Identifiler Plus Mixture Summaries:

Supplemental Validation for Mx and Tiered PHRs and Performance Check for MixMaster IDP version 1

This supplemental validation and performance check for Identifiler Plus has been placed into the Identifiler Mixture Validation, Binder 1.

Validation performed by:

Steven Myers _____

Brian Harmon _____

Validation reviewed by:

Jeanette Wallin _____

Approved for use in Casework by:

Gary Sims _____ **Date** 12/21/12

Identifiler Plus Mixture Summaries:

Supplemental Validation for Mx and Tiered PHRs

This work supplements the mixture and MixMaster validation performed previously using the Identifiler kit and addresses Standard 8.7 of the FBI Quality Assurance Standards for Forensic DNA Testing Laboratories document (Sept. 2011) as well as Section 3.5 of the SWGDAM Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories (Jan. 2010).

QAS Standard 8.7	*Modifications to software, such as an upgrade, shall require a performance check prior to implementation. New software or significant software changes that may impact interpretation or the analytical process shall require a validation prior to implementation.*
SWGDAM 3.5	*An individual's contribution to a mixed biological sample is generally proportional to their quantitative representation within the DNA typing results. Accordingly, depending on the relative contribution of the various contributors to a mixture, the DNA typing results may potentially be further refined.*
SWGDAM 3.5.3	*A laboratory may define other quantitative characteristics of mixtures (e.g., mixture ratios) to aid in further refining the contributors.*

Because the functionality of MixMaster was thoroughly evaluated (including stutter, peak height ratio, and stochastic thresholds) during validation, the version of MixMaster that was modified to accommodate Identifiler Plus settings (MixMaster IDP version 1) was evaluated using a performance check.

Mx for 2-person Mixtures

Mixture proportion (Mx) is the RFU contribution of one individual within the total RFU for all contributors. The purpose of the Identifiler Plus Mx study was: 1.) to assess the variation of Mx among the loci of a single amplification, 2.) to examine how often an Mx range includes the Mx of the true genotype pairing when the average Mx is estimated based upon an incomplete knowledge of the profiles, and 3.) to compare this data to the previously generated Identifiler Mx data.

To address these questions, 42 known mixture studies from analysts across BFS were evaluated on 3130/3130*xl* Genetic Analyzers, for a total of 362 ratios amplified with Identifiler Plus. The total DNA input was either approximately 0.5 or 1 nanogram (ng). A typical set of nine ratios (~ 1 ng) included 19:1, 9:1, 4:1, 2:1, 1:1, 1:2, 1:4, 1:9, and 1:19, whereas a typical set of five ratios (~ 0.5 ng) included 9:1, 4:1, 1:1, 1:4, and 1:9. Variation in the observed "true" Mx for each 2-person mixture was assessed using a residuals approach (Gill *et al.*, 1998, FSI).

For data analysis, three different Mx scenarios were considered:

- 4-allele loci
- 3-allele loci
- 3-allele and 4-allele loci

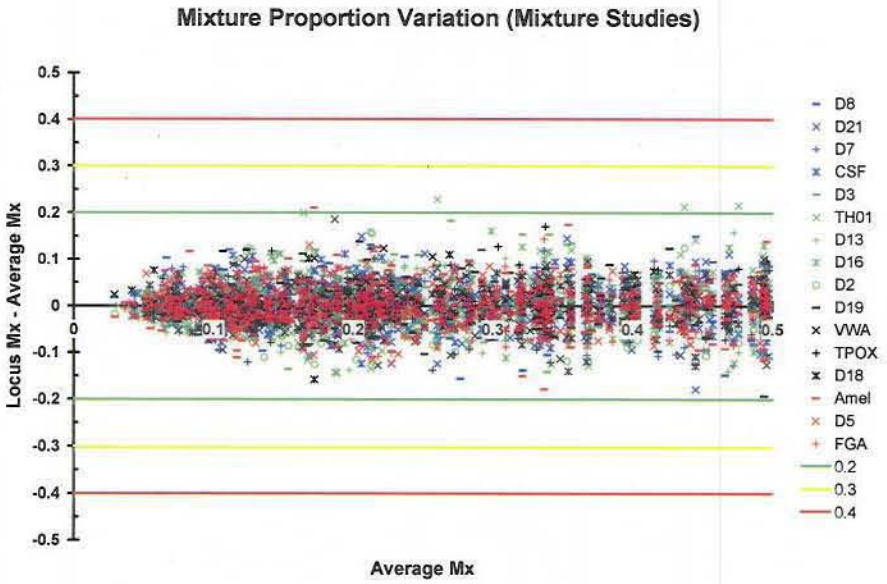
Considering 4-allele loci, this scenario is relied upon solely for the Mx calculation MixMaster when there is no assumed contributor. When there is an assumed contributor, the Mx is refined to include not only the true 4-allele locus pairings but also the 3-allele loci in which the assumed contributor is homozygous; note that this is still referred to as a "4-allele locus" throughout this validation. Both of these situations involve the calculation of an average Mx based solely upon an assessment of loci with no shared alleles.

