SWGDAM Guidelines for Validation of Probabilistic Genotyping Systems

TrueAllele® Casework System

Cybergenetics, Pittsburgh, PA

July 14, 2023

Table of Contents

Introduction 3

SWGDAM 2015 Guidelines 4

1. Validation of Probabilistic Genotyping Systems 4
2. System control 6
3. Developmental Validation 8
4. Internal Validation 12
5. Modification to Software 17

Appendix 1: TrueAllele Validation Summary

Appendix 2: TrueAllele Developmental Validations

Appendix 3: TrueAllele Peer-reviewed Papers

Appendix 4: Other Reports and Supporting Documentation

Introduction

This document describes how Cybergenetics TrueAllele® Casework system complies with all of the SWGDAM guidelines for the validation of probabilistic genotyping systems, as promulgated in their 2015 document.

The document embeds the SWGDAM guidelines, and gives a paragraph-by-paragraph description of system compliance. A separate appendix details the many TrueAllele validation studies that establish the system's reliability. There is also an appendix on the availability of the supporting documents referred to herein.

The SWGDAM guidelines are downloadable from:
http://media.wix.com/ugd/4344b0_22776006b67c4a32a5ffc04fe3b56515.pdf

Glossary

• *Cybergenetics* is a Pittsburgh-based company founded in 1994 that specializes in computer interpretation of DNA evidence data.
• *Peer review* is an assessment scientific research by a journal that has two (or more) independent workers review a manuscript before accepting it for publication.
• *Probabilistic genotyping* is any method that interprets DNA data and produces more than one genotype, assigning probabilities to the possibilities.
• *SWGDAM* is the Scientific Working Group on DNA Analysis Methods, a standing committee that helps establish guidelines of interest to the FBI.
• *TrueAllele* Casework is a computer system that accurately and automatically interprets DNA evidence data, producing reliable match statistics.
• *Validation* is a testing procedure for establishing the reliability of a method.
• *Validation study* is a scientific study that documents validation testing.

SWGDAM 2015 Guidelines

1. Validation of Probabilistic Genotyping Systems

1.1. The laboratory shall validate a probabilistic genotyping system prior to usage for forensic applications.

The TrueAllele Casework system has been extensively validated on both laboratory and casework DNA samples, with over 40 studies completed. Eight of these validation studies have been published in peer-reviewed journals. Appendix 1 (*TrueAllele Validation Summary*) describes the measures tested for each of the studies, both peer-reviewed and internal. Citations for the studies are also provided in that appendix.

1.2. The laboratory shall document all validation studies in accordance with the FBI Quality Assurance Standards for Forensic DNA Testing Laboratories.

Currently, TrueAllele validation studies have been completed on samples containing up to 10 unknown contributors with both high and low template samples tested across a range of conditions. Sensitivity, specificity, and reproducibility of the TrueAllele system have been thoroughly established, with other measures studied as well. These studies comply with the FBI QAS guidelines for validating systems, and performance checks are done when software updates are made.

1.3. The laboratory should document or have access to documentation that explains how the software performs its operations and activities, to include the methods of analysis and statistical formulae, the data to be entered in the system, the operations performed by each portion of the user interface, the workflow of the system, and the system reports or other outputs. This information enables the laboratory to identify aspects of the system should be evaluated through validation studies.

TrueAllele's methods of analysis and statistical formulae are described in various peer-reviewed publications. The *TrueAllele Methods:*

Statistical Model document summarizes those methods and citations. In addition, the *TrueAllele® VUIer™ Likelihood Ratio Calculation Application Note* describes the statistical formulae used by the TrueAllele system for calculating the LR match statistic.

TrueAllele's workflow, operation, and system inputs and outputs are described in the *TrueAllele® Visual User Interface (VUIer™)* user manuals and Cybergenetics' *TrueAllele® Casework Process: Standard Operating Procedures* document.

2. System control

2.1. The laboratory should verify that the software is installed on computers suited to run the software, that the system has been properly installed and that the configurations are correct.

TrueAllele Casework has two components: Client and Server. The TrueAllele Visual User Interface (VUIer™) client software is compatible with both Windows and Macintosh computers, with system requirements documented (see *TrueAllele® VUIer™: Getting Started* manual). A *TrueAllele® Server Quality Assurance Checklist* provides a list of tests that are performed before a server is deployed.

2.2. The laboratory should, where possible, ensure the following system control measures are in effect:

2.2.1. Every software release should have a unique version number. This version number should be referenced in any validation documentation or published results.

Each VUIer and server version has a unique version number that is documented both in the software (App > About for VUIer, Tools module for server) and in release documentation. Where applicable, version numbers appear in the validation papers or reports for a study.

2.2.2. Appropriate security protection to ensure only authorized users can access the results.

Data is stored in a secure database that is password protected at both the Operating System and Database levels. Administration features are performed by secure shell commands (SSH) secured by public key encryption (PKE).

2.2.3. Audit trails to track changes to system data and/or verification of system settings in place each time a calculation is run.

Each interpretation request performed by the system is automatically tracked via auditing information, such as date stamps and version information. Users are also provided a summary table, before

interpretation requests are uploaded, to allow for verification of the parameters.

