1358	-0.0279	0.00667
TH01	0.0216	-0.00042
D13S317	-0.0541	0.00921
D16S539	-0.0467	0.00918
D2S1338	-0.0189	0.00457
D19S433	-0.0799	0.0112
vWA	-0.117	0.0113
TPOX	-0.0237	0.00581
D18S51	-0.0374	0.00753
D5S818	-0.0249	0.00728
FGA	-0.0636	0.00613

A better explanatory variable for stutter ratio for loci with compound and complex structure has been shown to be the longest uninterrupted stretch of common repeats (LUS) within the allele [6-8] and not the allele designation itself. Values for LUS are determined by sequencing alleles. A number of common alleles for forensic loci have been typed. A summary of these appear on STRBase [9, 10]. A plot of *SR* versus LUS for compound and complex loci within the Identifier® Plus multiplex is provided within Appendix 2.

The third parameter within STRmix™ that determines expected stutter peak heights is an exception file based on either LUS or an average observed stutter ratio. LUS is used where it is a good explanatory variable for *SR* otherwise the average of the observed *SR* is used. A stutter exception file based on laboratory data has been created and was used in this analysis. Where alleles are not present in this file the expected stutter rates are calculated from the allele file (Table 1). A summary of the source of the predicted *SR* for each locus is given in Table 2.

Table 2: A summary of the explanatory variables for the predicted *SR* for each of the DFS loci

Locus	Source
D8S1179	Average observed
D21S11	Average observed
D7S820	Allele
CSF1PO	Allele
D3S1358	Average observed
TH01	LUS
D13S317	Allele
D16S539	Allele
D2S1338	Average observed
D19S433	LUS
vWA	Allele
TPOX	Allele
D18S51	Allele
D5S818	Allele
FGA	Allele

#### Drop-in parameters

The DFS laboratory has not observed drop-in peaks within their Identifiler® Plus profiles amplified at 28 cycles.

#### Saturation

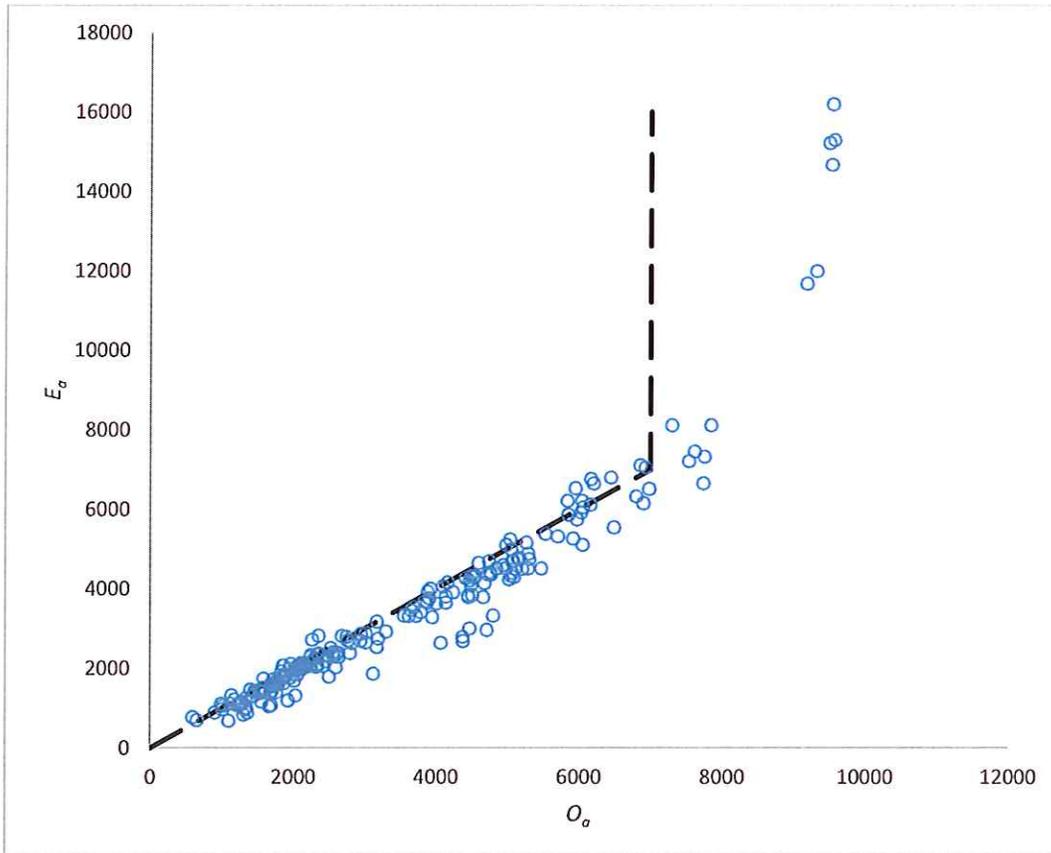
The peaks in a DNA profile are measured using fluorescence. The amount of fluorescence is proportional to the amount of DNA present. This fluorescence is captured by a camera. It is expected that as more DNA is added into a PCR the resulting peak height (measured in relative fluorescent units) in an electropherogram will increase. The camera can become saturated when there is too much fluorescence detected. This means we can no longer accurately measure the height of the peaks observed or estimate how much DNA is really represented by this result. Following this we can no longer accurately model over saturated peak heights using STRmix™. The saturation setting is the upper limit for a peak's height permitted in the software, beyond which the model is no longer optimal. The software will treat peaks in the input evidence data above this value as qualitative only. Saturation, like the analytical threshold, is mostly instrument related and not kit or method dependent.

