JOURNAL OF FORENSIC SCIENCES

J Forensic Sci, September 2015, Vol. 60, No. 5
doi: 10.1111/1556-4029.12810
Available online at: onlinelibrary.wiley.com

PAPER

CRIMINALISTICS

Susan A. Greenspoon[1], *Ph.D.; Lisa Schiermeier-Wood*[1], *M.S.; and Brad C. Jenkins*[1], *M.S.*

Establishing the Limits of TrueAllele® Casework: A Validation Study

ABSTRACT: The limits of the expert system, TrueAllele® Casework (TA), were explored using challenging mock casework profiles that included 17 single-source and 18 two-, 15 three- and 7 four-person DNA mixtures. The sensitivity (ability to detect a minor contributor) of the TA analysis process was examined by challenging the system with mixture DNA samples that exhibited allelic and locus dropout and other stochastic effects. The specificity (ability to exclude nondonors) was rigorously tested by interrogating TA derived genotypes with 100 nondonor profiles. The accuracy with which TA estimated mixture weights of contributors to the two-person mixtures was examined. Finally, first-degree relatives of donors were used to assess the ability of the system to exclude close relatives. TA demonstrated great accuracy, sensitivity, and specificity. TA correctly assigned mixture weights and excluded nearly all first-degree relatives. This study demonstrates the analysis power of the TrueAllele® Casework system.

KEYWORDS: forensic science, DNA typing, expert system, DNA mixtures, probabilistic modeling, likelihood ratio

While even compromised single-source profiles typically lend themselves readily to human interpretation, mixture analysis poses a greater challenge for the forensic examiner. Forensic mixture samples, those biological specimens compromised of DNA from more than one individual, constitute a large proportion of casework samples. In fact, the level of sophistication and complexity of the analysis methods applied to DNA mixture sample interpretation has increased steadily, as has the complex nature of the sample types profiled by the forensic examiner (1–4). In 2010, the Scientific Working Group on DNA Analysis Methods (SWGDAM) recommended along with other guidelines that stochastic thresholds be applied to mixture samples (5). A stochastic threshold is designed to alert the DNA analyst that all of the DNA typing information may not have been detected for a given sample, that is, that there is potential for allelic dropout. Alternate statistical approaches were suggested to accommodate the uncertainty of the data.

Frequently, with the application of stochastic thresholds to DNA mixture sample electropherogram data, the combined probability of inclusion/exclusion (CPI/CPE) is rendered impotent as a means of expressing the statistical value of a profile due to loss of data below the stochastic threshold. Thus, it comes as no surprise that a number of software programs described as expert systems have been developed to assist the forensic examiner in performing scientifically based and statistically sound interpretations of the mixed contributor DNA evidence. Such an expert system would utilize much of the allele data that fall below the stochastic threshold (6–8). The testing and evaluation of one of these systems, TrueAllele® Casework (Cybergenetics, Pittsburgh, PA), is the subject of the study reported herein. This study was designed by and undertaken at the Virginia Department of Forensic Science (VDFS) to test the performance and define the limits of the TrueAllele® Casework expert system.

TrueAllele® Casework is a continuous probabilistic modeling system that utilizes Markov chain Monte Carlo (MCMC) sampling of the joint distribution, a probability distribution that combines all of the random variables, to perform an exhaustive statistical modeling of the electropherogram data (8,9). Probabilistic modeling as a means to deconvolve or solve a complex problem is not a new invention and has been successfully utilized by many diverse disciplines since its advent post-WWII (10). A wide range of disciplines such as nuclear physics, psychology, computer learning, economics, biological systems, and more recently, DNA analysis, utilize probabilistic modeling to make sense of the patterns observed in complex data and predict likely outcomes for various tests (11–13). Moreover, computer modeling can allow for the trialing of thousands or even millions of different explanations for the observed data using large numbers of variables within a time frame that escapes a purely human endeavor (14–17).

The TrueAllele® Casework system utilizes MCMC analysis in order to try many thousands of different combinations of variables to explain the DNA profile data. The short tandem repeat (STR) data are displayed in the form of an electropherogram generated as a final product of DNA profiling. Following Bayes' theorem, the observed data are separated into derived contributor genotypes which are used to update prior probability into posterior probability (9,18). TrueAllele® Casework can then answer the question of whether there is statistical support for or against the person of interest being a contributor to a mixture or single-source DNA profile. Moreover, this modeling of the data to generate derived contributor genotypes occurs prior to and independent of any comparison to the person of interest's reference

[1]Virginia Department of Forensic Science, Richmond, VA 23219.

Received 24 Feb. 2014; and in revised form 27 June 2014; accepted 26 Aug. 2014.

profile and thus provides an objective and unbiased analysis of forensic DNA casework.

All testing of the system was performed at VDFS using compromised single-source profiles and challenging two-, three-, and four-person mixed donor DNA profiles created on site. The PowerPlex® 16 (Promega Corp., Madison, WI) STR profiles of all samples utilized had been documented and all genotypes were known. Thus, a critical assessment of software performance could be performed as all genotyping answers to the questions posed to the system were previously established. Moreover, samples were chosen to stress the system allowing for an evaluation of the performance of the modeling program when confronted with samples exhibiting allelic and locus dropout, artifacts and many alleles below the stochastic and analytical (limit of detection) thresholds. These are the very artifacts routinely encountered when performing DNA analysis on forensic casework.

The TrueAllele® Casework software system, like any other instrument, has limits. This body of work was designed to identify those limits and from that data, formulate policy and procedure for accurate and reproducible forensic DNA mixture analysis. The sensitivity (ability to detect trace donors), the specificity (ability to exclude nondonors), and the ability to exclude first-degree relatives to donors of a mixture were deemed most crucial to define, but other aspects to sample analysis were also evaluated.

Materials and Methods

DNA Sample Preparation

DNA samples were purified from previously collected dried blood and buccal samples obtained from volunteers. DNA was extracted using the DNA IQ® System (Promega Corp.) or organic purification, as described (19). All samples are listed in File S1. DNA samples were quantitated using the Plexor® HY System (Promega Corp.) and amplified using the PowerPlex® 16 System as described (19). Amplified samples were separated on the 3130*xl* Genetic Analyzer (Applied Biosystems, Foster City, CA) as described (20). Analysis was completed by the GeneMapper® ID v3.2.1 software (ABI). The stutter cutoffs were defined as well as the limit of detection (LOD; blue 73, green 84, yellow 75, and red 52 RFUs) (20).

Electropherogram data (.fsa files) were utilized from previously analyzed single-source and mixture DNA samples as described in the sample preparation section above. Challenging single-source profiles were obtained from amplified DNA used for establishing stochastic thresholds and for environmental studies. Two samples with a high degree of heterozygosity originally used to establish the stochastic threshold were analyzed by TrueAllele® Casework (two 30 pg samples and three 10 pg samples from sample S9; three 30 pg samples and two 10 pg samples from sample S1). The profiles subjected to stochastic effects from the two different donors were compared to both donor S9 and S1 reference profiles to generate a match statistic. Eight degraded samples from three different donors (S1, S11, and S3) were analyzed using TrueAllele® Casework and then compared to the reference profile for the donor and ten nondonors to generate the match statistic.

Eighteen two-person mixture samples were obtained from previously analyzed mixture studies as well as mock casework. The mixtures samples were derived from different combinations of donors and by differing the ratios of DNA from the donors. A total of five different samples were used to create the eighteen two-person mixture profiles. Fifteen three-person mixture samples were obtained from previously analyzed mixture studies and mock casework samples. A total of seven different donors were used to create the three-person mixture profiles. As with the two-person mixtures, the profiles were derived from different combinations of donors and by differing the ratios of DNA from the donors. Seven four-person mixture samples were obtained from previously analyzed mixture studies. A total of eight different donors were used to create the four-person mixture profiles and they also consisted of different combinations of donors and differing ratios of the DNA from the donors. For all analyses except for the specificity tests, eleven reference profiles (S1–S11) were used for comparison and generation of the match statistics. All donors used to create the mixture samples were contained within the set of 11 reference profiles. The reference samples were previously typed using the PowerPlex® 16 System and uploaded to TrueAllele® Casework by manually entering them as text files.

The TrueAllele® Casework System

TrueAllele® Casework (TA) is a genotype modeling system that uses probability to define the most likely explanations for the data. TA uses MCMC analysis to examine many different variables in order to account for the observed data (with STR data, a sample's electropherogram peak height and molecular weight data) (8,9,17). Variables such as genotype and mixture weight (each contributor's proportion in the mixture), among others, are mathematically combined in probability equations modeled to explain the data.

Each TA cycle sequentially tests the variables to accept or reject values. When TA proposes a new value for a variable, it compares the joint probability (of data and model) using that new value relative to the old probability. For example, a cycle might compare the probability of a 50:50 mixture ratio for a two-person mixture relative to that of a 55:45 ratio. When the joint probability is higher, the new value is accepted.

The reported cycle numbers (25K, 50K, 100K, or 200K) refer to the number of times TA sequentially tests all of the variables (25K refers to 25,000 cycles, 50K refers to 50,000 cycles, etc.). For this study, the same cycle value was utilized for both "burn-in" and "read-out". The "burn-in" phase moves the system into the posterior probability region (e.g., mixture weight values that better explain the data). In the MCMC "read-out" phase, the system statistically samples from that region (e.g., determining the mixture weight probability distribution). TA analyses of a sample run for different cycle numbers can still be concordant and are evaluated using the same metrics. Some complex mixtures will be better resolved using a greater number of cycles, but running mixture samples longer typically impacts minor contributors much more than more predominant contributors (pers. obs.).

Production of the Match Statistic

After the derived contributors are produced by the TA software system, a comparison is performed between the derived contributors of a sample and reference samples of interest selected by the user. The comparison is in the form of a likelihood ratio (LR) and also referred to as the match statistic. This value can also expressed as a logarithm of the LR, log(LR) (20,21). The match statistic for a comparison with a particular reference profile is the log(LR) which is the highest, most discriminating value for that analysis of the evidence sample. A

positive match statistic is where the log(LR) is positive, which means the LR is above one. A comparison which provides no statistical support for a match is where the log(LR) is negative, which means the LR is below one. Concordant analyses should produce match statistics that are within two log units (ban) of each other. For the study, a positive log(LR) is referred to as an inclusion and a negative log(LR) is referred to as an exclusion.

TrueAllele® Casework analyses were performed at 25, 50, and 100K cycles for the stochastic samples and at 25K twice and at 100K once for the degraded samples, except for one sample, S11 (UV treated for 3 months), which was analyzed once at 25K and twice at 100K cycles. TA analyses for the two-, three-, and four-person mixtures ranged from 25K to 200K, although 25K was determined to provide inadequate sampling for the complex three- and four- person mixtures and was discontinued.

