Further Exploration of TrueAllele® Casework

Susan Greenspoon, Ph.D., Lisa Schiermeier-Wood, M.S. and Brad Jenkins, M.S.

Virginia Department of Forensic Science, 700 N. 5th St., Richmond, VA 23219

Abstract

The Virginia Department of Forensic Science (VDFS) began casework analysis with TrueAllele® Casework (TA) in January of 2014. TA is used for analysis of complex mixtures comprised of up to four contributors on the most challenging of mixtures encountered by the laboratory. Since completion of validation, additional tests have been performed. These studies examined: 1) What effect does the use of the differential degradation feature have on the log(LR) of contributors to a differentially degraded mixture? 2) What happens to the TA analysis when either a greater or fewer number of contributors is solved for than are actually in the mixture? and 3) What happens to the log(LR) of contributors when the DNA sample is over-amplified/loaded?

The use of the differential degradation feature produced only a small (~1 log(LR) unit) change, if any, for contributors in differentially degraded mixtures. However, it may have affected how readily the sample was separated, resulting in fewer computer runs. This improvement may be due to a more accurate assessment of the mixture weights when differential degradation is taken into account.

When a greater number of contributors was hypothesized for TA analysis than was in the mixture, typically there was a small reduction, if any, effect on the log(LR) values generated. When a smaller number of contributors was hypothesized than was in the mixture, it often dramatically reduced the log(LR) of donors. This is consistent with how the TA modeling works; restricting the number of contributors also limits the potential genotype combinations that can explain the data, which may produce a reduction in the log(LR). Providing a greater number of potential contributors does not restrict the genotype combinations.

Without manual de-selection of excessive artifact peaks prior to TA analysis, the log(LR)s produced for contributor comparisons were typically decreased for over-amplified/loaded samples. The impact on log(LR) typically affected more minor contributors, rather than the predominate contributor. This result is consistent with how the TA process works; over-amplified/loaded samples exhibit excessive artifact which increases genotype uncertainty and thus can reduce the log(LR) values.

The results of the additional studies were consistent with the earlier work performed at VDFS and what has been reported in the literature for TA.

Introduction

TrueAllele® Casework (TA) is a continuous probabilistic modeling system that utilizes Markov chain Monte Carlo (MCMC) sampling in order to produce genotype probabilities for short tandem repeat (STR) data (1,2). Widely ranging disciplines such as physics, psychology, computer learning, economics, biological systems and DNA analysis, utilize probabilistic modeling to make sense of the patterns observed in complex data and predict likely outcomes for various tests (3,4,5). Following Bayes' theorem, the observed data is separated into derived contributor genotypes which are used to update prior probability into posterior probability (6,7). After modeling has been performed by TA, a comparison with reference profiles is performed, likelihood ratios calculated and are reported in logarithmic form, log(LR) (8). TA can then answer the question of whether there is statistical support for or against the person of interest (POI) being a contributor to a DNA profile.

The VDFS has been utilizing TA for analysis of complex casework profiles, however, the need for further understanding of the TA process was identified and additional testing has been performed and reported herein (9,10).

Methods and Materials

DNA Sample Preparation, Quantitation, STR Amplification, Capillary Electrophoresis (CE) and Detection – DNA was purified manually using the DNA IQ™ System (Promega Corp., Madison, WI) according to the manufacturer's protocol with minor modifications or robotically utilizing the Biomek® NXP Automation Workstation (Beckman Coulter, Inc., Fullterton, CA) as described (9). Samples were quantified with the Plexor® HY System using the Stratagene MX3005P as described (9). Mixtures were created as described (10). Mixture samples comprised of degraded DNA from one or more of the contributors were created using DNA samples that had been previously assessed for quality by STR profiling. DNA samples were amplified using the PowerPlex® 16 (Promega) STR amplification kit following manufacturer's recommendations, with minor modifications as described (9). Amplified samples were separated on the 3130*xl* Genetic Analyzer (Applied Biosystems, Foster City, CA) and analysis was performed using GeneMapper® ID v3.2.1 software (ABI) as described (9).

TrueAllele Casework analysis of CE data – .fsa files generated by the 3130*xl* CE analysis of the DNA samples were uploaded to the TA server following the procedure described in the TA User Manuals (8). Analyses were evaluated after TA mixture separation as described (9,10). An assessment was performed as to whether a TA analysis was satisfactory or not prior to any comparison to a reference sample. Only satisfactory analyses were utilized.