2.2.4. User-level security to ensure that system users only perform authorized actions.

Before data can be accessed or requests uploaded, users must connect to the TrueAllele database. Access limitations exist that allow only administrators to perform certain tasks.

3. Developmental Validation

Developmental validation of a probabilistic genotyping system is the acquisition of test data to verify the functionality of the system, the accuracy of statistical calculations and other results, the appropriateness of analytical and statistical parameters, and the determination of limitations. Developmental validation may be conducted by the manufacturer/developer of the application or the testing laboratory. Developmental validation should also demonstrate any known or potential limitations of the system.

3.1. The underlying scientific principle(s) of the probabilistic genotyping methods and characteristics of the software should be published in a peer-reviewed scientific journal. The underlying scientific principles of probabilistic genotyping include, but are not limited to, modeling of stutter, allelic drop-in and drop-out, Bayesian prior assumptions such as allele probabilities, and statistical formulae used in the calculation and algorithms.

The *TrueAllele Methods: Statistical Model* document describes the underlying scientific principles, prior assumptions, and citations to publications for the development and testing of the TrueAllele System.

3.2. Developmental validation should address, where applicable, the following:

3.2.1. Sensitivity – Studies should assess the ability of the system to reliably determine the presence of a contributor's(s') DNA over a broad variety of evidentiary typing results (to include mixtures and low-level DNA quantities). This should be evaluated using various sample types (e.g., different numbers of contributors, mixture proportions, and template quantities).

Appendix 1 (*TrueAllele Validation Summary*) describes the studies that address sensitivity.

3.2.1.1. Sensitivity studies should demonstrate the potential for Type I errors (i.e., incorrect rejection of a true hypothesis), in which, for example, a contributor fails to yield a LR greater than 1 and thus his/her presence in the mixture is not supported.

Appendix 1 (*TrueAllele Validation Summary*) describes the studies that address false exclusions.

3.2.1.2. Sensitivity studies should demonstrate the range of LR values that can be expected for contributors.

Appendix 1 (*TrueAllele Validation Summary*) describes the studies that address range of LRs.

3.2.2. Specificity – Studies should evaluate the ability of the system to provide reliable results for non-contributors over a broad variety of evidentiary typing results (to include mixtures and low-level DNA quantities). This should be evaluated using various sample types (e.g., different numbers of contributors, mixture proportions, and template quantities).

Appendix 1 (*TrueAllele Validation Summary*) describes the studies that address specificity.

3.2.2.1. Specificity studies should demonstrate the potential for Type II errors (i.e., failure to reject a false hypothesis), in which, for example, a non-contributor yields a LR greater than 1 and thus his/her presence in the mixture is supported.

Appendix 1 (*TrueAllele Validation Summary*) describes the studies that address false inclusions.

3.2.2.2. Specificity studies should demonstrate the range of LR values that can be expected for non-contributors.

Appendix 1 (*TrueAllele Validation Summary*) describes the studies that address range of LR values.

3.2.3. Precision – Studies should evaluate the variation in Likelihood Ratios calculated from repeated software analyses of the same input data. This should be evaluated using various sample types (e.g., different numbers of contributors, mixture proportions, and template quantities).

Appendix 1 (*TrueAllele Validation Summary*) describes the studies that address precision.

3.2.3.1. Some probabilistic genotyping approaches may not produce the same LR from repeat analyses. Where applicable, these studies

should therefore demonstrate the range of LR values that can be expected from multiple analyses of the same data and are the basis for establishing an acceptable amount of variation in LRs.

Appendix 1 (*TrueAllele Validation Summary*) describes the studies that address within-group standard deviation and LR variation.

3.2.3.2. Any parameter settings (e.g., iterations of the MCMC) that can reduce variability should be evaluated. For example, for some complex mixtures (e.g., partial profiles with more than three contributors), increasing the number of MCMC iterations can reduce variation in the likelihood ratio.

Appendix 1 (*TrueAllele Validation Summary*) describes the studies that address MCMC sampling variation.

3.2.4. Case-type Samples – Studies should assess a range of data types exhibiting features that are representative of those typically encountered by testing laboratories. These features include those derived from mixtures and single-source samples, such as stutter, masked/shared alleles, differential and preferential amplification, degradation and inhibition.

3.2.4.1. These studies should demonstrate sample and/or data types that can be reliably evaluated using the probabilistic genotyping system.

Appendix 1 (*TrueAllele Validation Summary*) describes the studies that address case-type samples.

3.2.5. Control Samples – If the software is designed to assess controls, studies should evaluate whether correct results are obtained with control samples.

Appendix 1 (*TrueAllele Validation Summary*) describes the studies that address control samples.

3.2.6. Accuracy – Studies should assess the accuracy of the calculations performed by the system, as well as allele designation functions, where applicable.

Appendix 1 (*TrueAllele Validation Summary*) describes the studies that address accuracy.

3.2.6.1. These studies should include the comparison of the results produced by the probabilistic genotyping software to manual calculations, or results produced with an alternate software program or application, to aid in assessing accuracy of results generated by the probabilistic genotyping system. Calculations of some profiles (e.g., complex mixtures), however, may not be replicable outside of the probabilistic genotyping system.