Operation of the System

The operation of the TrueAllele® Casework system (server v. 3.25.4441.1, VUIer v. 3.3.5148.1) was performed as described in the TrueAllele® VUIer™ manuals (22). Also utilized was the information and training provided by Cybergenetics for both Operator I and Operator II level training courses and in the literature. A theta correction value of 0.01 was employed for all analyses using VDFS allele frequencies.

Single-source profiles from degraded (7) and low template samples that exhibit stochastic effects (10) were analyzed using TrueAllele® Casework and compared to the donor reference as well as nondonor reference profiles. Two-, three-, and four-person mixtures were subjected to the TrueAllele® Casework analysis process and compared to a series of eleven reference profiles (named S1–S11 and including the true donors) for generation of the log(LR) match statistic. All single-source, mixture, and reference profile compositions are listed in File S1.

Evaluation of Data Output

The data produced by the TrueAllele® Casework system were evaluated for metrics listed in Table 1. The quality of the analysis which included the Markov chain sampling, the Gelman–Rubin convergence statistic value {$\leq$1.2, >1.2 and $\leq$1.5, >1.5} and histogram of derived mixture weights, was the initial quality aspects of the TA analysis evaluated after the software completed the deconvolution process (23). The MCMC provides a visible record and history of the statistical sampling of mixture weights for an analysis. Figs 1 and 2 display two independent analyses of a complex three-person mixture, Mix3_6. Fig. 1 depicts an ideal analysis and Fig. 2 depicts a nonideal (poor) analysis, given the complexity of the mixture.

TABLE 1—*Metrics assessed for TrueAllele® Casework Analysis. The three metrics listed are the first aspects to be assessed for a deconvoluted mixture sample.*

Metrics	Ideal	Acceptable	Poor
Markov chain (MC)	Good sampling of the "space". MCs with minimal/ no sampling for no more than ~20% of sampling time	MCs with minimal/no sampling for more than ~20% of sampling time	MCs stuck with no sampling. Rope-like appearance of the chain
MW Histogram	SD > 0.03 for complex mixtures	SD $\geq$ 0.03 for complex mixtures	SD < 0.03 for complex mixtures
Gelman-Rubin Convergence	$\leq$1.2	$\leq$1.5	>1.5*

MW, Mixture weight; SD, standard deviation.

*In some concordant samples, an analysis with poor convergence values may still be used for reporting.

After the initial assessment of the run metrics described above, other characteristics of the analyses were evaluated. The reproducibility of the results (genotype concordance and similar match statistics) was assessed. Reproducible match statistics were defined as a minimum of two ideal or acceptable analyses with the log(LR)'s within 2 ban (log units). Also assessed was if the correct individuals were included (generated a positive match statistic) and if nondonors were excluded (generated a negative match statistic). The derived mixture weights for concordant analyses should be similar, but do not need to be exact. For mixtures with very minor contributors (less than 15%), the mixture weights for the more minor contributors may show increased variability, even for concordant, ideal analyses. The Kullback–Leibler (KL) statistic (the information content of a derived contributor genotype) was also evaluated; however, it was not used for any determinations of concordance (24).

An example of good genotype concordance versus poor genotype concordance is shown in Fig. 3 for independent analyses of the same complex three-person mixture, Mix3_6. Excellent genotype concordance is depicted in Panel (a) with the predominant derived contributor for both the ideal and poor (nonideal) analyses. The correct genotype of the predominant contributor to the mixture is circled. Poor genotype concordance for the most minor derived contributor genotype (the correct minor contributor genotype is circled) is observed between the ideal and poor analyses. Genotype concordance that is deemed "fair" will typically fall between the two extremes of poor and good, showing good concordance at many loci and poorer concordance at other loci. The probability value assigned for each genotype is also considered when assessing the quality of the genotype concordance as more concordant genotypes typically display more similar genotype probabilities.

Typically for a poor TA analysis, the predominant derived contributor is concordant with the predominant derived contributor for an ideal analysis; however, a very minor contributor, as was the case for Mix3_6, may not be captured by the MCMC sampling process (personal obs.). As shown in Fig. 3, a lack of concordance was observed between the most minor derived contributor for the ideal and the poor analyses. A 100% probability for a nonconcordant genotype was produced for the most minor derived contributor for the poor analysis, whereas a distribution of genotypes was produced for the ideal analysis. The true genotype of the most minor contributor was included in that distribution for the ideal analysis (6,9.3). Thus, the poor analysis failed to capture the most minor contributor to the mixture due to insufficient sampling.

Mixture Ratio Assessment

Mixture weights for two-person mixtures were initially estimated based upon quantitation data and the input ratios of the quantitated DNA placed into the PowerPlex® 16 System amplification reaction. After generation of the electrophoretic data, manual estimates were created using loci for which there was no allele sharing between contributors (loci with four alleles visible or loci with two minor alleles and one major allele). The peak

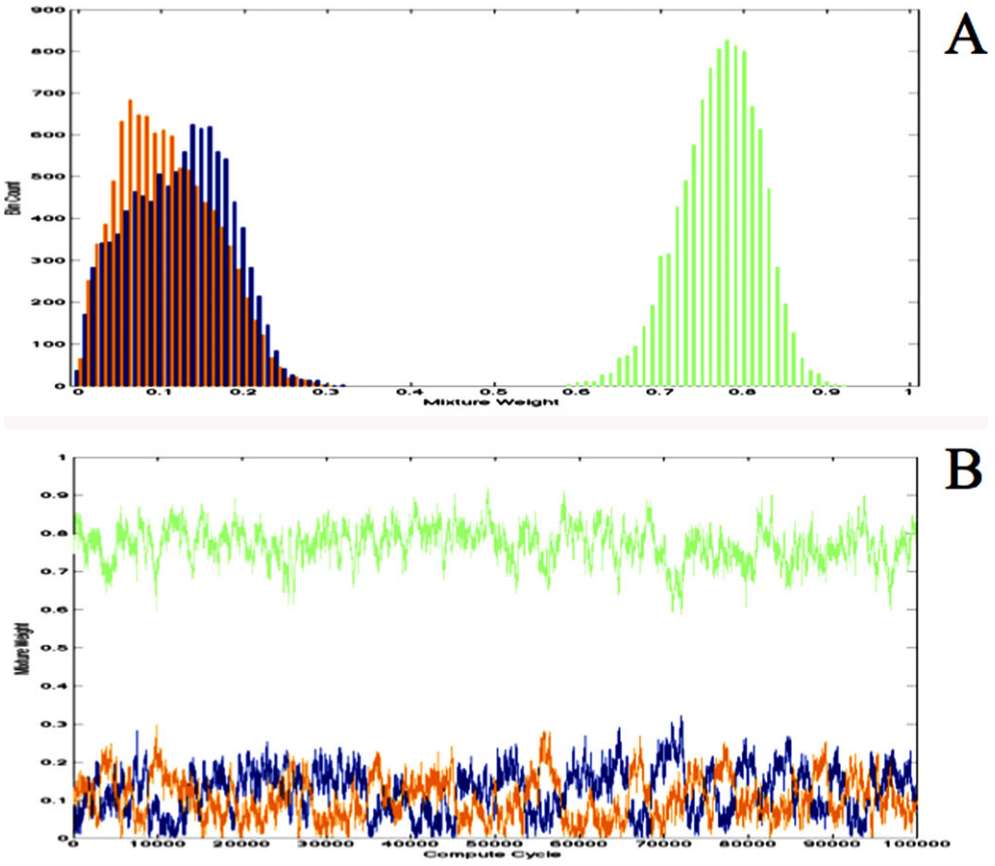

FIG. 1—*Markov chain and histogram of an ideal analysis of the complex three-person mixture, Mix3_6. (Panel A) Histogram of derived mixture weights for the three-person mixture. (Panel B) The corresponding ideal Markov chain history of the mixture weight sampling. The three colors indicate each derived contributor.*

height values for the minor alleles were summed and divided by the sum of the peak heights for all of the alleles.

Specificity of the TrueAllele® Casework System

Specificity, the ability to exclude noncontributors, of the TrueAllele® Casework analysis process was evaluated. The derived contributor genotypes of two-, three-, and four-person mixture samples were utilized for the test. Only ideal or acceptable TrueAllele® Casework analyses that were retained and used for genotype concordance were utilized for comparison with reference profiles.

All of the derived contributor genotypes from the two-, three-, and four-person mixture profiles were interrogated for the match statistic using 100 synthetically generated PowerPlex® 16 profiles kindly provided by Cybergenetics. None of the 100 profiles were donors to any of the mixtures tested. To form the reference profiles, a computer randomly sampled allele pairs at each locus from a representative human allele count database. The random profiles were saved as text files for subsequent upload to a TrueAllele® World and eventual match comparison.

The TrueAllele® Casework system allows the user to manage data in virtual worlds. A TA world will contain the STR data, interpretation requests and the MCMC joint distributions. The Cybergenetics representative population database (named CYB) is a multi-ethnic allele count database based on five thousand anonymous individuals (M. Legler, Cybergenetics, pers. comm.). The synthetically derived PowerPlex® 16 profiles were uploaded to TrueAllele® Casework as text files. Match statistics were performed for all three major population groups: Black, Caucasian, and Hispanic.

The ability of the TA system to distinguish relatives versus true donors to the two-, three-, and four-person mixture samples was assessed. Only first-degree relatives were tested, therefore, "sons" were manually created from seven of eleven reference profiles by selecting one of the reference profile alleles at each locus and randomly selecting a sister allele to create a "son". Of the eleven reference samples used for this validation study, ten of those were donors used for creation of the two-, three-, and four-person mixtures. Of the seven profiles that were used to synthesize "sons", six were donors to the two-, three-, and four-person mixtures. Match statistics for the mixture profiles were generated for all of the eleven reference profiles as well as the seven "sons".

Additionally, "brothers" were manually created from five profiles of donors to the two-, three-, and four-person mixtures. This was carried out by estimating the expected ratios given a sibling relationship of both alleles being shared, one allele shared and no alleles shared. The siblings were created in this manner to ensure that they shared many alleles and thus would challenge the TrueAllele® Casework system. Furthermore, the profiles of the references and the "brothers" were entered into Popstats (a module of the Federal Bureau of Investigation's CODIS software) to calculate a sibling index. All sibling indices surpassed the minimum of 33 used as an inclusion threshold at VDFS (25).

