NYSP TrueAllele® Validation

Prepared for

New York State Commission on Forensic Science
DNA Subcommittee
Meeting on 20 May 2011

This report was prepared with funding provided by New York State Contract C001220 between the New York State Police and Cybergenetics, Corp. (Pittsburgh, PA). Data was provided by the New York State Police Forensic Investigation Center.

Table of Contents

Property Crime Study: Match information

We conducted a TrueAllele® validation study of 25 NYSP/FIC property crime cases. The purpose was to determine the efficacy and reproducibility of the inferred genotype matches, and the suitability of TrueAllele for routine casework interpretation. For this purpose, it sufficed to make comparisons only to homeowner or other elimination genotypes, and not use suspect genotypes.

The NYSP/FIC laboratory generated data using the Identifiler STR panel, size separating the fluorescently labeled amplicons on an ABI 3130 DNA sequencer. TrueAllele was used to analyze the electronic .fsa data files, and uploaded quality checked quantified peak size data into the FIC TrueAllele database. Following request generation in the VUIer™ interface, TrueAllele interpreted the evidence. Each item was interpreted twice, assuming both one and two unknown contributors, since property crime samples are often mixtures and the computer can analyze the same data in different ways. The TrueAllele database was copied and reset, with reinterpretation of all requests producing duplicate genotype results.

TrueAllele objectively inferred genotypes from the electronic evidence data, without access to any reference genotypes. TrueAllele then compared these inferred evidence genotypes with the elimination references to compute a likelihood ratio (LR) match statistic, relative to three standard ethnic populations (Black, Caucasian, Hispanic). The smallest LR was reported, following NYSP practice. A co-ancestry coefficient of 1% was used throughout. The $\log_{10}$(LR) is a standard additive measure of information used throughout this report.

There were 91 evidence items analyzed, of which 33 yielded a positive (i.e., greater than zero) log(LR) TrueAllele match result. The FIC reported 6 Conditional Match Probability LR scores for single source items, since with property crime the lab usually assigns scores to just the probative items, and these cases did not have suspect references.

The mean TrueAllele log(LR) score on the 33 positive items was 10.72 (52.7 billion), and the median was 11.28 (190 billion). The information scores ranged from a minimum of 0.27 to a maximum of 21.84, with a standard deviation of 7.14.

Since the TrueAllele interpretations were run in duplicate, we could assess the reproducibility of genotype information in this data set via the log(LR) match scores. The within-item standard deviation was 0.364 log(LR) units. A bar chart (Figure 1) shows the efficacy and reproducibility of TrueAllele on these data. Note that more informative DNA items tend to have greater interpretation reproducibility.

The collated match statistic data can be found in the spreadsheet *match.xls*. There is a separate sheet for the inferred genotypes (Genotype), each evidence item (Item), computer-only results (TrueAllele), and human-only results (NYS).

Property Crime Study: Sensitivity and specificity

For each of the 25 cases, we looked at the sensitivity (within case) and specificity (between case) of the match information. This was done for comparisons of evidence genotypes to both references and evidence genotypes. The four-way design was:

log(LR) distribution	**reference comparison**	**evidence comparison**
within case	item-to-reference sensitivity	item-to-item sensitivity
between case	item-to-reference specificity	item-to-item specificity

We give examples for these four situations, based on the results for the first case, 09HL-1005. This case has 3 evidence items and 4 elimination references. Each item was interpreted in two different ways, assuming either 1 or 2 unknown contributors, and run in duplicate. Additional joint interpretations were done combining some items. Therefore, we have much log(LR) match data between various inferred genotypes and the elimination references.

Within case, evidence to reference

N = 88
file = Match_Tables/09HL-1005/Within Case/Display_All/09HL-1005_eviref_0.xls

The histogram (Figure 2) shows some very specific matches (log(LR) > 6) where the LR over a million clearly identifies an individual, demonstrating LR sensitivity within the case. There are nonspecific matches (0 < log(LR) < 3) due to cross-matching of 50:50 inferred genotypes, as well as the relatedness of the family members in the household. The preponderance of nonmatches (log(LR) < 0) shows LR specificity when the data do not support a match.

Within case, evidence to evidence

N = 169
file = Match_Tables/09HL-1005/Within Case/Display_All/09HL-1005_evievi_0.xls

The histogram (Figure 3) shows a positive association between the inferred evidence genotypes within a case. Starting on the right, we see the most specific (log(LR) > 9) evidence to evidence matches. There is then a right unimodal distribution, centered around log(LR) = 6, due to close matching of evidence item genotypes. The left unimodal distribution, centered around log(LR) = -3, comes from somewhat related evidence items. The overall effect is a rightward shift indicating greater association (than a random alternative) of the evidence genotypes in the case.