Appendix 1 (*TrueAllele Validation Summary*) describes the studies that address comparison with manual review.

3.2.6.2. If the software uses raw data files from a genetic analyzer as input data, the peak calling, sizing and allele designation functions should be compared to the results of another software system to assess accuracy. Allele designations should also be compared to known genotypes where available.

Appendix 1 (*TrueAllele Validation Summary*) describes the studies that address allele designation comparisons.

4. Internal Validation

Internal validation of a probabilistic genotyping software system is the accumulation of test data within the laboratory to demonstrate that the established parameters, software settings, formulae, algorithms and functions perform as expected.

If conducted within the same laboratory, developmental validation studies may satisfy some of the elements of the internal validation guidelines.

4.1. The laboratory should test the system using representative data generated in-house with the amplification kit, detection instrumentation and analysis software used for casework. Additionally, some studies may be conducted by using artificially created or altered input files to further assess the capabilities and limitations of the software. Internal validation should address, where applicable to the software being evaluated:

4.1.1. Specimens with known contributors, as well as case-type specimens that may include unknown contributors.

Appendix 1 (*TrueAllele Validation Summary*) describes the sample types used in studies.

4.1.2. Hypothesis testing with contributors and non-contributors

Appendix 1 (*TrueAllele Validation Summary*) describes the studies that address the use of known contributors during TrueAllele interpretation.

4.1.2.1. The laboratory should evaluate more than one set of hypotheses for individual evidentiary profiles to aid in the development of policies regarding the formulation of hypotheses. For example, if there are two persons of interest, they may be evaluated as co-contributors and, alternatively, as each contributing with an unknown individual. The hypotheses used for evaluation of casework profiles can have a significant impact on the results obtained.

Appendix 1 (*TrueAllele Validation Summary*) describes the studies that address the use of known contributors during TrueAllele interpretation.

4.1.3. Variable DNA typing conditions (e.g., any variations in the amplification and/or electrophoresis parameters used by the laboratory to increase or decrease the detection of alleles and/or artifacts)

Appendix 1 (*TrueAllele Validation Summary*) describes the sample types used in studies.

4.1.4. Allelic peak height, to include off-scale peaks

Appendix 1 (*TrueAllele Validation Summary*) describes the sample types used in studies.

4.1.5. Single-source specimens

Appendix 1 (*TrueAllele Validation Summary*) describes the sample types used in studies.

4.1.6. Mixed specimens

Appendix 1 (*TrueAllele Validation Summary*) describes the sample types used in studies.

4.1.6.1. Various contributor ratios (e.g., 1:1 through 1:20, 2:2:1, 4:2:1, 3:1:1, etc.)

Appendix 1 (*TrueAllele Validation Summary*) describes the sample types used in studies.

4.1.6.2. Various total DNA template quantities

Appendix 1 (*TrueAllele Validation Summary*) describes the sample types used in studies.

4.1.6.3. Various numbers of contributors. The number of contributors evaluated should be based on the laboratory's intended use of the software. A range of contributor numbers should be evaluated in order to define the limitations of the software.

Appendix 1 (*TrueAllele Validation Summary*) describes the sample types used in studies.

4.1.6.4. If the number of contributors is input by the analyst, both correct and incorrect values (i.e., over- and under-estimating) should be tested.

Appendix 1 (*TrueAllele Validation Summary*) describes the sample types used in studies.

4.1.6.5. Sharing of alleles among contributors

Appendix 1 (*TrueAllele Validation Summary*) describes the sample types used in studies.

4.1.7. Partial profiles, to include the following:

4.1.7.1. Allele and locus drop-out

Appendix 1 (*TrueAllele Validation Summary*) describes the sample types used in studies.

4.1.7.2. DNA degradation

Appendix 1 (*TrueAllele Validation Summary*) describes the sample types used in studies.

4.1.7.3. Inhibition

Appendix 1 (*TrueAllele Validation Summary*) describes the sample types used in studies.

4.1.8. Allele drop-in

Appendix 1 (*TrueAllele Validation Summary*) describes the sample types used in studies.

4.1.9. Forward and reverse stutter

Appendix 1 (*TrueAllele Validation Summary*) describes the sample types used in studies.

4.1.10. Intra-locus peak height variation

Appendix 1 (*TrueAllele Validation Summary*) describes the sample types used in studies.

4.1.11. Inter-locus peak height variation

Appendix 1 (*TrueAllele Validation Summary*) describes the sample types used in studies.

4.1.12. For probabilistic genotyping systems that require in-house parameters to be established, the internal validation tests should be performed using those same parameters. The data set used to establish the parameters should be different from the data set used to validate the software using those parameters.

This guideline is not applicable when using the TrueAllele System.

4.1.13. Sensitivity, specificity and precision, as described for Developmental Validation

Appendix 1 (*TrueAllele Validation Summary*) describes the studies that address these parameters.