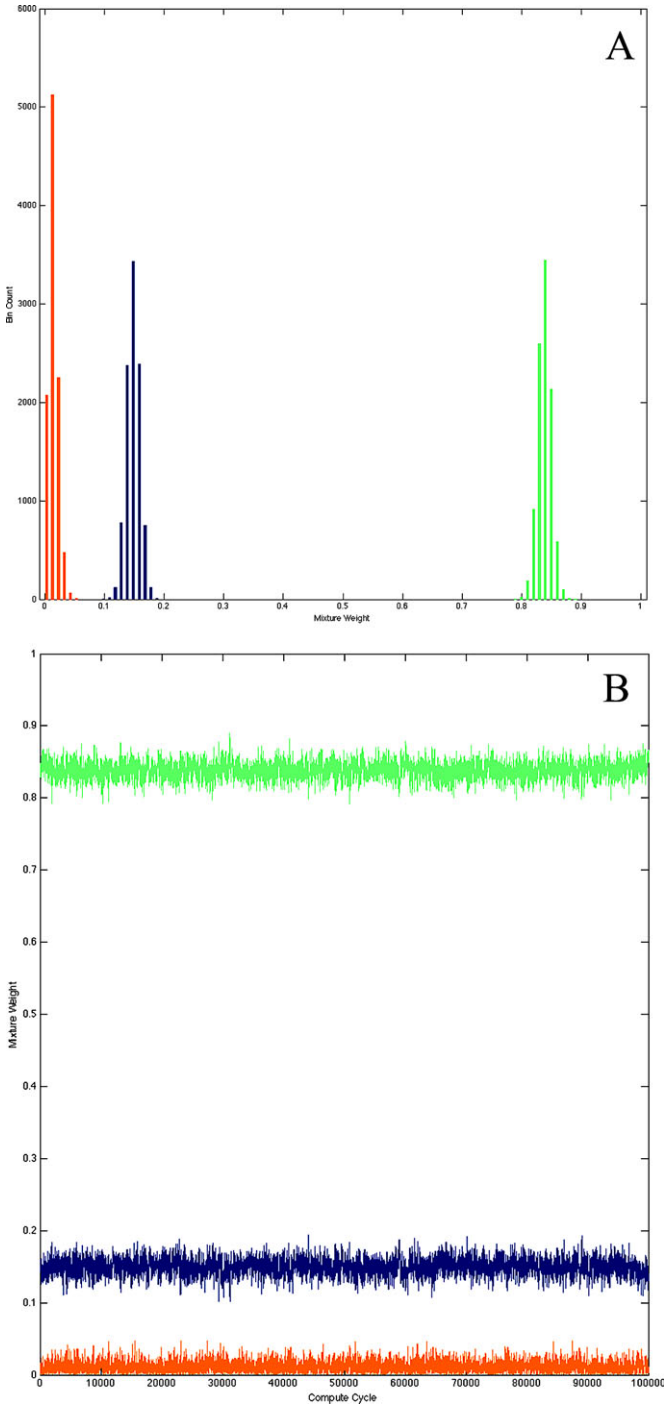

FIG. 2—*Markov chain and histogram of a poor analysis of the complex three-person mixture, Mix3_6. (Panel a) A nonideal histogram (the standard deviation is too small given the complexity of the mixture) of the derived mixture. (Panel b) The corresponding nonideal Markov chain history of the mixture weight sampling.*

The Use of Assumed Known Profiles

The use of assumed knowns was explored by analyzing seven different three-person mixture samples with TrueAllele® Casework and selecting one of the donor samples as an assumed known. Both the correct (assumed known was a donor to the mixture) and incorrect (assumed known was not a donor to the mixture) selection of an assumed known was tested. The match statistics produced when compared with eleven different reference profiles, of which three were the true donors in each mixture, were compared when no assumed known and when an assumed known was used.

Results and Discussion

Allelic Dropout, Locus Dropout, and Peak Imbalance: Single-source Samples

Eight degraded DNA samples were analyzed using TrueAllele® Casework (TA) and compared with their respective reference profiles for generation of the match statistic. Generally, there was a good correlation between the number of alleles observed both above and below the limit of detection (LOD) and the strength of the match statistic (Fig. 4, only S11 samples shown). The LOD values presented pertain solely to GMID as described in Materials and Methods. TA does not apply a LOD value, but instead samples electropherogram signal down to a selected value which was 10 rfu for analyses reported in this study. Baseline noise and peak uncertainty, which is proportional to peak height, among other variables are considered when the data are modeled (9,18,20). However, two samples provided negative log(LR) values when compared to their respective reference profiles (S11 UV and S3 80°C, S3 data not shown). Sample S11 subjected to 80°C produced a positive match statistic yet it displayed fewer alleles above and below the LOD than the sample S11 subjected to sunlight (referred to as UV exposed), which produced a negative match statistic. Thus, further investigation was necessary to determine the cause of such disparate match statistics.

Figure 5 displays electropherograms of the S11 samples incubated at 80°C and UV exposed at room temperature (RT). Six loci of S11 exposed to UV (Panel b) displayed allelic dropout (one allele of the heterozygous allele pair was not visible). Of these six, two loci showed the single visible sister allele below the LOD and unlabeled. The probability values ("p") generated by the TrueAllele® Casework analysis for the true heterozygous genotypes were all extremely low values, thus driving the overall match statistic lower than if neither allele of the heterozygote were present. However, TrueAllele® Casework was able to utilize allele data below the LOD, but distinguishable from baseline noise. An example of this is shown in Fig. 5, Panel (b) where an arrow points to two peaks at D21S11 that are imbalanced and below the LOD. The probability value for the 30,32.2 genotype at D21S11 was estimated at 0.8057. Another example is at the D7S820 locus in Panel (b) where both the 8 and 9 alleles are below the LOD, but TrueAllele® Casework assessed the probability of that genotype at 0.8878, thus demonstrating that TrueAllele® Casework was able to utilize more of the data than is currently available using a traditional threshold based approach. Conversely, the S11 sample subjected to 80°C (Panel a) did not display allelic dropout; instead, it displayed total locus dropout at multiple loci. The log(LR) match statistic (shown in the upper right hand corner of each panel) was significantly higher for S11 subjected to 80°C than subjected to UV even though fewer alleles were visible in the 80°C sample. This difference can be explained by the effect that false homozygotes had on the probability values for the heterozygote genotypes in the UV-treated sample.

Ten amplifications of two different samples (S9 and S1) using genomic template quantities in the stochastic range (30 pg and 10 pg) were analyzed using TrueAllele® Casework and

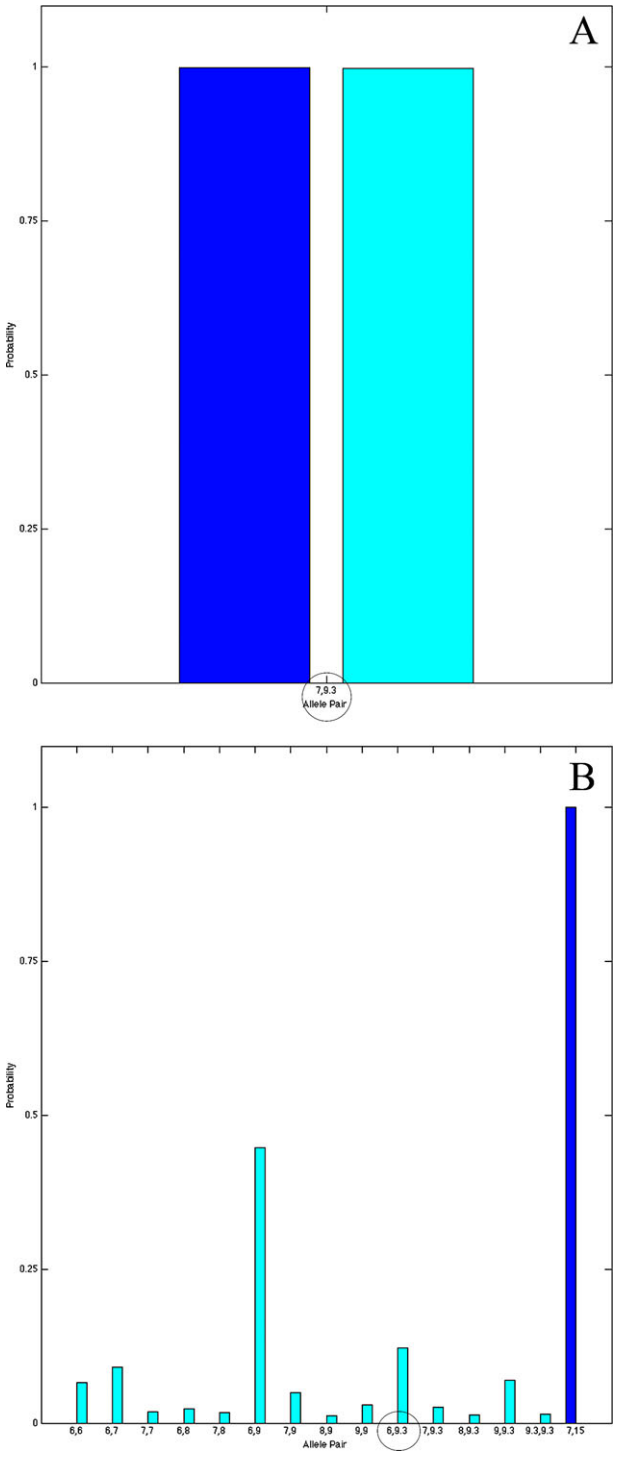

FIG. 3—*Genotype concordance at the TH01 locus for the complex three-person mixture, Mix3_6. (Panel A) Excellent and correct genotype concordance is observed between the predominant contributor for both the ideal and nonideal (poor) analyses shown in Figs 1 and 2. The dark blue column is the derived contributor genotype probability for the nonideal analysis, and the light blue column is the genotype probability derived by the ideal analysis. (Panel B) Poor genotype concordance is observed for the most minor contributor of the ideal and poor analyses. The dark blue column represents the derived contributor genotype probability for the nonideal analysis. It displays a 100% probability for a genotype that is not concordant with the minor contributor (true genotype is 6,9.3). The light blue columns show the derived contributor genotype distribution (due to uncertainty) for the most minor contributor of the ideal analysis. The correct genotypes for the most predominant and most minor contributors are circled (7,9.3 and 6,9.3, respectively).*

compared both to the donor reference profile and a nondonor reference profile for generation of the match statistic. A positive log(LR) was obtained when compared with the corresponding reference donor profile for all 30 pg samples tested, but negative log(LRs) were obtained for three of the five 10 pg samples (data not shown). An inspection of the electropherogram data for one of those 10 pg samples demonstrated the same phenomenon occurred as was described for the degraded samples: false homozygotes, due to allelic dropout, caused a dramatic reduction in the probability value down to zero for a heterozygote allele pair at those loci (data not shown).