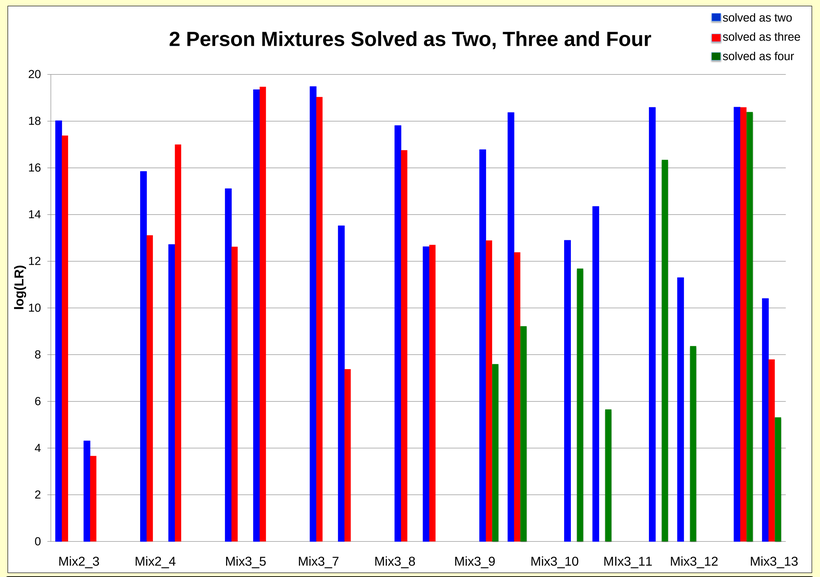
Results and Discussion

Sample	Number of Contributors	Original Match Score log(LR)	Match Score Using Differential Degradation Feature
1:10_S5:**S12**	2	9.559, **16.8116**	9.6849, **16.8516**
1:1_S5:**S12**	2	12.8887, **9.4711**	13.6069, **10.3843**
1:2.5_**S16**:S13	2	**-9.5441**, 20.9574	**-9.8209**, 20.9721
1:5_**S16**:S13	2	**-11.5818**, 20.9758	**-11.1655**, 20.9692
1:5_S5:**S12**	2	10.6646, **13.5557**	10.7905, **13.50**
1:10_**S16**:S13	2	**-13.1203**, 20.9725	**-10,4826**, 20.9312
1:2:1_**S14**:S13:S5	3	**-2.4908**, 20.8852, -1.8018	**-2.0581**, 20.8444, -0.5637
1:2:1_**S16**:S13:**S12**	3	**-10.1849**, 20.7741, 6.4646	**-9.2563**, 20.8886, 7.2565
1:3:1_**S14**:S13:S5	3	**-6.3185**, 20.959, 8.0922	**-4.8192**, 20.9537, 7.5003
1:2:2_**S14**:S13:S5	3	**-3.584**, 17.3762, 8.9592	ND**
1:1:1:1_S5:S15:S13:**S12**	4	4.2744, 0.6413, 10.6804, **2.3843**	4.5812, -0.1181, 9.4423, **1.1394**
1:2:2:2_**S16**:**S14**:S13:S5	4	ND*	**-7.8542**, 0.5242, 19.768, 9.3933
1:3:2:1_S12:S15:S13:**S14**	4	ND*	2.5992, 3.2491, 17.7018, **1.4099**

Table 1. The effect of the differential degradation feature on the log(LR) when the derived contributors were compared to their respective reference profiles. Degraded samples (in bold) have different levels of degradation (S12<S14<S16), with S16 being the most degraded sample (extremely degraded).
Key: S# = Sample name (ratios listed indicate approximate ratios of the contributors); log(LR) in bold = degraded sample.
No satisfactory analysis obtained out of two runs (ND*) or three runs (ND**) performed at the settings used.

Test of Differential Degradation Feature

Two, three and four person mixtures were created containing at least one degraded contributor DNA sample. This was designed to test the TA differential degradation feature (degrd.). Table 1 displays the log(LR) values of TA analyses performed with and without the use of the differential degradation feature. The use of degrd resulted in little to no difference in the log(LR) values generated when derived contributors genotypes were compared to contributor reference profiles. However, the use of degrd appeared to increase the number of satisfactory runs with fewer analyses required (16 satisfactory runs using degrd, 4 unused runs; 11 satisfactory runs not using degrd, 9 unused runs; data not shown).