The expected height of every allele within the stutter ratio dataset was calculated using the formula:

$$E_a = \frac{O_{a-1}}{\epsilon SR_a}$$

Where ( $E_a$ ) is the expected peak height calculated from the observed stutter height ( $O_{a-1}$ ) and  $\epsilon SR_a$  is the expected stutter ratio for allele  $a$  calculated using the equation described above. A plot of  $O_a$  versus  $E_a$  is provided in Figure 1. A vertical line at  $O_a = 7000$  rfu indicates a common saturation limit for a 3130 instrument. The points should deviate from the  $x = y$  line at the saturation value. After inspection of Figure 1 we recommend a saturation threshold setting of 7000 rfu is applied.

Figure 1: Observed versus expected peak heights



#### Peak height variance and LSAE using Model Maker

Empirical observations and experience suggests that profiles differ in variance (hereafter “quality”). Within STRmix™ the variability of peaks within profiles is described using a model containing a variance constant. Within V2.3 allele and stutter peaks have separate variances,  $c^2$  and  $k^2$ , respectively. The  $c^2$  and  $k^2$  terms are variables which are determined through the MCMC process. The starting position for these values within the MCMC is the mode of a gamma distribution based on empirical values from the DFS laboratory.

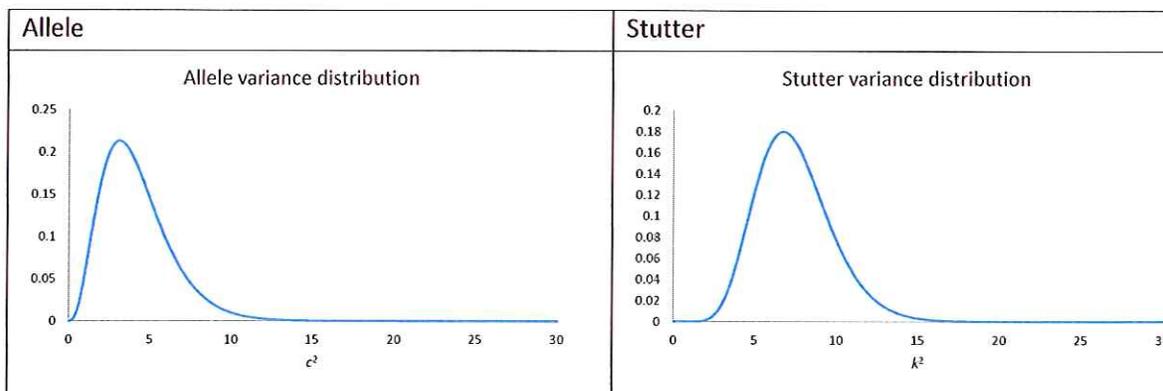
100 single source profiles of varying quality were analysed using the Model Maker function within STRmix™. Variances were calculated separately for the high and low AT samples however, the high AT variances will be used for all samples interpreted using STRmix™. This more reliably allowed Model Maker to create appropriate variances due to the lack of stutter detected in low AT samples. A summary of the results for both  $c^2$  and  $k^2$  for the dataset is provided in Table 3. Plots of the allele and stutter gamma distributions are provided in Figure 2.