Mock Casework Mixture Samples: Two-person Mixtures

Two-person mixture samples were utilized to evaluate how well TA includes the true donors to the mixtures (the sensitivity) and excludes nondonors (the specificity). Mixture samples were chosen purposefully to define the limits of the TA system. The contributor proportions varied from equal to a very tiny (less than 10%) contribution of the minor contributor. All metrics for the TA analyses were considered as described in Materials and Methods. Only the TA analyses that were deemed ideal or acceptable were utilized to assess genotype concordance between independent runs.

Eighteen two-person mixture samples (Mix2_1–Mix2_18) were analyzed with TrueAllele® Casework and interrogated using 11 references profiles. The derived contributors from the mixture profiles were compared to the true donor references and nine nondonor reference profiles. The quality of the TrueAllele® Casework analysis results was evaluated using the metrics as described in Table 1 and Materials and Methods. One requirement of the TrueAllele® Casework review process is to assess the reproducibility; thus, results were compared between two or more independent analyses of the same mixture that were deemed acceptable: Deconvolved mixture weights for the derived contributors were compared to ascertain whether or not they were similar in value, and genotype concordance for both contributors was assessed between analyses. A detailed description of the concordance for all two-person mixtures can be viewed in File S2.

The analyses of all of the eighteen two-person mixtures produced at least two ideal analyses. All analyses provided good or good/fair genotype concordance between the major contributors. Thirteen of the two-person mixture samples provided good or good/fair genotype concordance for the minor contributor. The minor contributor proportion of the mixture for the majority of those samples was greater than 15%, but less than 30%, so a clear distinction between the major and the minor contributors was possible. These samples also showed concordance for the other metrics, such as mixture weights and the log(LRs).

Five samples provided a fair or fair/poor genotype concordance for the minor contributor and for these samples, and the minor contributor proportion was less than 15%. Three of the five samples (Mix2_1, Mix2_3, and Mix2_7) failed to provide reproducible log(LR)'s for the minor contributor. However, Mix2_3 and Mix2_7 did provide consistent mixture weights for both the minor and major contributors. Mix2_1 failed to yield a positive log(LR) for the minor contributor; however, upon examination of the electropherogram, only two small alleles at D3S1358 and TH01 (144 rfu and 92 rfu, respectively) were observed that were solely attributable to the minor contributor (Fig. 6), and thus, the negative log(LR) appears to be appropriate.

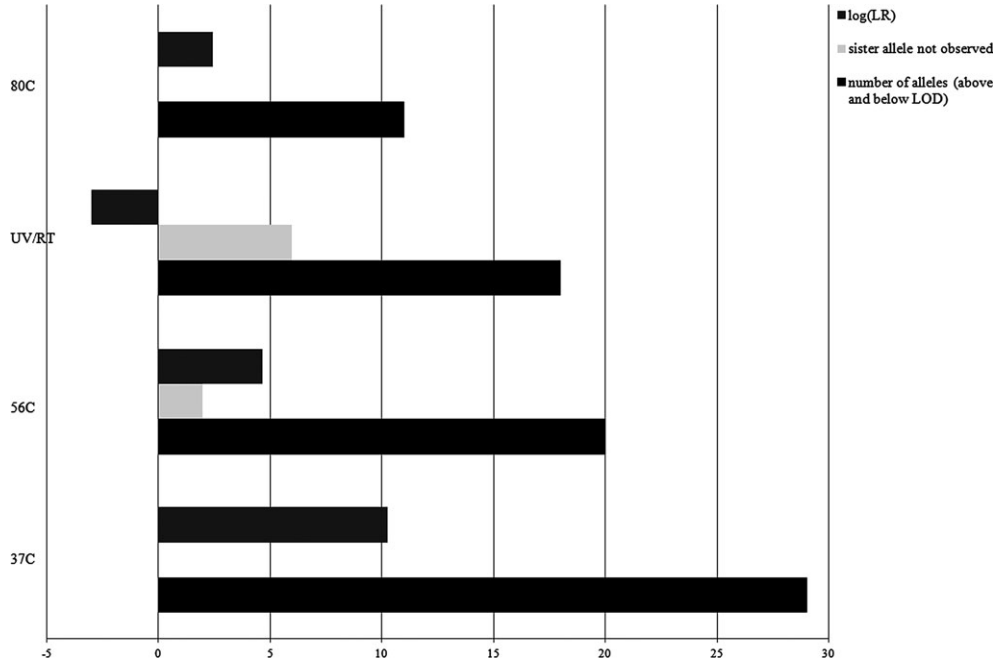

FIG. 4—*Relationship between the number of alleles above and below the LOD, the sister allele not observed and the log(LR) (match statistic). Only sample S11 is shown, and all analyses depicted were performed for 100K cycles. The maximum log(LR) for S11 (match probability) is 19.4927. All samples were incubated for 3 months. Key: 37C, 56C, and 80C = temperature incubated in degrees centigrade, UV = ultra violet light exposure at room temperature (RT).*

It should be noted that the three samples lacking good log (LR) reproducibility for the minor contributor had one analysis performed at 25K and the other at 100K. Thus, additional runs at 50K or more may be merited to produce more consistent match statistics for the minor contributor. The two-person mixture samples with a low-level minor contributor were deconvolved with great accuracy in that no nondonors were falsely included. The minor contributors displayed more genotype uncertainty, as would be expected with such low-level proportions.

Mixture Weight Accuracy

The accuracy with which TrueAllele® Casework deconvolutes mixture weights for two-person mixtures was assessed. Figure 7 displays a comparison between the targeted mixture weights of 17 mixture samples based upon the quantitation data, the estimated mixture weights assessed by manual calculation and the TrueAllele® Casework deconvolved mixture weight derivations. An inspection of the graph reveals that the manual and TrueAllele Casework derived mixture weights were extremely similar, but somewhat different from the targeted mixture weights based upon the DNA quantitation data.

No manual calculation was performed for the Mix2_5 sample as no clear minor contributor could be identified. The TrueAllele® Casework mixture weight value for the minor contributor was far from the targeted mixture weight for Mix2_5 (49% vs. 20%, respectively), but a review of the electropherogram data demonstrates that the TrueAllele® Casework derived mixture weight was more accurate as it is clear that the mixture was very close to a 1:1 combination of the two components (Fig. 8). The Mix2_8 and Mix2_9 samples were dehydrated and not re-quantitated, so the DNA concentrations were unknown.

Three-person Mixture Samples

Fifteen three-person mixture samples (Mix3_1–Mix3_15) were analyzed with TrueAllele® Casework and interrogated using 11 reference profiles; however, only ten of these mixtures were assessed for genotype concordance and reproducibility. The other five mixture samples were not repeatedly analyzed and thus were utilized solely for the specificity test. Detailed information about the samples can be found in File S1 and detailed assessments of the TA analysis for each sample can be viewed in File S3. The reference profile population contained the three donors for each of the mixtures in addition to eight nondonors. The quality of the TrueAllele® Casework analysis results was evaluated using the metrics as described in Materials and Methods.

The three-person mixtures present a far more complex analysis for either a human or the TrueAllele® Casework process. The three-person mixtures utilized were challenging and purposefully chosen for this study to assess the limitations of the TrueAllele® Casework process. Given the complexity of the mixture samples, the 25K cycle number was abandoned and those analyses are not included in the File S3. In general, when all of the metrics provided values within the desired ranges, for example, Mix3_4 (50K, 100K, and 100K2X runs), the concordance observed between runs was very good. Analyses that showed examples of the convergence value exceeding 1.2 were observed for all cycle numbers employed (50K, 100K, and 200K). This may indicate that longer sampling (more cycles) might be merited or it may be that the challenging nature of the mixture makes it recalcitrant to an ideal resolution, even at a very high cycle number. While convergence values below 1.2 are ideal, many examples of concordant runs were observed with higher than ideal convergence values.

Mix3_10 proved to be a challenging sample. Seven independent analyses were initially performed, consisting of five 100K

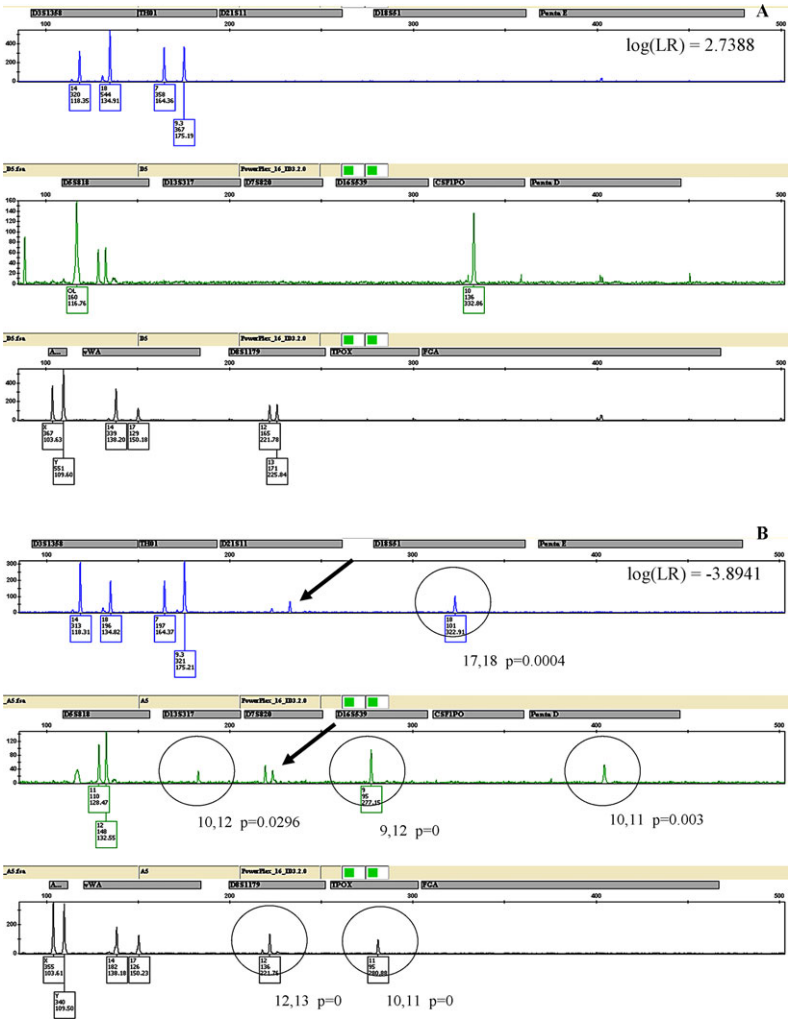

FIG. 5—*PowerPlex® 16 System typing data for sample S11 incubated at 80°C (Panel A) and for sample S11 incubated at RT with UV exposure (Panel B). Circled peaks indicate loci where the sister allele is missing. The correct, heterozygous genotype is indicated below and to the right of the peaks. Probability values ("p") for the true genotypes are adjacent to the genotypes. An arrow points to two peaks at the D21S11 locus that have poor allele balance (39%) and are both below the LOD. An arrow points to D7S820 where both peaks are below the LOD. The match statistic is displayed in the upper right-hand corner of panels (A) and (B).*

and two 200K runs. Only one of those seven provided an ideal analysis. Upon re-inspection of the electropherogram data, it was noted that a large spike at a size of approximately 399 bases was evident (data not shown). The allele calls associated with that spike were removed using the Request module of TrueAllele® Casework and the sample re-analyzed at 100K two times and once at 200K ("edited" appears in the name of the follow-up analyses, File S3). One of the 100K analyses and the 200K analysis provided concordant results. It was noted that the two concordant runs with the spike removed provided larger match statistics for the three contributors than the single ideal analysis that included the spike. This result is consistent with an increase in genotype certainty once the spike was removed.