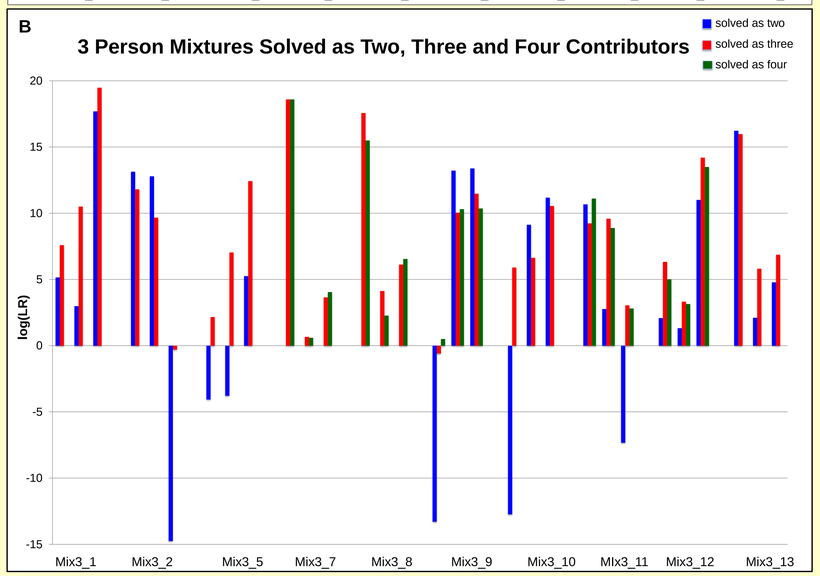


Figure 2. Two and three person mixtures solved as two, three and four person mixtures. Panel A (two person mixtures) and Panel B (three person mixtures) display the log(LR) values for the contributors when solved as two, three and four person mixtures. Each bar of a color corresponds to a derived contributor and the log(LR) generated when compared to its corresponding reference profile. Not every mixture was tested for the three different numbers of contributors.

Mixture Analysis Testing Differing Numbers of Contributors

In order to ascertain the effect that hypothesizing the incorrect number of a contributors to a mixture would have on the TA analysis, two, three and four person mixtures were analyzed using the correct and incorrect number of contributors. Figure 2 displays the effects on the log(LR) when a two person mixture was solved for two, three and four contributors (panel A). Panel B shows the effects on the log(LR) when a three person mixture was solved for two, three and four contributors. Four person mixtures solved as three and four contributors are not shown. Hypothesizing a greater number of contributors than is in the mixture typically had little effect or reduced the log(LR). Hypothesizing too few either had little effect or resulted in a pronounced reduction in the log(LR) for less predominant contributors. More predominant contributors sometimes showed an increase in log(LR) (Figure 2, Panel B, Mix3_9).

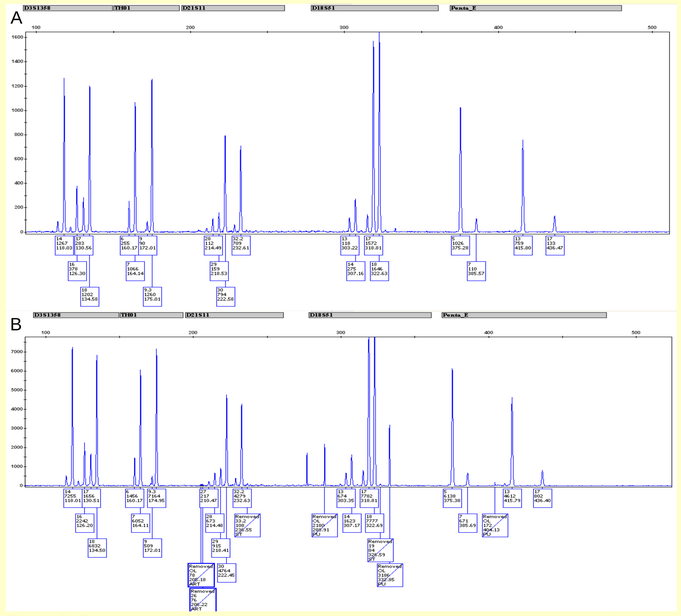


Figure 3. Three person mixture which has been analyzed at both normal and over-amplified/loaded Levels (only Fam channel shown). Panel A. Mix3_4. Panel B. Mix3_4 over-amplified/loaded data.