4.1.14. Additional challenge testing (e.g., the inclusion of non-allelic peaks such as bleed-through and spikes in the typing results

Appendix 1 (*TrueAllele Validation Summary*) describes the sample types used in studies.

4.2. Laboratories with existing interpretation procedures should compare the results of probabilistic genotyping and of manual interpretation of the same data, notwithstanding the fact that probabilistic genotyping is inherently different from and not directly comparable to binary interpretation. The weights of evidence that are generated by these two approaches are based on different assumptions, thresholds and formulae. However, such a comparison should be conducted and evaluated for general consistency.

Appendix 1 (*TrueAllele Validation Summary*) describes the studies that address comparison to manual review.

4.2.1. The laboratory should determine whether the results produced by the probabilistic genotyping software are intuitive and consistent with expectations based on non-probabilistic mixture analysis methods.

Appendix 1 (*TrueAllele Validation Summary*) describes the studies that address this guideline.

4.2.1.1. Generally, known specimens that are included based on non-probabilistic analyses would be expected to also be included based on probabilistic genotyping.

Appendix 1 (*TrueAllele Validation Summary*) describes the studies that address this guideline.

4.2.1.2. For single-source specimens with high quality results, genotypes derived from non-probabilistic analyses of profiles above the stochastic threshold should be in complete concordance with the results of probabilistic methods.

Appendix 1 (*TrueAllele Validation Summary*) describes the studies that address this guideline.

4.2.1.3. Generally, as the analyst's ability to deconvolute a complex mixture decreases, so do the weightings of individual genotypes within a set determined by the software.

Appendix 1 (*TrueAllele Validation Summary*) describes the studies that address this guideline.

5. Modification to Software

5.1. Modification to the system such as a hardware or software upgrade that does not impact interpretation or analysis of the typing results or the statistical analysis shall require a performance check prior to implementation.

Upon a VUIer or server software update, where interpretation is not affected, performance checks are conducted with in-house validation data to ensure proper function and LR calculation as performed by the software program.

5.2. A significant change(s) to the software, defined as that which may impact interpretation or the analytical process, shall require validation prior to implementation.

When server code updates affect interpretation, validation is done before the new version is distributed and used in routine processing.

5.3. Data used during the initial validation may be re-evaluated as a performance check or for subsequent validation assessment. The laboratory must determine the number and type of samples required to establish acceptable performance in consideration of the software modification.

Established data sets are used during performance checks, and the number of samples is determined before testing is done. Once sufficient testing is complete, the software or server version is deployed for use in casework.

Appendix 1: TrueAllele Validation Summary

Introduction

The TrueAllele Casework system has been thoroughly validated across a range of conditions. Cybergenetics and other groups have conducted over 40 validation studies. These studies have been presented either as peer-reviewed papers, or as written reports or presentations. Additional validation studies are currently being conducted.

This section contains a table describing the validation studies that fulfill the various developmental and internal validation guidelines presented in sections 3 and 4 of the 2015 SWGDAM Guidelines for Validation of Probabilistic Genotyping Systems. The table contains the SWGDAM *Guideline* number, a *Description* of the guideline, and a *Study* number that corresponds to the study fulfilling the guideline. These *Study* numbers correspond to both the *TrueAllele Validation Citations* section in this document as well as the study information contained in the *TrueAllele Validation Reports and Papers (ReadMe)* document. Many of these guidelines appear in other standards and guideline documents. Thus, this appendix can be used to show how TrueAllele complies with those standards and recommendations as well.

A Dropbox link to all of the papers and reports can be provided upon request. It should be noted that this table may not list every topic covered in a study but is representative of the major points covered in each study.

Note: SWGDAM guideline 4.1.12 (establishing in-house parameters) is not applicable to TrueAllele analysis.