For the 3-allele loci, this scenario includes those loci in which there may be a homozygous-heterozygous pairing or a heterozygous- heterozygous pairing with a shared allele. Both possibilities are considered when there is no assumed contributor. The use of 3-allele loci is invoked when there is a lack of 4-allele loci. An example of when this might occur is a mixture of a parent and child, where it is expected that the two will share an allele identical by descent at each locus. Another situation where this may be useful is with a mixture that has only one locus with four alleles, especially if this locus has the minor alleles in the stutter positions of the major alleles.

The 3-allele and 4-allele locus scenario simply refers to the use of both 4-allele loci and 3-allele loci from which to calculate an average Mx.

Mx variation when the genotypes of both contributors were known (a best-case scenario) is shown in Figure 1.

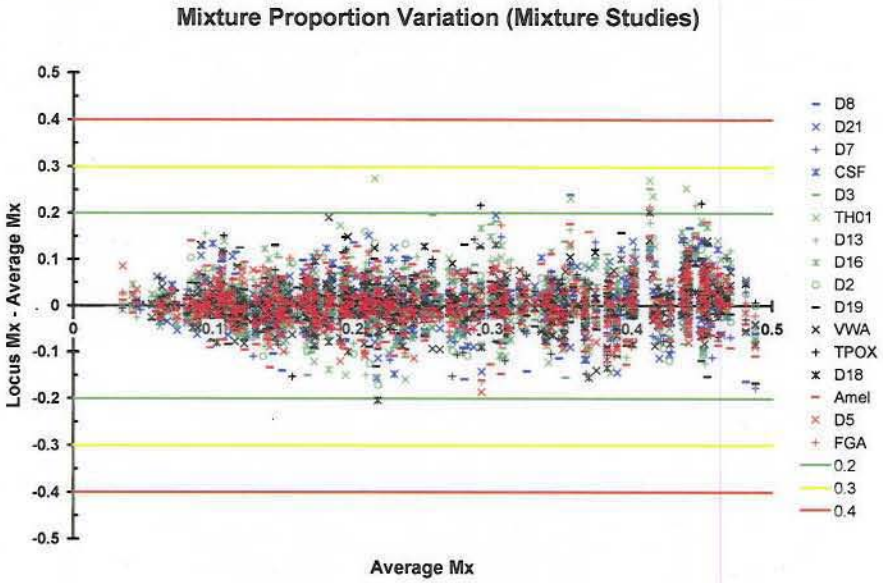
Figure 1



As shown in Figure 1, when the genotypes of both contributors were known, the true intra-profile inter-locus variation was evenly distributed around the average. Most values were within a range of ±0.2 from the average, regardless of average Mx, and all were less than 0.23 away from the average. This overall range was identical to that observed for the Identifiler kit study.

Mx variation shown in Figure 2 is from when the genotypes of neither contributor were known and only the 4-allele loci were used.

Figure 2



When the genotypes of neither contributor were known, and the average Mx was based solely upon an RFU assessment of 4-allele loci, the intra-profile inter-locus variation was evenly distributed around the average for most of the range. Most values were within a range of ±0.2 from the average, regardless of average Mx, and all were less than 0.28 away from the average.