Between case, evidence to reference

N = 559
file = Match_Tables/09HL-1005/Across Cases/Display_All 09HL-1005_evivsallref_0.xls

The histogram (Figure 4) shows the specificity of the inferred evidence genotypes, relative to elimination samples. The TrueAllele inferred genotypes in this case were compared with the reference samples in all the other 24 cases in the study. The unimodal log(LR) distribution has mean -16.95 (under a quadrillionth) and median -17.15. The range is from -30 to a maximum of -4.24. These negative log(LR) values demonstrate that the evidence genotypes are specific to their own case, and do not identify individuals in other cases.

Between case, evidence to evidence

N = 3978
file = Match_Tables/09HL-1005/Across Cases/Display_All/09HL-1005_evievi_0.xls

The histogram (Figure 5) shows the specificity of the inferred evidence genotypes, relative to other evidence genotypes. The 13 TrueAllele inferred genotypes in this case were compared with the 306 evidence genotypes across the other 24 study cases. All but 15 of the 3,978 evidence-to-evidence genotype comparisons had a negative log(LR). Eleven of them (0.28%) had score between 0 and 1, while 4 of them (0.01%) had a log(LR) between 1 and 2, with no higher match scores observed. These log(LR) values demonstrate that the computer inferred evidence genotypes are specific to their own case, and do not associate with evidence genotypes in other cases.

Other case results

The same sensitivity and specificity analysis was conducted for the other 24 cases. The spreadsheets for these cases can be found in the *Match_Tables* folder. That folder also contains a *description.pdf* file that describes the folder organization for the 200 spreadsheets (8 files for each of the 25 cases).

Mixture Study Update

Preserving identification information

At the May, 2010 DNA subcommittee meeting, we described the match results (both computer and human) for all 86 mixture items in the validation study presented in March, 2010. Figure 6 lists these items on the x-axis, ordered by decreasing match information. The y-axis again shows DNA match information, measured by log likelihood ratio.

The amount of information that the TrueAllele computer inferred for each item is shown as the large (blue) background. Quantitative computer interpretation successfully determined a LR score for every mixture item in the study, with a median log(LR) score of 15 (quadrillion).

The analyst's LR result in a case is shown as a bar in the foreground. The bar's height gives the log(LR) information, while the method is indicated by the color: RMP (gray), CLR (green) and CPI (orange). When no match score was assigned, no bar appears.

The qualitative threshold-based review assigned a match score to only 30% of the items. Moreover, even when human review did find an answer, the qualitative method lost about half of the data's match information.

Solving unsolved cases

A question arises: examining the 41 cases, does this 70% loss in reportable mixture item information ever lead to a solvable case not being solved? For example, perhaps the case already had an informative single source or other mixture item, so that not all mixture items had to be interpreted with a match statistic.

To address this question (Figure 7), we organized the LR item results by case. Shown in a stacked bar chart are the number of numerical match results for computer interpretation of mixture (blue) and single source (green) samples, along with the corresponding human mixture (red) and single source (yellow) interpretation count.

We can identify 5 cases where the computer obtained a reportable mixture LR result, but qualitative human review obtained no match score on any item, whether mixture or single source. This is indicated (numbers 1, 15, 18, 32 and 36) by the presence of a blue bar, but no red or yellow bars. So the computer can indeed solve cases that human review cannot. In this study, the rate of human unsolved (but computer solvable) was 12% (5/41).

Workflow Study: Complex case resolution

Current human data review of DNA evidence in a case having many items can be an arduous undertaking.

- Each of the J evidence items STR data must be individually examined to infer a genotype (allele pair listing or probabilities). Mixed or damaged items entail more consideration.
- Similarly, the K known items are separately reviewed to produce reference genotypes.
- After genotyping all the items, JxK comparisons are made between every pair of evidence and reference genotypes to assess their degree of match.
- Likelihood ratio (LR) match scores are calculated for every matching genotype pair, under various reference population and coancestry assumptions.
- These LR scores are then entered into the case folder.

A computer-based genetic calculator that automates genotype inference and LR match comparison can appreciably streamline the DNA data review process.

- A user asks the computer to examine each of the J evidence and K reference item data. The user indicates that these items belong to the same case, so that they will be matched later on.
- The computer independently assesses each DNA item, inferring a genotype for each item. With mixtures, the computer infers a genotype for each unknown contributor. This objective genotype inference never uses any suspect information.
- The computer automatically compares the inferred evidence and known genotypes in the case to form LR match scores that can be saved into a report table.