TrueAllele Studies and SWGDAM Guidelines

Guideline	Description	Study
3.2.1, 4.1.13	Sensitivity	4, 5, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 28, 29, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43
3.2.1.1	Type I errors (False exclusions)	16, 21, 22, 23, 24, 27, 28, 32, 34, 36, 37, 39, 40, 42, 43
3.2.1.2	Sensitivity range of LR values expected for contributors	4, 5, 7, 8, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 28, 31, 32, 33, 34, 35, 36, 37, 39, 40, 43
3.2.2, 4.1.13	Specificity	7, 8, 12, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 28, 29, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 43
3.2.2.1	Type II errors (False inclusions)	16, 18, 19, 20, 21, 22, 23, 24, 27, 28, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 43
3.2.2.2	Specificity range of LR values expected for non-contributors	12, 15, 16, 18, 19, 20, 21, 22, 23, 24, 25, 27, 28, 31, 32, 33, 34, 35, 36, 37, 39, 40, 43
3.2.3, 4.1.13	Precision	2, 5, 7, 8, 9, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 28, 29, 31, 32, 33, 34, 35, 36, 37, 39, 40, 43
3.2.3.1	Range of LR values expected between multiple analyses (σ_w)	5, 7, 8, 13, 15, 16, 17, 19, 20, 21, 22, 23, 24, 25, 27, 28, 31, 32, 33, 34, 35, 36, 37, 39, 40, 43
3.2.3.2	Reducing the variability of LR variation (e.g., increasing MCMC iterations)	15, 16, 18, 28, 29, 31, 33, 34, 37, 39, 42
3.2.4, 3.2.4.1, 4.1.1	Case-type samples (reliable evaluation)	5, 6, 7, 9, 10, 13, 17, 19, 25, 27, 31, 33, 37, 38, 40, 43
3.2.5	Control samples	1, 9, 25
3.2.6	Accuracy	2, 4, 5, 6, 8, 9, 13, 15, 17, 19, 21, 24, 26, 27, 29, 31, 34, 35, 38, 39, 40, 43
3.2.6.1, 4.2	Comparison with manual review	1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 13, 15, 17, 19, 25, 29, 31, 33, 35
3.2.6.2	Comparison of allele calling of raw data (.fsa) files	1, 17
4.1	Data from kits, instruments, and analysis software used in casework	1, 3, 4, 5, 7, 8, 9, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 43
4.1.1	Known contributor samples	4, 8, 9, 12, 14, 15, 16, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43
4.1.2, 4.1.2.1	Hypothesis testing with contributors and non-contributors	4, 5, 9, 11, 12, 13, 17, 18, 19, 25, 26, 28, 29, 31, 32, 37, 38, 39, 40, 42, 43
4.1.3	Variable DNA typing conditions	9, 16, 18, 19, 22, 24, 28, 31, 32, 36, 37, 40, 43
4.1.4	Allelic peak height	3, 9, 16, 18, 19, 22, 24, 28, 30

4.1.5	Single-source samples	1, 5, 6, 8, 9, 12, 15, 25, 28, 29, 31, 35, 37, 38, 40, 43
4.1.6	Mixture samples	2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43
4.1.6.1	Various contributor ratios	4, 7, 8, 9, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 34, 35, 36, 37, 39, 40, 41, 42, 43
4.1.6.2	Various total DNA template quantities	4, 7, 8, 9, 11, 12, 15, 17, 18, 19, 20, 21, 27, 28, 32, 35, 36, 37, 40, 41, 43
4.1.6.3	Various numbers of contributors in samples	7, 10, 11, 12, 15, 16, 17, 18, 19, 21, 23, 24, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43
4.1.6.4	Over- and under- estimating of number of contributors input	8, 27, 28, 30, 32, 34, 39
4.1.6.5	Allele sharing among contributors	8, 11, 12, 18, 20, 26, 29, 38, 40
4.1.7	Partial profiles	5, 8, 9, 14, 15, 18, 28, 29, 35
4.1.7.1	Allele and locus drop-out	5, 8, 15, 18, 29, 34, 35, 39
4.1.7.2	DNA degradation	8, 12, 28, 29, 30, 32, 36, 37, 40, 43
4.1.7.3	Inhibition	30, 32, 36, 43
4.1.8	Allele drop-in	14
4.1.9	Forward and reverse stutter	1, 8, 13
4.1.10	Intra-locus peak height variation	1, 3, 29, 41
4.1.11	Inter-locus peak height variation (mixture weight modeling)	4, 5, 13, 14, 15, 17, 27, 41
4.1.14	Additional challenge testing (spikes, etc.)	1, 29
4.2.1	Determination if results produced are intuitive and consistent with expectations	1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 13, 15, 17, 18, 19, 25, 29, 31, 33, 35
4.2.1.1	If included manually, also included with probabilistic genotyping	1, 2, 4, 5, 6, 7, 8, 9, 10, 13, 15, 17, 19, 25, 29, 31, 33, 35
4.2.1.2	Single-source concordance between manual and probabilistic genotyping methods	1, 5, 6, 8, 9, 15, 17, 25, 31, 35
4.2.1.3	Weightings given to individual genotypes decrease with increasing mixture complexity	5, 8, 11, 15, 16, 17, 18, 21, 22, 23, 24, 26, 27, 28, 31, 32, 33, 34, 35, 36, 37, 39, 42, 43

TrueAllele Validation Citations

This section lists the citations for all TrueAllele validation studies.

1. Kadash K, Kozlowski BE, Biega LA, Duceman BW. Validation study of the TrueAllele® automated data review system. *J Forensic Sci.* 2004;49(4):1-8.

2. Perlin MW. Scientific validation of mixture interpretation methods. *Promega's Seventeenth International Symposium on Human Identification*, 2006 Oct 10-12; Nashville, TN.

3. Cybergenetics. "TrueAllele® System 2 and Genotyper/Genescan Peak Heights and Orchid UK Data." *Cybergenetics (Pittsburgh, PA),* May 2007.

4. Perlin MW, Sinelnikov A. An information gap in DNA evidence interpretation. *PLoS ONE.* 2009;4(12):e8327.

5. B.W. Duceman, M.W. Perlin, and J.L. Belrose. "New York State TrueAllele® Casework Developmental Validation." *New York State Police Forensic Investigation Center (Albany, NY), Cybergenetics (Pittsburgh, PA), and Northeast Regional Forensic Institute (Albany, NY),* February 2010.