Note that an Mx average based solely upon 4-allele loci when there is no assumed contributor may be underestimated. Because the Mx is calculated for the minor contributor, the two peaks with the lowest RFU are used in the numerator and may not truly pair. This may lead to an underestimation of the true Mx. The effect of this is most pronounced when approaching an Mx of 0.5 as the minor contributor may truly be the major contributor at another locus. Nevertheless, the average Mx was shown to closely model the respective individual locus Mx values, with the majority of data points falling within a ±0.2 range and all within a ±0.3 range.

In mixtures lacking four detected alleles, Mx may be estimated from 3-allele loci. The 3-allele approach was previously validated using Identifiler data; the calculations may be performed manually or using MxCalculator. Shown in Figure 3 is the Mx variation when the average Mx was based solely upon 3-allele loci and the genotypes of neither contributor were known.

Figure 3

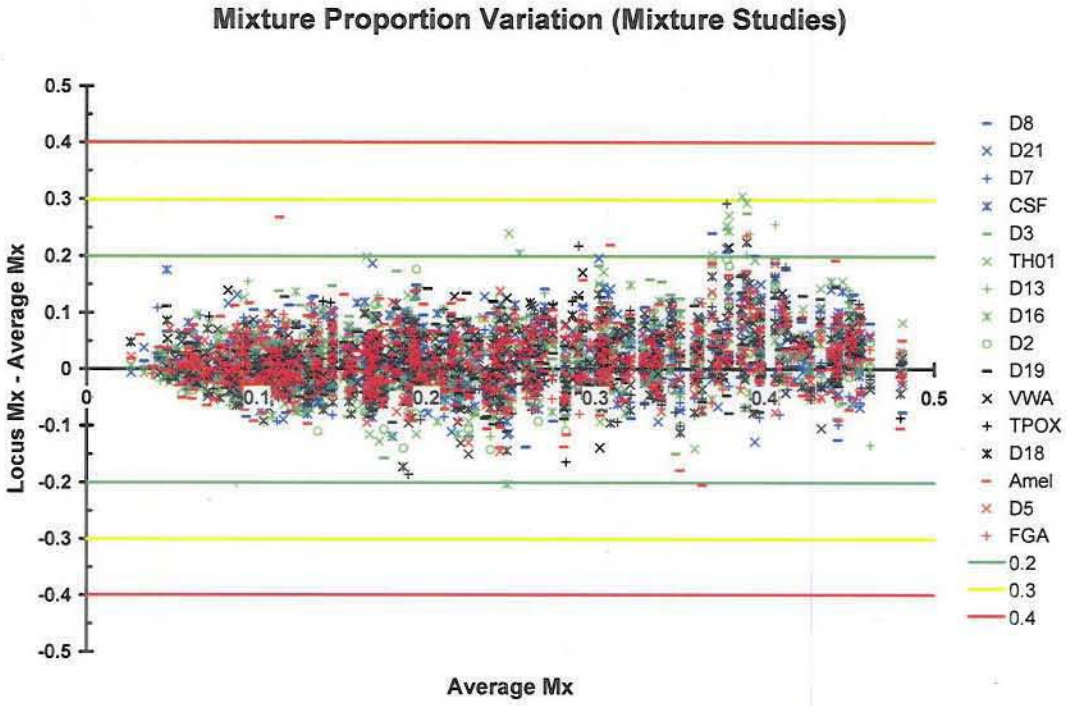


Figure 3 shows the intra-profile inter-locus variation was evenly distributed around the average for the lower part of the range. As the average Mx increased above ~0.3, it began to underestimate the true average. Despite this, most values were still within a range of ±0.2 from the average, regardless of average Mx, and all but one were within a ±0.3 range from the average. The lone outlier was a TH01 2:1 mixture result ~0.305 greater than the average; note that interestingly, a majority of the outermost values in each figure are attributed to TH01.

When the average Mx was based solely upon an RFU assessment of 3-allele loci and the genotypes of one of the contributors was known, all data points were within the ±0.3 range from the average, as shown in Figures 4 (female contributor assumed) and 5 (male contributor assumed).