With many DNA items, this computer-driven workflow reduces the human endeavor from J+K genotype manual genotype inferences and JxK match comparisons to simply asking J+K calculator questions.

We compared the human and computer workflows on a case having J=30 evidence items and K=6 knowns. The manual review entailed months of human effort to genotype the 36 (30+6) items, and then compare the 180 (30x6) genotype pairs. Indeed, data examination revealed the presence of two additional "Jane/John Doe" individuals, increasing the reference number up to K=8, so that ultimately 240 (30x8) genotype pair match comparisons were made. LR scores (such as CMP, CPI and CLR) were computed for every match.

TrueAllele provided a simpler workflow that was faster, more informative and easy to use. Cybergenetics began by taking two hours of their time to ask the computer to genotype the J=30 evidence items (including mixtures) and K=6 reference items. The parallel computer proceeded to infer genotypes and match them (Table 1). Afterwards, they reviewed the LR match table results, augmenting the table with the two inferred "Jane/John Doe" genotypes that increased the number of references to K=8 (Table 2).

We also asked the computer some additional focused questions. For example, on two evidence items that had yielded no individual match results, we conducted a joint computer analysis that assumed four contributors (Table 3). The computer solved for three unknown contributor genotypes, producing a match from evidence to suspect with a LR of over a million to one. Amongst the 240 (30x8) genotype comparisons, the computer found 27 matches with reportable LR scores (Table 4).

Case Report Formats

Genotype

For both nonsuspect and suspect cases, the genotype probability represents the computer's objective results prior to making any genotype comparisons. This is similar to the current laboratory practice of listing observed allele peaks, independently of any match comparison. Here is a typical genotype report, showing for each locus the allele pairs and their probabilities. This format is consistent with the new ANSI/NIST-ITL forensic exchange standard for DNA laboratories when reporting probabilistic genotypes.

locus	allele 1	allele 2	probability
AMELO	1	1	1
CSF1PO	10	11	0.95
D13S317	9	13	0.99
D16S539	12	12	1
D18S51	12	18	0.96
D19S433	11	13	0.99
D21S11	30	31.2	0.95
	30	30	0.05
D2S1338	21	24	0.9
	17	24	0.07
D3S1358	17	18	0.98
D5S818	12	15	1
D7S820	9	11	0.92
	8	9	0.05
D8S1179	11	13	0.99
FGA	20.2	23	0.89
	20	23	0.04
	23	23	0.01
	21	23	0.01
TH01	7	9	0.99
TPOX	8	11	0.99
vWA	14	19	1

Match

When a comparison is made between an evidence genotype and another genotype (reference or evidence), a LR value can be determined. Once there is an LR value between two genotypes (relative to a reference population), a match report can be generated. A table of loci (rows) versus populations (columns) can give a log(LR) value in each entry, as shown. The joint (total) match information can be given for each population. This table is consistent with current CODIS Popstats reports, which produce similarly formatted LR information.

log(LR)	US_BLK_FBI	US_CAU_FBI	US_HIS_FBI
CSF1PO	0.92	0.78	0.84
D13S317	2	1.7	1.19
D16S539	1.37	0.9	1.04
D18S51	1.72	1.54	1.82
D19S433	1.34	2.25	2.35
D21S11	1.49	1.28	1.18
D2S1338	1.49	2.06	2.41
D3S1358	1.57	1.13	1.6
D5S818	2.15	2.06	2.14
D7S820	1.1	1.17	1.46
D8S1179	1.68	1.32	1.33
FGA	2.44	2.27	2.31
TH01	0.87	1.21	1.12
TPOX	0.77	0.55	0.51
vWA	1.92	1.7	1.95
Total	**22.84**	**21.92**	**23.27**

Following current lab practice, the smallest log(LR) value would be reported, here equal to 21.92. Other information can be supplied in case report, in addition to this locus match table. For example, the joint log(LR) value of 21.92 can be translated into words via exponentiation (i.e., $10^{21.92}$) as 8.32 sextillion.

Accompanying Materials

The accompanying materials contain useful contextual information, most of which has already been presented to the DNA subcommittee.

DNA Subcommittee Presentations

These presentations to the DNA Subcommittee from NYSP about TrueAllele done in the year 2010 are found in the *Presentations* folder.