Table 3: Summary of Model Maker results for DFS dataset

Multiplex	Number profiles analysed	Allele variance parameters (Mode)	Stutter variance parameters (Mode)	Mean LSAE variance
Identifiler® Plus 28 cycle	100	gamma(3.949,1.060) (3.126*)	gamma(10.391,0.717) (6.733*)	0.00688

\*Due to rounding, these values are slightly different than the supporting documentation. No measurable effect on the data is expected.

Figure 2: A plot of the allele and stutter gamma distributions



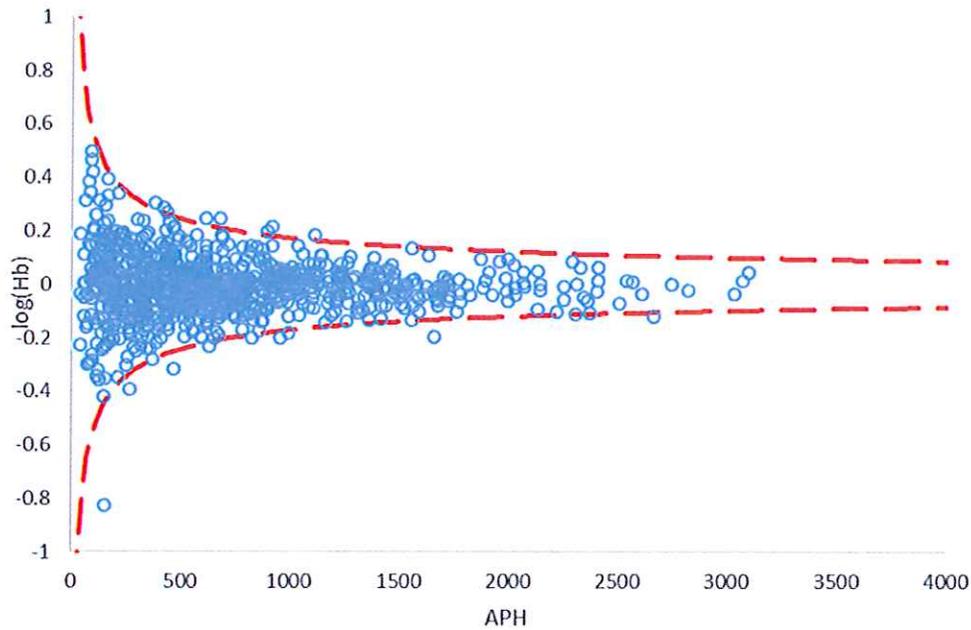
Heterozygote balance was calculated for all heterozygote loci for the Model Maker profiles. Heterozygote balance ( $Hb$ ) was calculated as:

$$Hb = \frac{O_{HMW}}{O_{LMW}}$$

Where  $O_{HMW}$  refers to the observed height of the high molecular weight allele and  $O_{LMW}$  the observed height of the low molecular weight allele. Previous work has suggested that there is a relationship between the variation in peak height and the variation in  $Hb$  [11, 12]. In single source profiles, variability in  $Hb$  reduces as the average peak height (APH) at a locus increases. The variance of  $Hb$  is expected to be twice the variance of the individual allelic peaks assuming the variance of each peak is the same. This allows an approximate comparison between the variance from the STRmix™ MCMC approach and a readily determined variable from empirical data.