6. Cybergenetics and Orchid Cellmark. "TrueAllele® Volume Crime Validation Study." *Cybergenetics (Pittsburgh, PA) and Orchid Cellmark (Abingdon, Oxfordshire, UK),* February 2010.

7. Cybergenetics. "NYSP TrueAllele® Validation." *Cybergenetics (Pittsburgh, PA),* May 2011.

8. M. Perlin, M. Legler, and J. Galdi. "Suffolk County TrueAllele® Validation." *Cybergenetics (Pittsburgh, PA) and Suffolk County Crime Laboratory (Hauppauge, NY),* May 2011.

9. NSW Review Team. "Phase 1 Evaluation Report of Cybergenetics TrueAllele® Expert System." *NSW Police Force (Lidcombe, New South Wales, Australia),* July 2011.

10. J. Sgueglia and K. Harrington. "Phase I: Internal Validation of TrueAllele Genetic Calculator as an Expert Assistant for Reads and Review of Data from Reported Sexual Assault Evidence." *Massachusetts State Police Forensic and Technology Center (Maynard, MA),* August 2011.

11. M.D. Coble and J.M. Butler. "Exploring the Capabilities of Mixture Interpretation Using True Allele Software." *National Institute for Standards and Technology (Gaithersburg, MD),* September 2011.

12. Cybergenetics. “Australia TrueAllele® Validation Report.” *Cybergenetics (Pittsburgh, PA),* September 2011.

13. Perlin MW, Legler MM, Spencer CE, Smith JL, Allan WP, Belrose JL, Duceman BW. Validating TrueAllele® DNA mixture interpretation. *J Forensic Sci.* 2011;56(6):1430-1447.

14. Ballantyne J, Hanson EK, Perlin MW. DNA mixture genotyping by probabilistic computer interpretation of binomially-sampled laser captured cell populations: Combining quantitative data for greater identification information. *Sci Justice.* 2013;53(2):103-114.

15. J. Caponera. “New York State Police Crime Laboratory System TrueAllele® Casework Validation Addendum.” *New York State Police Forensic Investigation Center (Albany, NY),* June 2013.

16. M.W. Perlin, J. Hornyak, J. Caponera, and B. Duceman. “New York State TrueAllele® Validation on DNA Mixtures of Known Composition.” *Cybergenetics (Pittsburgh, PA) and New York State Police Forensic Investigation Center (Albany, NY),* October 2013.

17. Perlin MW, Belrose JL, Duceman BW. New York State TrueAllele® Casework validation study. *J Forensic Sci.* 2013;58(6):1458-1466.

18. J. Caponera. “New York State Police Crime Laboratory System TrueAllele® Casework Validation Addendum.” *New York State Police Forensic Investigation Center (Albany, NY),* December 2013.

19. 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.

20. M.A. Clarke, J. Hornyak, W.P. Allan, and M.W. Perlin. “TrueAllele® Casework Separates DNA Mixtures that Share Alleles.” *Cybergenetics (Pittsburgh, PA),* March 2014.

21. J. Hornyak, W.P. Allan, and M.W. Perlin. “TrueAllele® Casework Validation on PowerPlex® 21 Mixture Data.” *Cybergenetics (Pittsburgh, PA),* March 2014.

22. J. Hornyak, W.P. Allan, and M.W. Perlin. “TrueAllele® Validation on Minifiler™ Mixture Data.” *Cybergenetics (Pittsburgh, PA),* July 2014.

23. J. Hornyak, M. Bowkley, and M.W. Perlin. “TrueAllele® Validation on PowerPlex® 16 HS Mixture Data.” *Cybergenetics (Pittsburgh, PA),* July 2014.

24. J. Hornyak, W.P. Allan, and M.W. Perlin. “TrueAllele® Validation on Identifiler® Plus Mixture Data.” *Cybergenetics (Pittsburgh, PA),* August 2014.

25. G. Amick. “TrueAllele Validation.” *Richland County Sheriff's Department (Columbia, SC),* March 2015.

26. K. Guest, L. Ludvico, L. Ferrara, and M. Perlin. "Development of Kinship Mixtures and Subsequent Analysis Using TrueAllele® Casework." *Master's Thesis*, *Duquesne University (Pittsburgh, PA),* April 2015.

27. Perlin MW, Hornyak J, Sugimoto G, Miller K. TrueAllele® genotype identification on DNA mixtures containing up to five unknown contributors. *J Forensic Sci*. 2015; 60(4):857-868.

28. J.M. Hornyak, T. Hebert, W.P. Allan, and M.W. Perlin. “Baltimore Police Department TrueAllele® Validation.” *Cybergenetics (Pittsburgh, PA) and Baltimore City Police Department Laboratory Section (Baltimore, MD),* August 2015.

29. Greenspoon SA, Schiermeier-Wood L, Jenkins BA. Establishing the limits of TrueAllele® Casework: a validation study. *J Forensic Sci.* 2015;60(5):1263-1276.

30. S. Greenspoon, L. Schiermeier-Wood, and B. Jenkins. “Further Exploration of TrueAllele® Casework.” *Promega's Twenty Sixth International Symposium on Human Identification,* Grapevine, TX, October 2015.