Figure 4

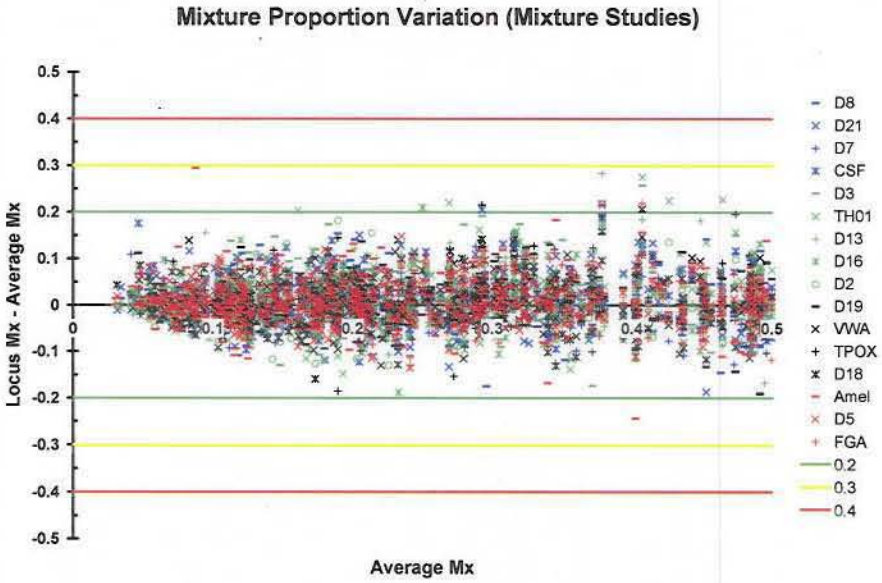
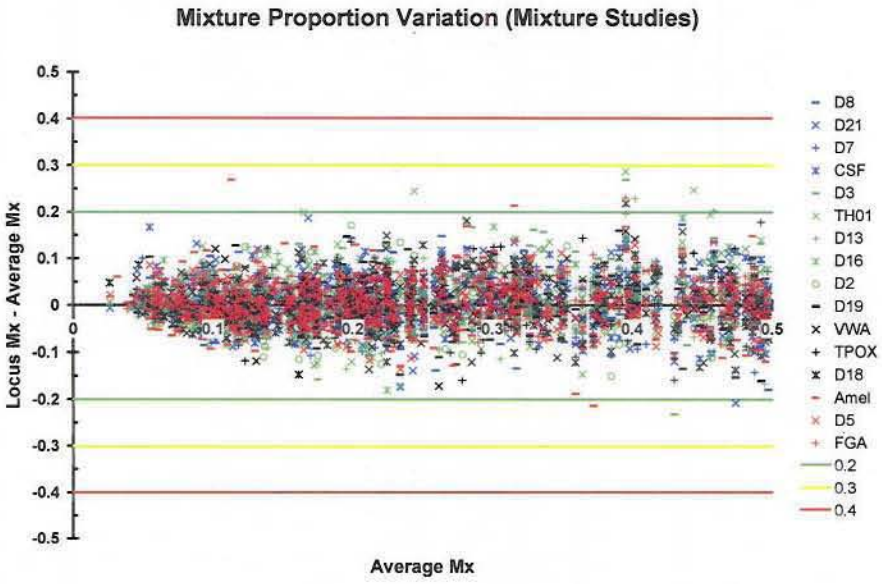
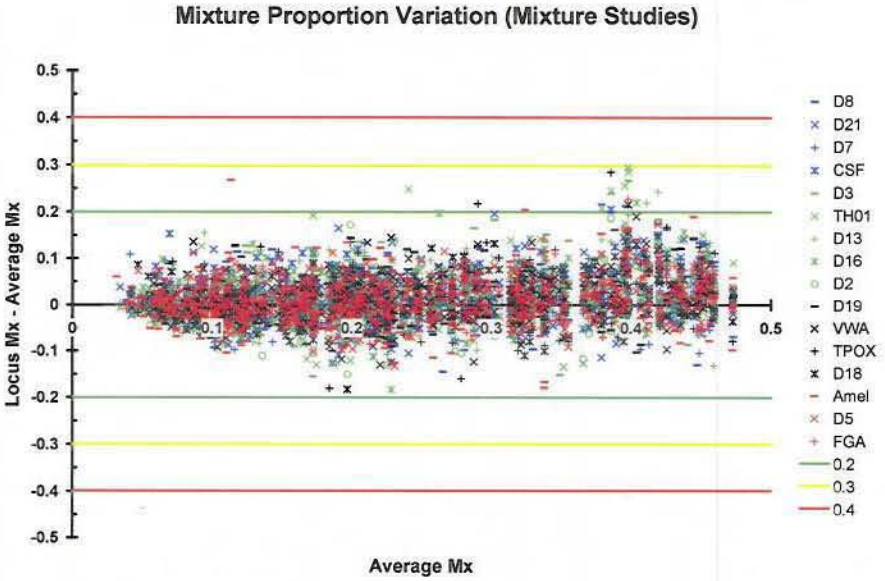