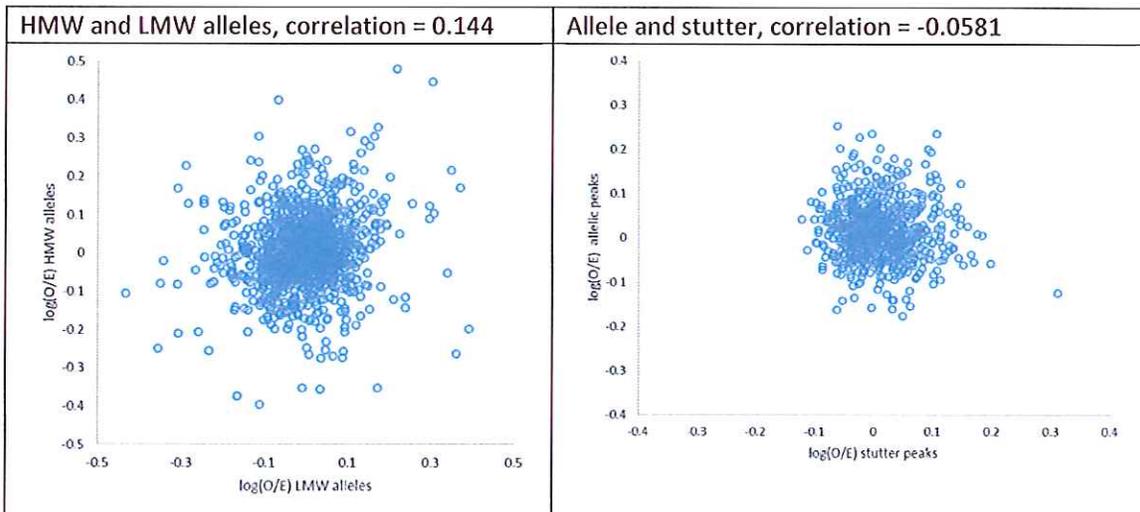
The plot of  $\log Hb$  versus APH for each of the datasets described above and the expected 95% bounds (plotted as dotted lines) calculated at  $\pm\sqrt{2} \times 1.96 \times \sqrt{\frac{c^2}{APH}}$  where  $c^2 = 3.84$ , the 50<sup>th</sup> percentile from the gamma distribution from the combination data set. The 95% bounds encapsulate sufficient data as demonstrated in the graphs (coverage = 97.7%) demonstrating that the values for variance are sufficiently optimised. This graph also verifies that the use of the high AT variances is also appropriate for low AT samples. The plot in Figure 3 is an approximate check of Model Maker.

Figure 3: Log(*H<sub>b</sub>*) versus APH for single source profiles



In Figure 4 we plot the correlation plots for LMW versus HMW allele and allele versus stutter peaks for the Model Maker dataset. The distribution of the points within the figures is as expected, with no observed correlation. There are some outliers observed in the logarithm of the observed over expected stutter peak height versus log(*O/E*) allelic peak height plot. These are larger than expected stutter peaks that were labelled at analysis however they do not affect the results.

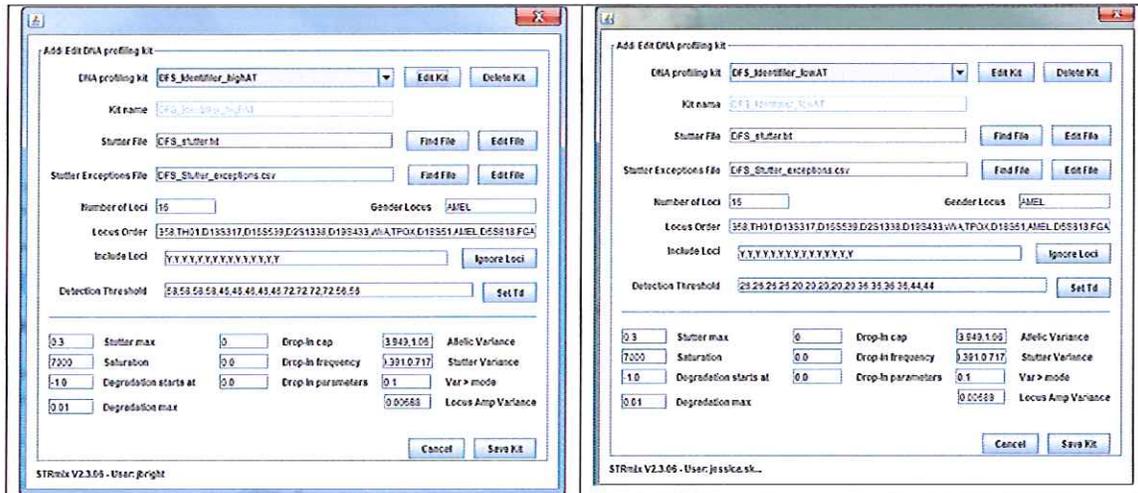
Figure 4: DFS Identifiler® Plus correlation plots