31. J. Donahue. “TrueAllele Casework Validation.” *Beaufort County Sheriff's Office (Beaufort, SC),* January 2016.

32. J.M. Hornyak, E.M. Schmidt, and M.W. Perlin. “Georgia Bureau of Investigation Forensic Biology Unit TrueAllele® Validation.” *Cybergenetics (Pittsburgh, PA) and Georgia Bureau of Investigation Forensic Biology Unit (Decatur, GA),* September 2016.

33. M.M. Legler, B.L. Harris, C.L. Booker, and M.W. Perlin. “Acadiana Criminalistics Laboratory TrueAllele® Casework Validation.” *Cybergenetics (Pittsburgh, PA) and Acadiana Criminalistics Laboratory (New Iberia, LA),* October 2016.

34. D.W. Bauer, N. Butt, and M.W. Perlin. “Cuyahoga County TrueAllele® Validation Study.” *Cybergenetics (Pittsburgh, PA) and Cuyahoga County Regional Forensic Science Laboratory (Cleveland, OH),* September 2016.

35. B.L Harris. “Acadiana Criminalistics Laboratory TrueAllele® Casework Validation Using Investigator® 24plex Kits & 2017 Server Upgrade Performance Check.” Acadiana Criminalistics Laboratory (New Iberia, LA), May 2017.

36. E.M. Schmidt. “TrueAllele® GlobalFiler Performance Check.” *Georgia Bureau of Investigation Forensic Biology Unit (Decatur, GA)*, August 2017.

37. J.M. Hornyak, C.L. Brown, and M.W. Perlin. “TrueAllele® Casework Validation of the PowerPlex® Fusion 6C STR Kit.” Cybergenetics (Pittsburgh, PA) and Louisiana State Police Crime Laboratory (Baton Rouge, LA), July 2018.

38. G. Sugimoto. “Validation of the TrueAllele® Casework VUIer™ Kinship Application.” Kern Regional Crime Laboratory (Bakersfield, CA), August 2019.

39. Bauer DW, Butt N, Hornyak JM, Perlin MW. Validating TrueAllele® interpretation of DNA mixtures containing up to ten unknown contributors. *J Forensic Sci*, 2020; 65(2):380-398.

40. B.A. Pujols, B.M. Browning, J.M. Bracamontes, M.M. Legler, D.W. Bauer, and M.W. Perlin. “TrueAllele® Casework Validation on Greenville County DNA Lab GlobalFiler™ Data.” Cybergenetics (Pittsburgh, PA) and Greenville County Department of Public Safety Forensic DNA Laboratory (Greenville, SC), March, 2020.

41. S. Antillon. “Deconvolution of DNA mixtures using replicate sampling and TrueAllele® mixture interpretation [master’s thesis].” George Mason University (Fairfax, VA), 2020.

42. H.S. Chaudhry. “Peeling away uncertainty: A probabilistic approach to DNA mixture deconvolution [master’s thesis].” George Mason University (Fairfax, VA), 2020.

43. E.E. Mole, J.M. Bracamontes, I. Fleming, M.M. Legler, and M.W. Perlin. "Metro Nashville Police Department Crime Laboratory TrueAllele® Casework Validation on PowerPlex® Fusion 6C data." Cybergenetics (Pittsburgh, PA) and Metro Nashville Police Department Crime Laboratory (Nashville, TN), June 2023.

Appendix 2: TrueAllele Developmental Validations

This section lists the citations for TrueAllele developmental validation studies.

1. Perlin MW, Sinelnikov A. An information gap in DNA evidence interpretation. *PLoS ONE.* 2009;4(12):e8327.

2. Perlin MW, Legler MM, Spencer CE, Smith JL, Allan WP, Belrose JL, Duceman BW. Validating TrueAllele® DNA mixture interpretation. *J Forensic Sci.* 2011;56(6):1430-1447.

3. Ballantyne J, Hanson EK, Perlin MW. DNA mixture genotyping by probabilistic computer interpretation of binomially-sampled laser captured cell populations: Combining quantitative data for greater identification information. *Sci Justice.* 2013;53(2):103-114.

4. Perlin MW, Belrose JL, Duceman BW. New York State TrueAllele® Casework validation study. *J Forensic Sci.* 2013;58(6):1458-1466.

5. 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.

6. Perlin MW, Hornyak J, Sugimoto G, Miller K. TrueAllele® genotype identification on DNA mixtures containing up to five unknown contributors. *J Forensic Sci*. 2015; 60(4):857-868.

7. Greenspoon SA, Schiermeier-Wood L, Jenkins BA. Establishing the limits of TrueAllele® Casework: a validation study. *J Forensic Sci.* 2015;60(5):1263-1276.

8. Bauer DW, Butt N, Hornyak JM, Perlin MW. “Validating TrueAllele® interpretation of DNA mixtures containing up to ten unknown contributors.” *J Forensic Sci*, 2020; 65(2):380-398.

Appendix 3: TrueAllele Peer-reviewed Papers

This section lists citations for TrueAllele-related peer-reviewed papers.