Figure 5



In instances where it may be desirable to estimate the Mx using the maximum number of loci (*e.g.*, in a mixture where only one locus has four alleles, especially if this locus has the minor alleles in the stutter positions of the major alleles), an average Mx incorporating all 3-allele and 4-allele loci can be calculated by combining the aforementioned approaches. Figure 6 shows the observed Mx variation when based upon 3-allele and 4-allele loci and no contributor has been assumed.

Figure 6



Most values were still within a range of ±0.2 from the average, regardless of the average Mx, and all were within a ±0.3 range.

An average Mx based on using both 3-allele and 4-allele loci might be advantageous over one based solely upon 3-allele loci, but the differences observed here were subtle. Compared to a 4-allele locus Mx, allowing for the incorporation of 3-allele loci with an overlapping allele will improve the number of mixtures that can be tested while achieving a similar level of accuracy and precision.

Overall Mx Conclusions

Overall, the locus Mx variation generally varied within ±0.3 from the profile average Mx, with almost all of the data points falling between ±0.2 from the profile average.

An Mx average based solely upon 4-allele loci when there is no assumed contributor may be underestimated. Because the Mx is calculated for the minor contributor, the two peaks with the lowest RFU are used in the numerator and may not truly pair. This may lead to an underestimation of the true Mx. The effect of this is most pronounced when approaching an Mx of 0.5 as the minor contributor may truly be the major contributor at another locus. Nevertheless, the average Mx was shown to closely model the respective individual locus Mx values, with the majority of data points falling within a ±0.2 range.

In instances where it may be desirable to estimate the Mx using the maximum number of loci, an average Mx incorporating all 3-allele and 4-allele loci can be calculated by combining the different scenarios. Incorporating 3-allele loci into the average Mx will increase the number of mixtures that can be deconvolved.

Compared to the Identifiler study, where most of the typing was performed using the 310 Genetic Analyzer, the use of the 3130 for all Identifiler Plus studies may have led to some increase in the perceived variability in Mx. This could be due to the more sensitive 3130s detecting results from lower template quantity samples. Increased variability may also be attributed to the enhanced sensitivity of the Identifiler Plus kit.

Close examination of the plots shows TH01 having locus Mx values varying more from the profile average (TH01 often > average) than the other loci; the TH01 values were amongst the most variable in the study. This included the only observation outside the ±0.3 range. Nevertheless, because that example was a 3-allele locus with one shared allele, allowance for varied contributions to the shared allele in MixMaster brought this back into the ±0.3 range.

Based upon these mixtures studies, which involved intact DNA (*i.e.,* not degraded), it appeared to be acceptable to estimate the average Mx using the 4-allele, 3-allele, and 4&3-allele approaches, with or without an assumed contributor, when using a ±0.3 range. These approaches captured all but one of the STR results in the specified range. These results are similar to those previously obtained with the Identifiler kit. As with the Identifiler study, though, degradation and/or differential degradation could become severe enough to cause a true genotype pairing to fall outside the ±0.3 range.