**Conclusions**

The recommended STRmix™ V2.3 default parameters for the interpretation of the DFS 28 cycle Identifiler® Plus profiles run on a 3130 CE instrument are given in Figure 5. The high AT parameters are on the left and the low AT parameters on the right.

Figure 5: STRmix™ recommended default parameters for DFS Identifiler® Plus profile interpretation

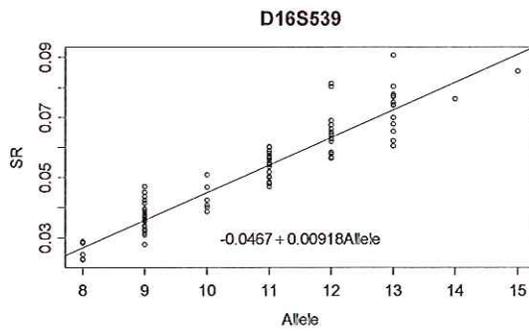
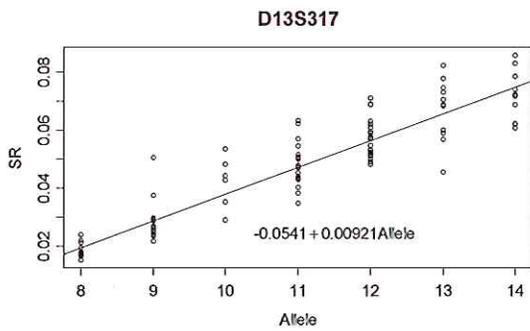
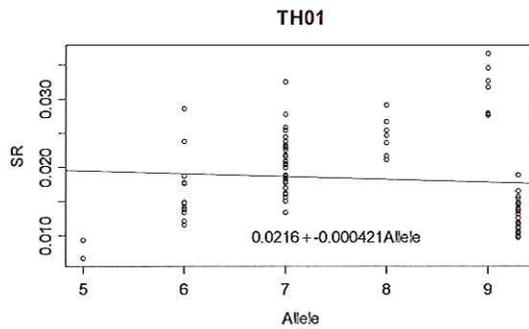
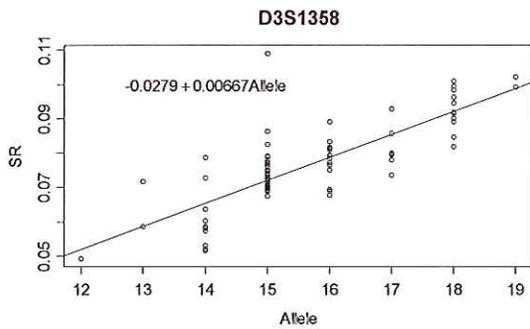
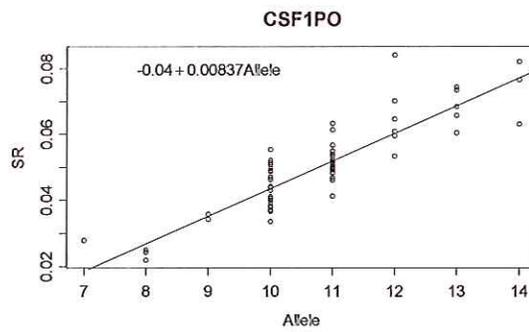
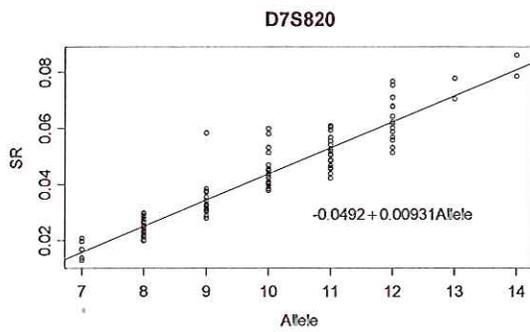
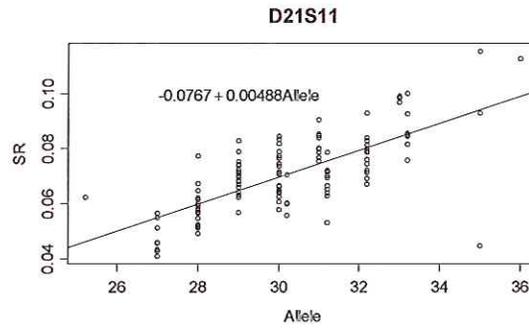
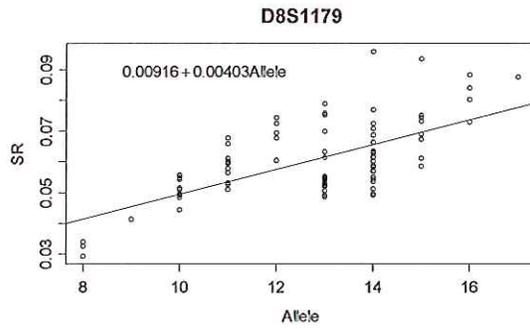