In nine of the ten three-person mixtures, all nondonors for every ideal or acceptable and even poor analysis were excluded (consistently provided negative log(LR) match statistics). Mix3_6 did display a small positive match statistic for a noncontributor (S6; 3.057 times more likely {log(LR) 0.485}) for the under-sampled 50K analysis; however, this was rated a poor analysis prior to comparison with any reference samples and more importantly, this positive match statistic for comparison to S6 was not reproducible. The two ideal analyses provided log (LRs) of −1.0538 and −1.0291, reproducibly providing no statistical support for inclusion of the nondonor, S6 (data not shown). An examination of the electropherogram data demonstrates the selectivity of the TrueAllele® Casework analysis process as nearly every allele of reference S6 is shared with the Mix3_6 mixture profile (Fig. 9), yet no statistical support was generated for reference S6 as a contributor to the mixture.

Three-person Mixtures with an Assumed Known

The use of an assumed known for three-person mixtures was explored with respect to its effect on the TrueAllele® Casework analysis process. Assumed knowns are frequently utilized in forensic DNA analysis and mixture de-convolution as some samples, such as intimate ones, might reasonably be expected to contain DNA from the source of the sample (e.g., a vaginal swab would be expected to contain victim DNA). To demonstrate this effect, an assumed known was designated for one of the true donors for each of the seven-three-person mixtures selected for this demonstration. A minor contributing donor was chosen for

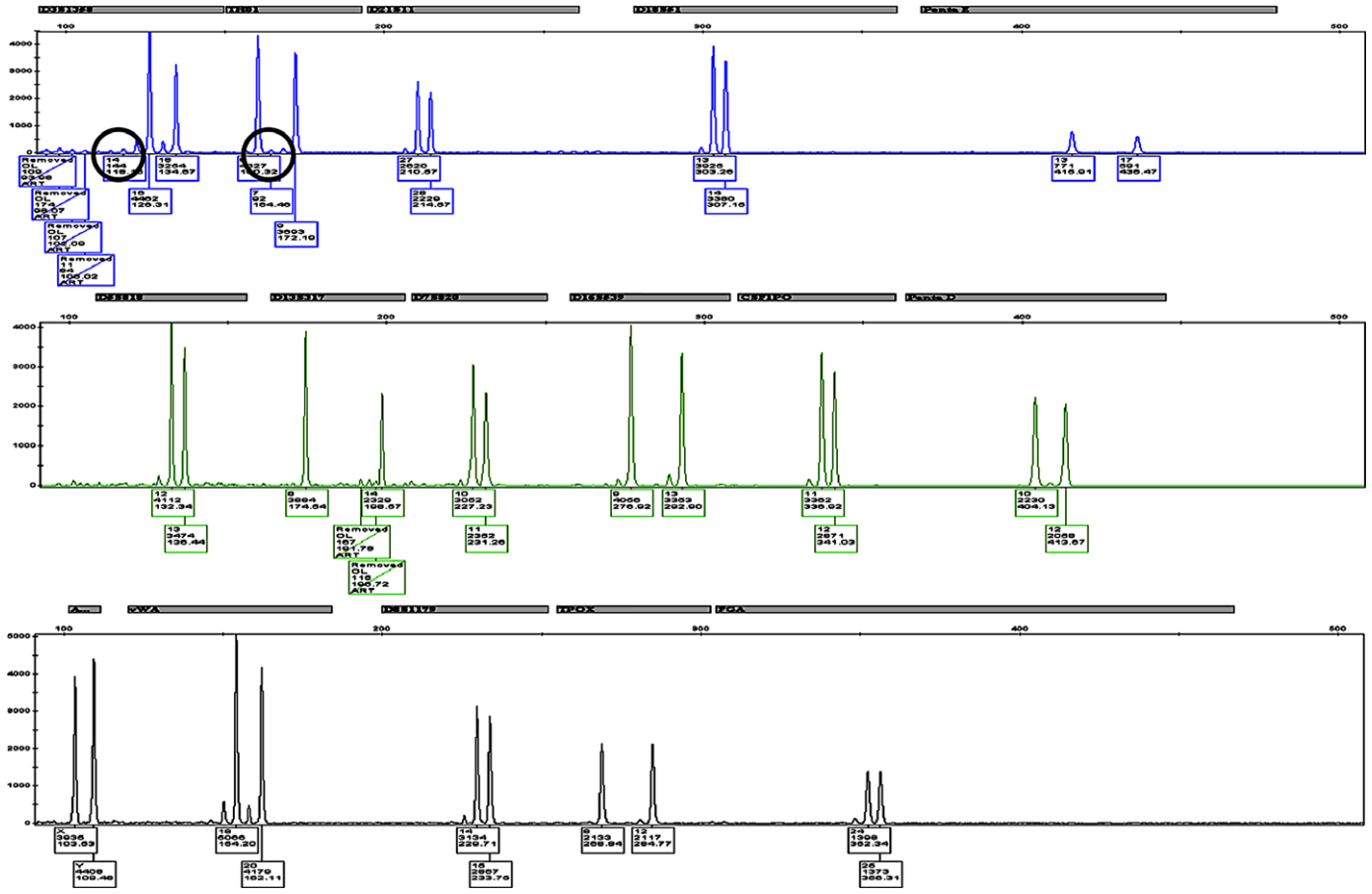

FIG. 6—*PowerPlex® 16 System profile of the Mix2_1 sample. Obligate alleles to the minor contributor are circled (14 at D3S1358 and 7 at TH01).*

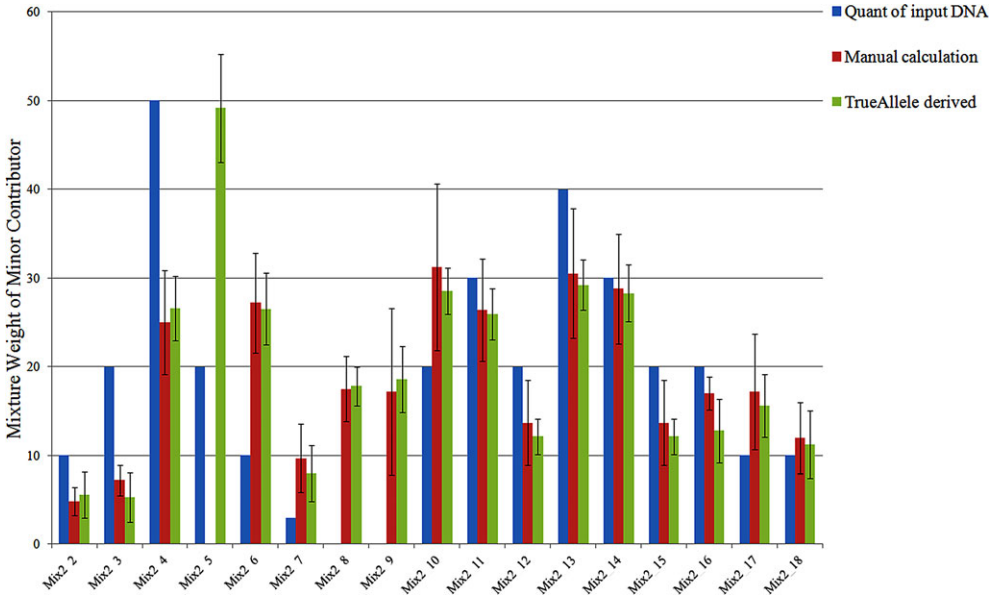

FIG. 7—*Accuracy of mixture weight assessment by TrueAllele® Casework for the minor contributor of two-person mixture samples. The "*n*" ranged from 2 to 9, with the average being 6.4 loci for manual mixture weight estimates.*

designation as an assumed known except for mixtures Mix3_3 and Mix3_4, where the predominant donor was designated. Table 2 provides examples of the use of a correct (individual was a donor to the mixture) and incorrect (individual was not a donor to the mixture) assignment of an assumed known. Mix3_4 and Mix3_8 were tested using assumed knowns that were actual