1. Perlin MW. Transforming conjunctive match into RETE: a call-graph caching approach, *International Journal of Software Engineering and Knowledge Engineering*, 1991;1(4):373:408.

2. Perlin MW, Burks MB, Hoop RC, Hoffman EP. Toward fully automated genotyping: allele assignment, pedigree construction, phase determination, and recombination detection in Duchenne muscular dystrophy. *Am J Hum Genet*. 1994;55(4):777-87.

3. Perlin MW, Lancia G, Ng S-K. Toward fully automated genotyping: genotyping microsatellite markers by deconvolution. *Am J Hum Genet.* 1995;57(5):1199-210.

4. Andrews C, Devlin B, Perlin M, Roeder K. Binning clones by hybridization with complex probes: statistical refinement of an inner product mapping method. *Genomics*, 1997;41(2):141-154.

5. Lancia G, Perlin M. Genotyping of pooled microsatellite markers by combinatorial optimization techniques. *Discrete Applied Math*. 1998;88(1-3):291-314.

6. Pálsson B, Pálsson F, Perlin M, Gubjartsson H, Stefánsson K, Gulcher J. Using quality measures to facilitate allele calling in high-throughput genotyping. *Genome Research*. 1999;9(10):1002-12.

7. Perlin MW, Szabady B. Linear mixture analysis: a mathematical approach to resolving mixed DNA samples. *J Forensic Sci*. 2001;46(6):1372-7.

8. Kadash K, Kozlowski BE, Biega LA, Duceman BW. Validation study of the TrueAllele® automated data review system. *J Forensic Sci.* 2004;49(4):1-8.

9. Hill SY, Shen S, Zezza N, Hoffman EK, Perlin M, Allan W. A genome wide search for alcoholism susceptibility genes. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*. 2004;128B(1):102-13.

10. Perlin MW, Kadane JB, Cotton RW. Match likelihood ratio for uncertain genotypes. *Law, Probability and Risk*. 2009;8(3):289-302.

11. Perlin MW, Sinelnikov A. An information gap in DNA evidence interpretation. *PLoS ONE.* 2009;4(12):e8327.

12. Perlin MW, Legler MM, Spencer CE, Smith JL, Allan WP, Belrose JL, Duceman BW. Validating TrueAllele® DNA mixture interpretation. *J Forensic Sci.* 2011;56(6):1430-1447.

13. Ballantyne J, Hanson EK, Perlin MW. DNA mixture genotyping by probabilistic computer interpretation of binomially-sampled laser captured cell populations: Combining quantitative data for greater identification information. *Sci Justice.* 2013;53(2):103-114.

14. Perlin MW, Belrose JL, Duceman BW. New York State TrueAllele® Casework validation study. *J Forensic Sci.* 2013;58(6):1458-1466.

15. 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.

16. Perlin MW. Inclusion probability for DNA mixtures is a subjective one-sided match statistic unrelated to identification information. *Journal of Pathology Informatics*, 6(1):59, 2015.

17. Perlin MW, Hornyak J, Sugimoto G, Miller K. TrueAllele® genotype identification on DNA mixtures containing up to five unknown contributors. *J Forensic Sci*. 2015; 60(4):857-868.

18. Greenspoon SA, Schiermeier-Wood L, Jenkins BA. Establishing the limits of TrueAllele® Casework: a validation study. *J Forensic Sci.* 2015;60(5):1263-1276.

19. Stokes NA, Stanciua CE, Brocatoa ER, Ehrhardta CR, Greenspoon SA. Simplification of complex DNA profiles using front end cell separation and probabilistic modeling. *Forensic Science International: Genetics*. 2018;36:205-212.

20. Perlin MW. Efficient construction of match strength distributions for uncertain multi-locus genotypes. *Heliyon*, 4(10):e00824, 2018.

21. Bauer DW, Butt N, Hornyak JM, Perlin MW. “Validating TrueAllele® interpretation of DNA mixtures containing up to ten unknown contributors.” *J Forensic Sci*, 2020; 65(2):380-398.

Appendix 4: Other Reports and Supporting Documentation

Several supporting reports and other materials are mentioned throughout this document. These materials give additional support for TrueAllele's compliance with various guidelines and standards. A Dropbox link to these documents can be provided upon request.

TrueAllele reports

Perlin MW. Scientific validation of mixture interpretation methods. Promega's Seventeenth International Symposium on Human Identification, 2006; Nashville, TN.

Perlin MW. Explaining the likelihood ratio in DNA mixture interpretation. Promega's Twenty First International Symposium on Human Identification, 2010; San Antonio, TX.

Other supporting documents:

- *TrueAllele® Methods: Statistical Model*
- *TrueAllele® VUIer™* user manuals:
    - *Workflow Introduction*
    - *Getting Started*
    - *Analyze Module*
    - *Data Module*
    - *Request Module*
    - *Review Module*
    - *Report Module*
    - *Tools Module*
    - *Tutorial*
    - *Database Application Note*
    - *Specificity Application Note*
    - *Likelihood Ratio Calculation Application Note*
- Cybergenetics' *TrueAllele® Casework Process: Standard Operating Procedures*
- *TrueAllele® Server Quality Assurance Checklist*