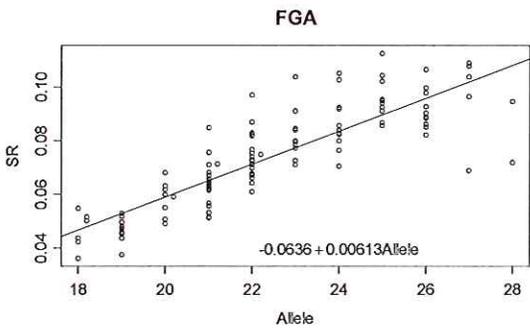
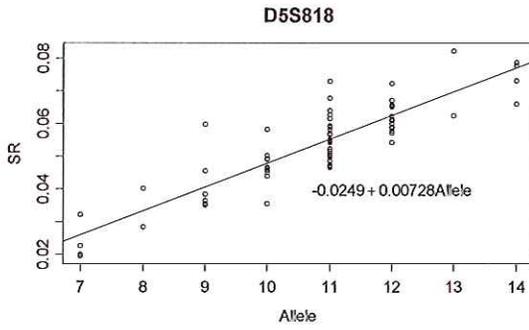
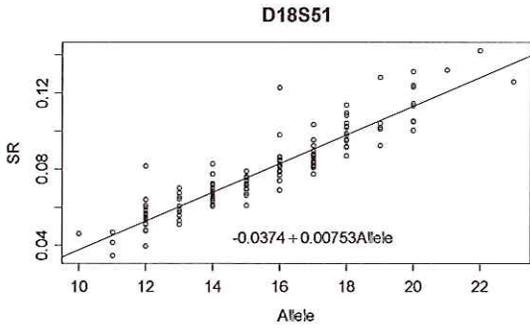
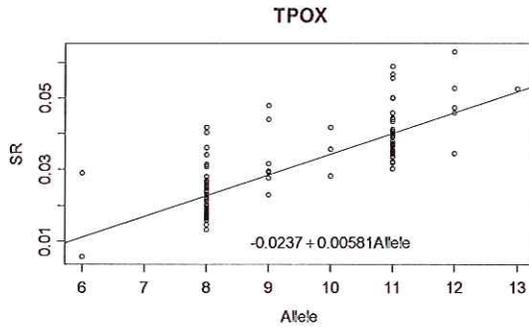
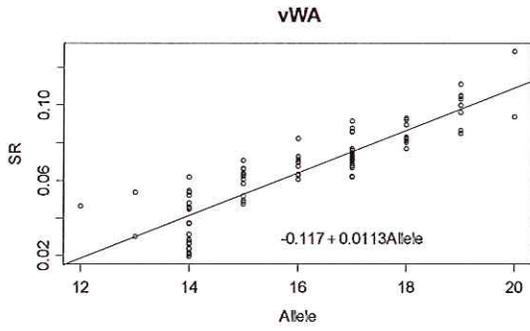
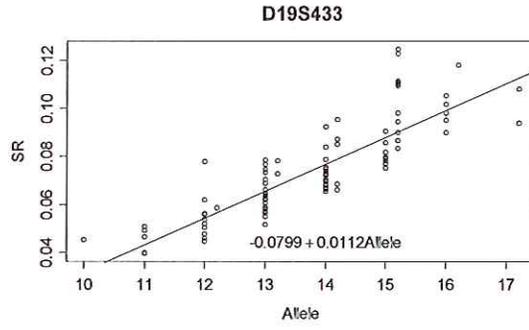
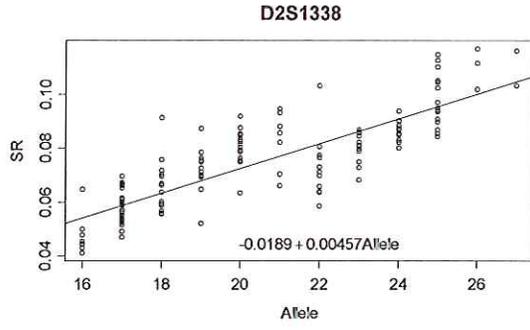
## References