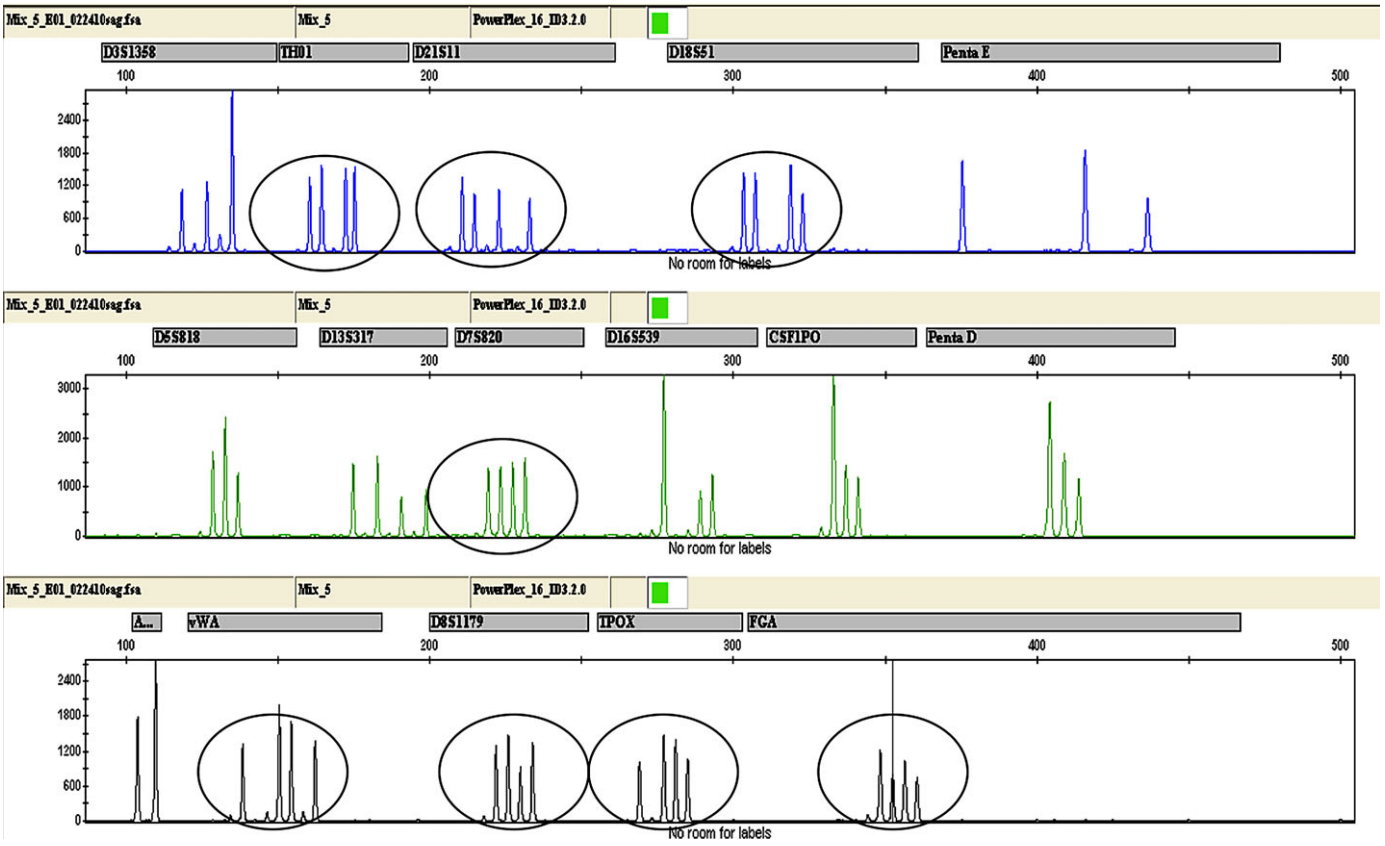

FIG. 8—*PowerPlex® 16 System profile of the Mix2_3 two-person mixture. Loci displaying four alleles nearly equal in height representing the 50:50 mixture are circled.*

donors and assumed knowns that were not donors to the mixtures. The log(LR) for an assumed known is the maximum value for that profile and is shown in bold.

In general, the use of a correct assumed known increased the match statistic for the remaining contributors and can increase the KL value (the information content of the derived contributors); however, for the Mix3_4, Mix3_6, and Mix3_7 samples, the log (LR) was only slightly changed. The impact of an assumed known is greater when teasing apart contributors similar in mixture weight or when applied to a minor contributor (pers. obs.). Deconvolved contributor genotypes with little to no uncertainty typical of a major contributor demonstrate only marginal gain in match score as a result of assigning an assumed known.

The question of whether or not the software could be confused by an operator error when assigned an assumed known was addressed by choosing a nondonor as an assumed known. An example of this is shown for samples Mix3_4 and Mix3_8. The impact on the log(LR) for the remaining minor contributors was a reduction in the value, but value for the predominant contributor was relatively unaffected. The use of an incorrect assumed known did not result in the inclusion of noncontributors to the mixtures among the eleven reference profiles tested (data not shown) nor in the exclusion of true donors.

Four-person Mixture Samples

Seven four-person mixtures (Mix4_1–Mix4_7) were analyzed using the TrueAllele® Casework system. Supplementary File 4 provides a detailed summary of the results. Although the 25K cycle number analysis was initially performed for these complex four contributor mixture profiles, 25K cycles were deemed insufficient and those analyses are not included in File S4. As with the three-person mixtures, the four-person mixtures required multiple analyses to produce reproducible and concordant results. Mix4_3 was a very challenging sample and eight independent analyses were performed generating four ideal or acceptable analyses. The concordance between the 100K2X, 200K, 200K3X, and 200K4X analyses was good except for the match statistic produced for the nondonor, S4, which fluctuated around zero giving small negative (−0.0536 and −0.0104, 200K and 200K4X, respectively) and small positive (0.686 and 0.0869, 100K2X and 200K3X, respectively) log(LR)'s. An examination of the electropherogram for Mix4_3 demonstrated that as with the three-person mixture (Mix3_6), the nondonor shared nearly every allele with the mixture profile (data not shown). The match statistic for S4 was not reproducible among the four analyses utilized for genotype concordance.

Of the seven samples analyzed, six provided at least two ideal and concordant analyses. The analysis of one sample, Mix4_1, did not produce more than one ideal analysis of the seven performed, so genotype concordance was not assessed. Two samples, Mix4_4 and Mix4_5, provided small yet reproducible negative log(LRs) for the most minor of the minor contributors. An examination of the electropherogram data provided an explanation for this statistical result as the mixture displayed allele drop-out at three or more loci, peaks below the stochastic threshold, masking of alleles, and alleles falling in the stutter position which corresponded to the minor contributor, S8 (data not shown). Given the complexity of the four-person mixture samples, additional analyses would be merited for casework samples

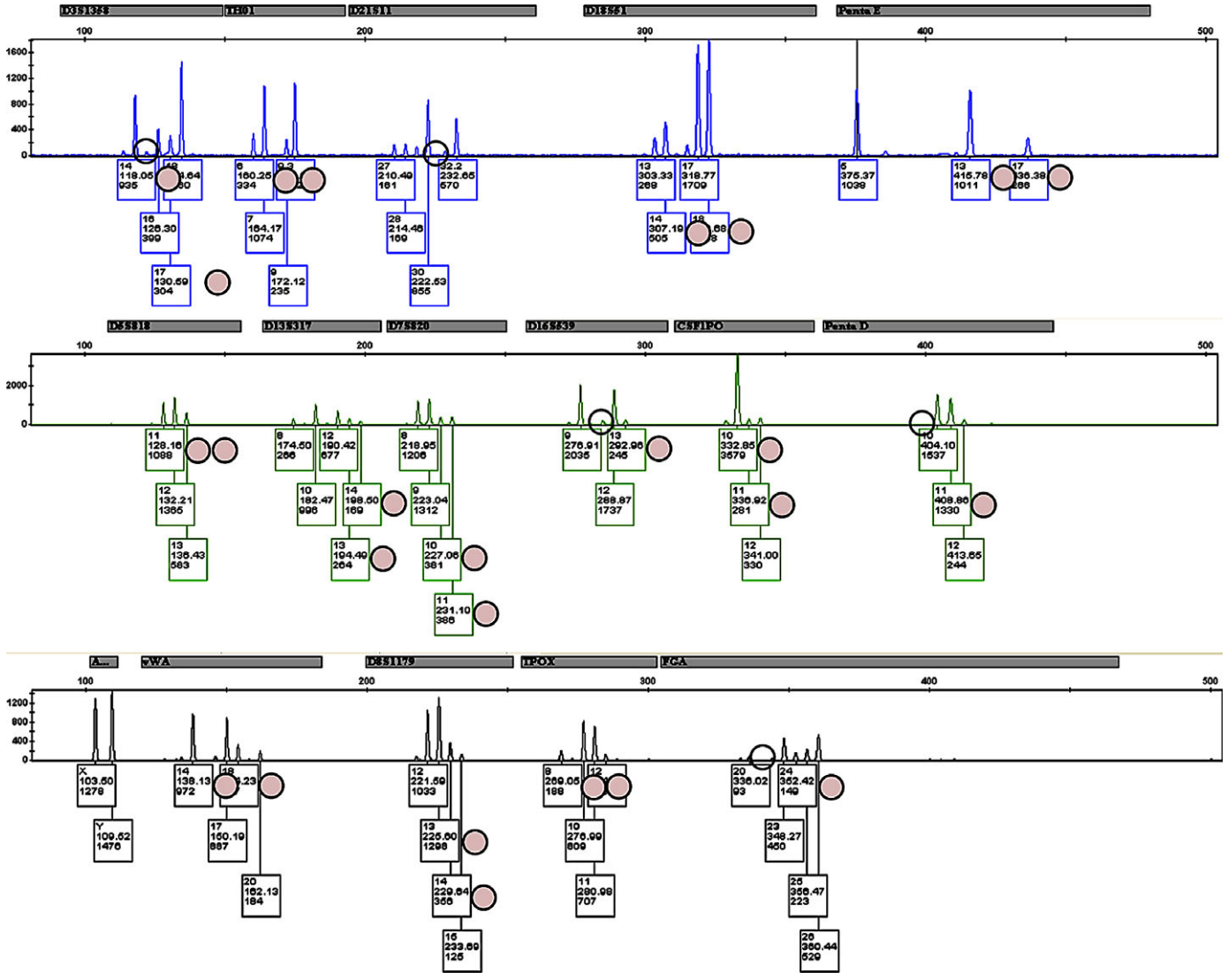

FIG. 9—*Mix3_6 PowerPlex® 16 System profile. The pink dots are adjacent to allele calls that are consistent with sample S6. Two dots indicate that S6 was homozygous for the allele. The black circles at the baseline encircle either a stutter allele consistent with an allele for S6 at that locus, or the position at which S6 would have an allele (i.e., no peak was observed in the Mix3_6 mixture profile).*

of similar intricacy; however, further testing was not conducted given that four contributor mixtures are not routinely interpreted at VDFS (18,19).

Specificity of Differentiating Relatives

The ability of the TrueAllele® Casework system to differentiate between closely related people was tested. First-degree relatives ("sons") were manually synthesized for seven reference profiles. Six of the reference profiles for which "sons" were created were donors to the mixture samples. Thus, it would not be unexpected to observe small positive match statistics for such close relatives. Only ideal or acceptable analyses of the two-, three-, and four-person mixtures were utilized for this test. The match statistic for all synthetic relatives was negative for the two-person mixtures (data not shown). Table 3 displays the log (LR) values generated for three- and four-person mixture analyses which produced a positive match statistic when compared with the "sons".