Tiered Peak Height Ratios (PHRs)

Peak height ratios (PHRs) were calculated from 42 Identifiler Plus sensitivity studies compiled from BFS laboratories and run on 3130/3130*xl* Genetic Analyzers. This data is shown in Figures 7 and 8.

There were 351 instances of allelic drop-out (*i.e.,* one allele of a heterozygous pair at <50 RFU) observed in the data. The detected sister allele ranged from 209 down to 50 RFU. No instances of allelic drop-out were observed above the stochastic threshold of 365 RFU.

The sensitivity study data were grouped by RFUs of the taller peak and the minimum PHR to arrive at the following PHR threshold recommendations. Each PHR threshold was set to capture all of the observed data within each RFU range.

Taller Peak Height RFUs	Minimum Peak Height Ratio
1200 or greater	50%
750 to 1199	40%
365 to 749	30%
< 365	Any

Figure 7 ID Plus PHR vs Taller Peak RFUs

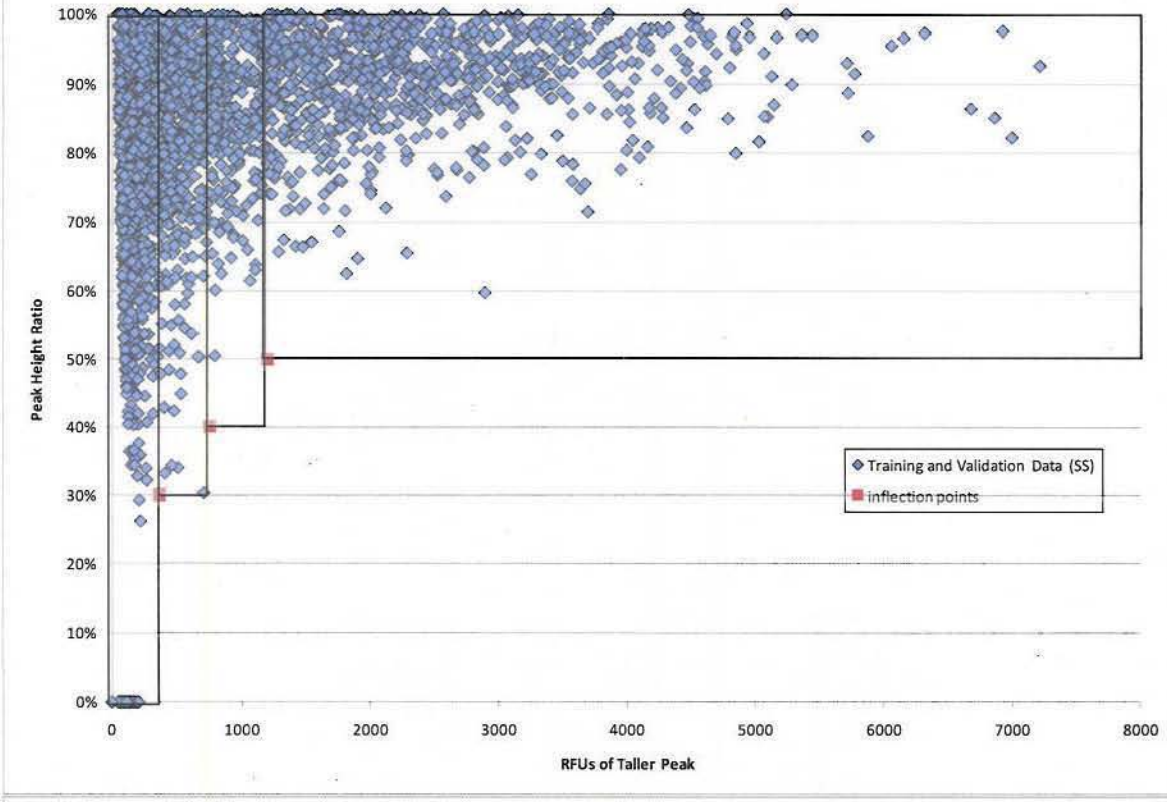


Figure 8 ID Plus PHR vs Taller Peak RFUs, Closeup, PHR ranges