- [1] Skillman J, Eschinger L. Applied Biosystems® AmpFISTR® Identifiler® Plus PCR Amplification Kit Validation Using the Applied Biosystems® 3130xL Genetic Analyzer. [http://dfs.dc.gov/sites/default/files/dc/sites/dfs/page\\_content/attachments/Identifiler%20Plus%20Using%20the%203130xl%20Validation%20Summary.pdf](http://dfs.dc.gov/sites/default/files/dc/sites/dfs/page_content/attachments/Identifiler%20Plus%20Using%20the%203130xl%20Validation%20Summary.pdf) Accessed.
- [2] Litofsky IB. Science Advisory Board Letter to Mayor. <http://dfs.dc.gov/page/letter-mayor-dfs-science-advisory-board-regarding-audits-april-26-2015> Accessed.
- [3] Bencivenga P. ANSI-ASQ National Accreditation Board Surveillance and Remote Surveillance Audit Report. <http://dfs.dc.gov/page/anab-report-april-2015> Accessed.
- [4] Scientific Working Group on DNA Analysis Methods (SWGDM). SWGDAM Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories. 2010.
- [5] Budowle B, Beiber FR. Final Report on Review of Mixture Interpretation in Selected Casework of the DNA Section of the Forensic Science Laboratory Division (FSL), Department of Forensic Sciences (DFS), District of Columbia. [http://dfs.dc.gov/sites/default/files/dc/sites/dfs/page\\_content/attachments/USAO%20Report%20Final.pdf](http://dfs.dc.gov/sites/default/files/dc/sites/dfs/page_content/attachments/USAO%20Report%20Final.pdf) Accessed 2015.
- [6] 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.
- [7] Brookes C, Bright J-A, Harbison S, Buckleton J. Characterising stutter in forensic STR multiplexes. *Forensic Science International: Genetics*. 2012;6:58-63.
- [8] Walsh PS, Fildes NJ, Reynolds R. Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA. *Nucleic Acids Res*. 1996;24:2807-12.
- [9] Butler JM, Reeder DJ. Short Tandem Repeat DNA Internet DataBase. [www.cstl.nist.gov/biotech/strbase](http://www.cstl.nist.gov/biotech/strbase) Accessed.
- [10] Ruitberg CM, Reeder DJ, Butler JM. STRBase: a short tandem repeat DNA database for the human identity testing community. *Nucleic Acids Research*. 2001;29:320 - 2.
- [11] Bright J-A, Huizing E, Melia L, Buckleton J. Determination of the variables affecting mixed MiniFiler™ DNA profiles. *Forensic Science International: Genetics*. 2011;5:381-5.
- [12] Bright J-A, Turkington J, Buckleton J. Examination of the variability in mixed DNA profile parameters for the Identifiler multiplex. *Forensic Science International: Genetics*. 2009;4:111-4.

Appendix 1

SR versus Allele





Appendix 2

SR versus LUS