The analysis of three of the three-person and one of the four-person mixture samples produced derived contributors that resulted in positive log(LRs) when compared to a synthetic son of a donor to the mixture. Mix3_7 and Mix3_8 displayed reproducible small positive match statistics for a "son" of S1. While, the match statistic for the "son" of S1 was significantly lower than the match statistic for S1 in Mix3_8, it was approximately the same as S1 in Mix3_7. Reference S1 is the most minor contributor in both mixtures and exhibited allelic dropout at several loci (data not shown). There was also a nonreproducible small log(LR) produced for a "son" of S8 in Mix3_10-ed. The analysis of one-four-person mixture sample, Mix4_4, provided a positive match statistic for a "son" of a donor to the mixture. The positive log(LR) for the "son" was very small (1.002–1.79 times more likely). It is interesting to note that the donor (father of the son) was the most minor contributor to Mix4_4 and a reproducible negative log(LR) was produced by the analysis of the sample.

When synthetic "brothers" of contributors to the same two-, three-, and four-person mixture samples were compared, only one sample, three-person mixture Mix3_2, displayed a positive match statistic for the "brother". Only one analysis of Mix3_2 (analyzed at 100K2X) displayed a small positive match score of 1.0659 when the comparison to the synthetic brother of one of

TABLE 2—*The effect of an assumed known on the match statistic for three contributor mixtures. All samples were analyzed at 100K cycles. The assumed knowns used for Mix3_3 and Mix3_4 were the most predominant contributors, whereas a minor contributor was utilized for all others. The match statistic for the assumed known (maximum value) is shown in bold. The corresponding match statistic for the same donor when not selected as an assumed known in a different analysis is italicized. The match statistics for all three contributors are listed in order (e.g., S1, S5, S7) for both analyses (no assumed known and with an assumed known).*

Mixture Sample	Assumed Known	AK Donor to Mixture?	log(LR)
Mix3_1	No	–	*8.08*, 10.41, 19.45
	Yes	Yes	**16.33**, 13.21, 19.49
Mix3_2	No	–	−0.75, *11.01*, 12.13
	Yes	Yes	3.18, **20.96**, 15.85
Mix3_3	No	–	5.43, 10.73, *14.63*
	Yes	Yes	10.47, 11.23, **19.49**
Mix3_4	No	–	6.42, 5.50, *19.49*
	Yes	Yes	6.41, 5.96, **19.49**
	Yes	No	4.6, 1.23, *19.49*
Mix3_6	No	–	2.93, *11.85*, 19.47
	Yes	Yes	2.58, **20.96**, 19.49
Mix3_7	No	–	0.67, *3.65*, 18.59
	Yes	Yes	1.44, **21.96**, 18.6
Mix3_8	No	–	4.13, *6.13*, 17.57
	Yes	Yes	8.39, **21.96**, 18.25
	Yes	No	2.03, *5.94*, 17.69

AK, assumed known.

the donors was performed (data not shown). This was not reproducible.

Specificity

The specificity of the TrueAllele® Casework analysis process was more thoroughly addressed using 100 synthetic PowerPlex® 16 reference profiles, kindly provided by Cybergenetics, to compare with the derived contributors genotypes of two-, three-, and four-person mixtures. Multiple analyses of eighteen two person, fourteen three-person and seven-four-person deconvolved mixture samples were utilized. Only ideal or acceptable TrueAllele® Casework analyses were utilized. A total of 21,400 comparisons were performed for the derived contributors. No positive log(LR)s were produced for the comparisons performed for the two- and three-person derived contributors (data not shown). Of all of the derived contributors (214) for all of the analyses performed of the 39 total samples analyzed, only one provided a small (2.9 times more likely) and nonreproducible match statistic. Results for the most common match scores for the four-person mixtures are displayed in Fig. 10. The results of this test indicate that the TrueAllele® Casework analysis process is highly specific, even for complex three- and four-person mixtures.

Conclusion

The TrueAllele® Casework system accurately inferred problematic single-source sample profiles and generated positive match statistics. Generally, the greater the number of loci with alleles above the limit of detection, the more discriminating the match statistic; however, exceptions were observed if the single-source profile contained multiple false homozygotes. In those instances, a negative match statistic was observed due to a very low or zero probability being generated for the true heterozygote genotype at those loci. TrueAllele® Casework analysis was demonstrated to take advantage of additional information not utilized in a traditional threshold based analysis, such as alleles below the LOD and assigned probability values greater than zero to the correct genotypes.

Two-person mixture samples were easily resolved with the TrueAllele® Casework system with great specificity and discriminating match statistics unless the minor contributor was less than a 10% contributor to the mixture. When the minor contributor provided only a very small proportion of the DNA in the mixture, the match statistic reflected that weak contribution with

TABLE 3—*Three and four-person mixture samples which provided positive match scores for synthetic sons. Dark gray fill and bolded number indicates the match statistic for the donor included in the mixture (highest value generated for the comparison with the reference sample). Light gray fill and italicized number indicates a positive match statistic for a "son" of a donor to the mixture. S1–S11 refers to sample name. Not shown are the results for comparisons to "brothers" of the donors to the mixtures.*

	S1	S1_son	S4	S4_son	S5	S5_son	S6	S8	S8_son	S10
3 person mixtures										
Mix3_7 100K2X	**0.668**	−0.150	−9.323	−9.015	−5.840	−6.054	−9.515	3.387	−3.195	−8.424
Mix3_7 100K2X	−24.490	−26.157	−23.057	−25.163	**18.590**	−26.413	−26.890	−23.428	−26.870	−26.725
Mix3_7 100K2X	0.468	*1.342*	−9.659	−8.628	−5.851	−5.950	−10.009	**3.647**	−2.853	−6.734
Mix3_7 200K	**1.404**	*0.791*	−9.070	−8.012	−5.728	−5.712	−10.560	**3.672**	−2.743	−8.496
Mix3_7 200K	0.945	0.140	−9.410	−8.636	−5.775	−5.330	−9.748	3.605	−2.641	−8.203
Mix3_7 200K	−24.490	−26.188	−23.717	−25.835	**18.588**	−27.073	−26.565	−23.299	−26.878	−26.380
Mix3_8 100K2X	**4.129**	*0.234*	−10.648	−8.149	−3.554	−5.224	−7.045	**6.126**	−0.338	−6.790
Mix3_8 100K2X	−21.214	−21.625	−21.935	−17.852	**17.575**	−22.281	−25.652	−18.319	−21.712	−24.039
Mix3_8 100K2X	3.090	0.199	−8.567	−5.418	−4.499	−4.756	−6.088	5.033	−0.773	−7.022
Mix3_8 100K3X	3.341	0.182	−10.585	−7.533	−2.875	−5.385	−7.209	5.975	−0.538	−6.687
Mix3_8 100K2X	**3.864**	*0.818*	−11.299	−7.934	−0.678	−4.045	−8.987	**6.399**	−0.840	−7.254
Mix3_8 100K2X	−20.640	−18.904	−21.404	−16.901	**17.346**	−20.465	−25.347	−15.784	−19.532	−21.679
Mix3_10-ed 100K	0.906	−3.998	−18.764	−8.834	4.753	−5.347	−12.771	9.551	−0.223	−9.649
Mix3_10-ed 100K	−6.701	−8.054	−23.744	−13.919	**6.636**	−11.562	−19.780	**10.552**	−4.562	−11.529
Mix3_10-ed 100K	**5.904**	−1.487	−15.747	−9.026	2.377	−3.790	−10.635	6.174	*0.936*	−8.448
4 person mixtures										
Mix4_4 100K2X	−24.400	−24.337	−30	−30	−27.073	−27.073	−19.002	−27.145	−20.593	**17.159**
Mix4_4 100K2X	−8.983	−7.388	1.585	−6.232	−8.339	−5.727	2.539	−1.886	−0.111	−4.441
Mix4_4 100K2X	−8.285	−6.954	1.453	−6.981	−6.680	−6.364	2.859	−2.674	*0.001*	−4.439
Mix4_4 100K2X	−10.065	−7.549	**1.705**	−6.665	−7.805	−5.116	**3.099**	**−1.735**	−0.101	−4.475
Mix4_4 100K4X	−24.297	−24.337	−30	−30	−27.073	−27.073	−21.493	−27.145	−20.490	**17.262**
Mix4_4 100K4X	−8.128	−7.071	**1.446**	−7.189	−7.101	−5.990	2.530	−2.511	0.124	−4.550
Mix4_4 100K4X	−8.005	−6.789	1.376	−6.610	−8.198	−5.380	2.535	−1.998	−0.135	−5.003
Mix4_4 100K4X	−8.390	−7.476	1.314	−6.363	−7.079	−4.944	**3.170**	**−1.680**	*0.252*	−4.383

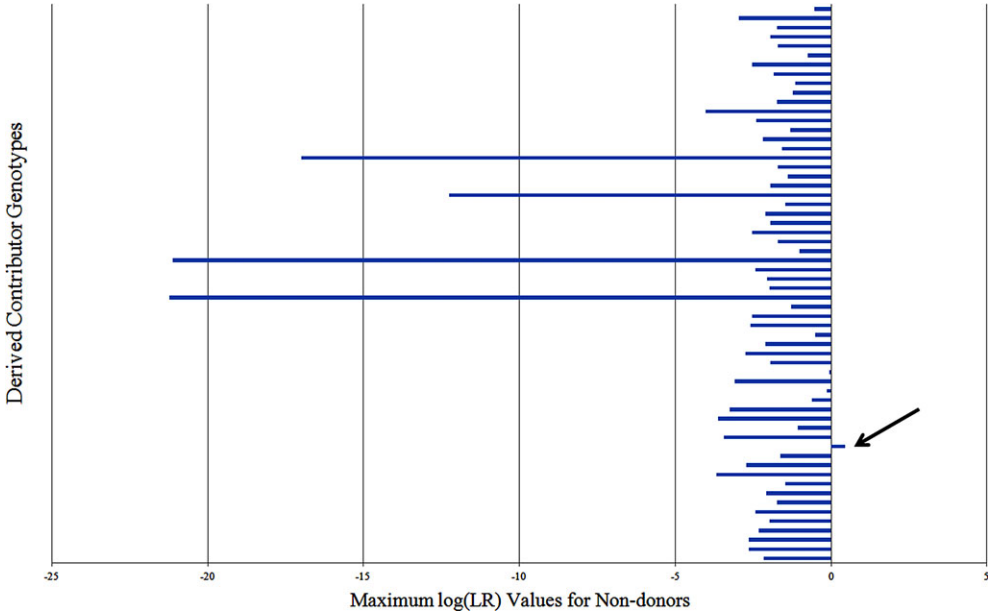

Fig. 10—*Maximum log(LR) values obtained by comparison of the derived contributors for the four-person mixture samples to 100 synthetic profiles (non-donors). Each bar represents a derived contributor of the four-person mixture samples. Only one derived contributor for mixture Mix4_3 (arrow) provided a positive log(LR).*

uncertainty resulting in lower match statistics. Of the eighteen samples, only Mix2_1 provided a negative match statistic for the minor contributor. An inspection of the electropherogram demonstrated that only two very small alleles were solely attributable to the minor contributor. All nondonors to the mixtures were definitively excluded (generated negative log(LR) values).