Analysis of Over-amplified/loaded Samples

Over-amplified/loaded STR profiles (Figure 3, Panel B) for four different three person mixtures were previously identified when amplified samples were analyzed using routine DNA profiling conditions (e.g. 1 µL amplified product injected for 5 sec.; 11). All samples were re-loaded and analyzed under different conditions (0.5 µL amplified product injected for 2 sec.) to eliminate artifacts (Panel A). Analysis by TA resulted in a reduction in the log(LR) for the contributors (S1, S9 and S11) for all over-amplified/loaded mixture samples when compared to "normal" profiles (minimal artifacts, no off-scale data). However, S11, the most predominate contributor, showed little to no change in log(LR) except for Mix3_5 (Table 2).

Sample	Over-amp/load	Contributor S1	Contributor S9	Contributor S11
Mix3_1	No	8.0861	10.4196	19.4898
	Yes	6.89.58	9.0601	19.3052
Mix3_4	No	6.4262	5.5026	19.4923
	Yes	5.0110	2.4905	19.1883
Mix3_5	No	3.0779	7.8335	10.7062
	Yes	-0.2888	2.0048	7.8663
Mix3_6	No	2.9363	11.8571	19.4701
	Yes	3.01621	9.9005	18.7164

Table 2. The effect of un-edited excessive PCR artifacts on the log(LR) of true contributors when the over-amplified/loaded samples were compared with normal profiles.

Conclusions

The assessment of the differential degradation feature, the hypothesis of the incorrect number of contributors and the use of over-amplified/loaded samples demonstrated that the principles by which TrueAllele® Casework statistically models mixture profiles remained consistent. Greater uncertainty (such as excessive PCR artifact) can reduce the log(LR) of true contributors when compared to the analyzed data. Hypothesizing too few contributors to a mixture restricts the possible genotype combinations tried and thus can drastically reduce the log(LR) for true lower level contributors, however the use of a greater number typically has little impact (12). And lastly, applying the differential degradation feature to a mixture with at least one degraded contributor may render the mixture more readily solvable, but leave the log(LR) unchanged.

References

1. Perlin MW, Legler MM, Spencer CE, Smith JL, Allan WP, Bellrose, JL, et al. Validating TrueAllele® DNA mixture interpretation. J Forensic Sci 2011 Nov;56:1430-47.
2. Perlin MW, Belrose JL, Duceman BW. New York State TrueAllele® Casework validation study. J Forensic Sci 2013 Nov;58(6):1458-66
3. Richey M. The evolution of Markov chain Monte Carlo methods. American Mathematical Monthly 2010; 117:383-413.
4. Poirer DJ. The growth of Bayesian methods in statistics and economics since 1970. Bayesian Analysis 2006;1(4):969-80.
5. Gelfand AE, Smith AFM. Sampling-Based Approaches to Calculating Marginal Densities. J American Statistical Assoc 1990;85:398-409.
6. Perlin MW, Legler MM, Spencer CE, Smith JL, Allan WP, Bellrose, JL, et al. Validating TrueAllele® DNA mixture interpretation. J Forensic Sci 2011 Nov;56:1430-47.
7. Perlin MW, Sinelnikov A. An information gap in DNA evidence interpretation. PLoS ONE 2009;4:e8327.
8. TrueAllele® VUIer™ manuals (Analyze, Data, Request, Review and Report modules).
9. http://www.dfs.virginia.gov/wp-content/uploads/2014/11/210-D1500-FB-PM-TrueAllele-Casework-System.pdf (accessed Oct. 1, 2015)
10. Greenspoon SA, Schiermeier-Wood L, Jenkins BC. Establishing the limits of TrueAllele® casework: A validation study. J Forensic Sci 2015;60(5):1263-76
11. http://www.dfs.virginia.gov/wp-content/uploads/2014/11/210-D1100-FB-PM-VIII-Capillary-Electrophoresis.pdf (accessed Oct. 1, 2015)
12. Perlin MW, Hornyak JM, Sugimoto G, Miller KWP. TrueAllele_ Genotype Identification on DNA Mixtures Containing up to Five Unknown Contributors.J Forensic Sci, 2015;60(4):857-868, 2015

Acknowledgements

The authors thank Dr. Mark Perlin for helpful comments.