TrueAllele® Casework accurately assessed the mixture weights for two-person mixtures. When compared with the estimated mixture weights based upon DNA quantitation and template input quantities, TrueAllele® Casework provided a more accurate estimate based on an evaluation of the electropherogram data, comparable to the manually measured values calculated using peak heights.

Three- and four-person mixtures greatly increased the complexity and the genotype uncertainty of the analysis which was reflected in the match statistics for the minor contributors. For the 10 three-person mixture samples repeatedly analyzed using the TrueAllele® Casework process, none of the runs used for concordance provided a positive match statistic for a nondonor to the mixture. One sample (Mix3_6) provided a small (3.057 times more likely) match statistic for a nondonor, but it was not reproducible and only observed in a nonideal analysis (50K); thus, it could safely be excluded when drawing conclusions based on ideal runs for that sample. An inspection of the electropherogram demonstrated that the aforementioned nondonor shared nearly every allele with the mixture profile; therefore, the successful exclusion of the nondonor provides evidence supporting the specificity of the TrueAllele® Casework analysis process. One sample, Mix3_10, appeared recalcitrant to obtaining reproducible analyses. However, upon re-inspection of the electropherogram data, a large polymer spike was evident and once the allele information associated with that spike was deleted, additional ideal analyses were obtained with higher match statistics for the contributors, reflecting an increase in genotype certainty once the spike was removed.

Of the seven-four-person mixture samples repeatedly analyzed by the TrueAllele® Casework process, only one sample, Mix4_3, provided small, but nonreproducible positive match statistics for a nondonor. Ideal analyses at 100K and 200K (100K2X, 200K, 200K3X, and 200K4X) provided both positive and negative log(LRs) hovering around an uninformative log (LR) of zero for the nondonor. As with the Mix3_6 three-person mixture discussed above, the nondonor shared nearly every allele at all loci with the mixture profile and thus not unexpectedly, provided a difficult challenge for the TrueAllele® Casework analysis process. Analysis of two of the four-person mixture samples, Mix4_4 and Mix4_5, produced small, but reproducible negative log(LRs) for the most minor contributor. An inspection of the electropherogram data provided a reason for these exclusions as the donor displayed allelic dropout at multiple loci, masking of alleles, alleles in the stutter position, and alleles below the stochastic threshold. This demonstrates that TrueAllele® Casework analysis process requires sufficient evidential support for a true donor to derive a positive match statistic.

The use of an assumed known was explored with respect to its effect on the TrueAllele® Casework analysis process. Generally, the use of a correct assumed known, especially for a minor contributor, increased the match statistic for the remaining contributors by one or more bans and strengthened the KL value for the derived contributors; however, for some samples, the match statistic remained little altered. The use of an incorrect assumed known did reduce the match statistic for the true donors; however, it did not result in the inclusion of nondonors to the mixtures among the eleven reference profiles tested. Only a small study was conducted as it is unlikely that a nondonor would be selected as an assumed known.

Two person, three-person and four-person mixture runs were used to assess the ability of the TrueAllele® Casework system to differentiate between closely related people. First-degree relatives ("sons") were successfully excluded for all 35 mixture samples tested except for three-three-person and one-four-person mixtures. The positive match statistic for the "son" of a donor for two of the three-person mixtures was reproducible and small. The third example of a positive match statistic for a "son" of a donor to a three-person mixture was not reproducible. One

four-person mixture sample (Mix4_4) provided a positive match statistic for a "son" of a donor to the mixture; however, the LR was not reproducible and was small. When synthetic "brothers" were compared with the same two-, three-, and four-person mixture samples analyzed by the TrueAllele® Casework system, only a single analysis of one sample (Mix3_2), a three-person mixture, provided a small positive match statistic when compared to a "brother" of one of the donors to the mixture. That result was not reproducible. The potential for rendering a relatively small positive match statistic for a first degree relative of a contributor to a complex mixture is to be expected.

The specificity of the TrueAllele® Casework analysis process was tested using 100 synthetic PowerPlex® 16 reference profiles which were compared to the derived contributors of two-, three-, and four-person mixtures. Of the 214 derived contributors from the analyses performed, 21,400 comparisons were completed. Only one provided a very small and nonreproducible match statistic, indicating that the TrueAllele® Casework analysis process is highly specific, even for complex three- and four-person mixtures.

The TrueAllele® Casework process has been demonstrated to be sensitive and specific in its ability to include true donors and exclude or find no statistical support for nondonors. STR data displaying a great deal of allelic dropout and false homozygotes may produce a negative match statistic when compared to a true donor. This typically reflects the weakness of the profile and the conservativeness of the TrueAllele® Casework process.

Based upon this body of work and extensive training, the VDFS has implemented the use of TA in selected cases beginning in 2014. An inconclusive log(LR) range of ± 1 log unit (ban) has been established for casework analysis based on the specificity studies. Interestingly, a small number of cases where the human review of a mixture sample resulted in a finding of inconclusive with regard to the person of interest were also analyzed by TA and the match scores supported that finding (L. Schiermeier-Wood, pers. obs., 26). While the VDFS does not routinely interpret complex mixtures in which there is evidence supporting four or more donors, the analysis of four-person mixtures is not precluded. The studies reported herein demonstrate that even with complex four-person mixtures, TA is capable of performing an accurate, sensitive, and specific analysis.

Acknowledgments

The authors greatly appreciate helpful comments and guidance provided by Dr. Mark Perlin, Matthew Legler, and William Allan of Cybergenetics.

References

1. Clayton TM, Whitaker JP, Sparkes R, Gill P. Analysis and interpretation of mixed forensic stains using DNA STR profiling. Forensic Sci Int 1998;91:55–70.
2. SWGDAM. Interpretation guidelines for autosomal STR typing by forensic DNA testing laboratories, 2010. http://swgdam.org/Interpretation_Guidelines_January_2010.pdf (accessed June 12, 2014).
3. Gill P, Brenner CH, Buckleton JS, Carracedo A, Krawczak M, Mayr WR, et al. DNA commission of the International Society of Forensic Genetics: recommendations on the interpretation of mixtures. Forensic Sci Int Genet 2006;160:90–101.
4. Gill P, Buckleton J. A universal strategy to interpret DNA profiles that does not require a definition of low-copy-number. Forensic Sci Int Genet 2009;4(4):221–7.
5. http://www.fbi.gov/about-us/lab/biometric-analysis/codis/swgdam.pdf (accessed May 21, 2014).
6. http://www.nichevision.com/index.php/forensics/armedxpert (accessed May 21, 2014).
7. http://strmix.com (accessed May 21, 2014).
8. Perlin MW, Szabady B. Linear mixture analysis: a mathematical approach to resolving mixed DNA samples. J Forensic Sci 2001;46:1372–7.
9. Perlin MW, Legler MM, Spencer CE, Smith JL, Allan WP, Bellrose JL, et al. Validating TrueAllele® DNA mixture interpretation. J Forensic Sci 2011;56:1430–47.
10. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of state calculations by fast computing machines. J Chem Phys 1953;21(6):1087–92.
11. Richey M. The evolution of Markov chain Monte Carlo methods. Am Math Mon 2010;117:383–413.
12. Poirer DJ. The growth of Bayesian methods in statistics and economics since 1970. Bayesian Anal 2006;1(4):969–80.
13. Gelfand AE, Smith AFM. Sampling-based approaches to calculating marginal densities. J Am Stat Assoc 1990;85:398–409.
14. Gill P, Haned H. A new methodological framework to interpret complex DNA profiles using likelihood ratios. Forensic Sci Int Genet 2013;7(2):251–63.
15. Perlin MW, Belrose JL, Duceman BW. New York State TrueAllele® Casework validation study. J Forensic Sci 2013;58(6):1458–66.
16. Taylor D, Bright J, Buckleton J. The interpretation of single source and mixed DNA profiles. Forensic Sci Int Genet 2013;7:516–28.
17. Curran J. A MCMC method for resolving two person mixtures. Sci Justice 2008;48(4):168–77.
18. Perlin MW, Sinelnikov A. An information gap in DNA evidence interpretation. PLoS ONE 2009;4:e8327.
19. http://www.dfs.virginia.gov/wp-content/uploads/2013/07/210-D1100-Forensic-Biology-Section-Procedures-Manual-Section-III.pdf (accessed May 21, 2014).
20. http://www.dfs.virginia.gov/wp-content/uploads/2013/07/210-D1100-Forensic-Biology-Section-Procedures-Manual-Section-VIII.pdf (accessed May 21, 2014).
21. Perlin M. Explaining the likelihood ratio in DNA mixture interpretation. Proceedings of the Twenty-First International Symposium on Human Identification; 2010 Oct 11-14; San Antonio, TX. Madison, WI: Promega Corporation, 2010.
22. TrueAllele® VUIer™ manuals (analyze, data, request, review and report modules). Pittsburgh, PA: Cybergenetics.
23. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. Stat Sci 1992;7(4):457–511.
24. Kullback S, Leibler RA. On information and sufficiency. Ann Math Stat 1951;22(1):79–86.
25. http://www.dfs.virginia.gov/wp-content/uploads/2013/07/210-D1100-Forensic-Biology-Section-Procedures-Manual-Familial DNA Testing.pdf (accessed May 21, 2014).
26. Perlin MW, Dormer K, Hornyak J, Schiermeier-Wood L, Greenspoon S. TrueAllele casework on Virginia DNA mixture evidence: computer and manual interpretation in 72 reported criminal cases. PLoS ONE 2014;9(3):e92837.

Additional information and reprint requests:
Susan A. Greenspoon, Ph.D.
Virginia Department of Forensic Science
700 N. 5th St.
Richmond, VA 23219, USA
E-mail: susan.greenspoon@dfs.virginia.gov

Supporting Information

Additional Supporting Information may be found in the online version of this article:

File S1. Table listing all of the samples used for testing the TrueAllele® Casework software program except for the synthetically created profiles obtained from Cybergenetics (100 profiles uploaded to the system as text files).

File S2. Two person mixture results.

File S3. Three person mixtures.

File S4. Four person mixtures.