- March: validation report submitted to the committee
- March: Dr. Mark Perlin's PowerPoint presentation
- March: Jamie Belrose's PowerPoint presentation
- May: Dr. Barry Duceman's PowerPoint presentation

TrueAllele Publications

These recent TrueAllele publications are found in the *Publications* folder.

- December, 2009: PLoS ONE TrueAllele methods and validation paper
- December, 2010: Promega conference paper on explaining the likelihood ratio
- November, 2011: JFS TrueAllele methods and validation paper

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

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

Perlin MW, Legler MM, Spencer CE, Smith JL, Allan WP, Belrose JL, et al. Validating TrueAllele® DNA mixture interpretation. Journal of Forensic Sciences. 2011;56(November):in press.

Figures

Figure 1. A bar chart of item information, sorted by LR, that shows efficacy and reproducibility on a logarithmic scale.

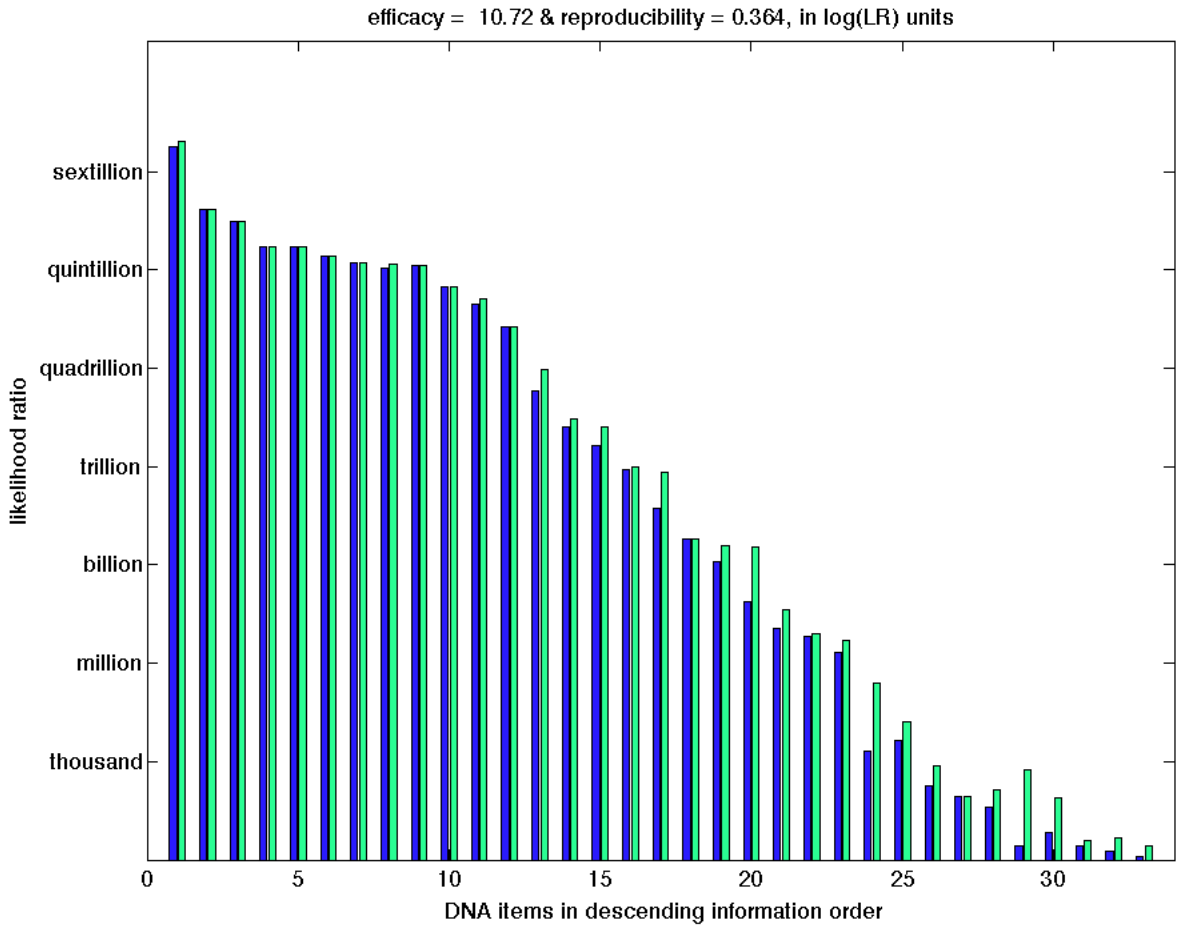


Figure 2. A histogram of log(LR) information for within-case evidence to reference genotype comparisons.

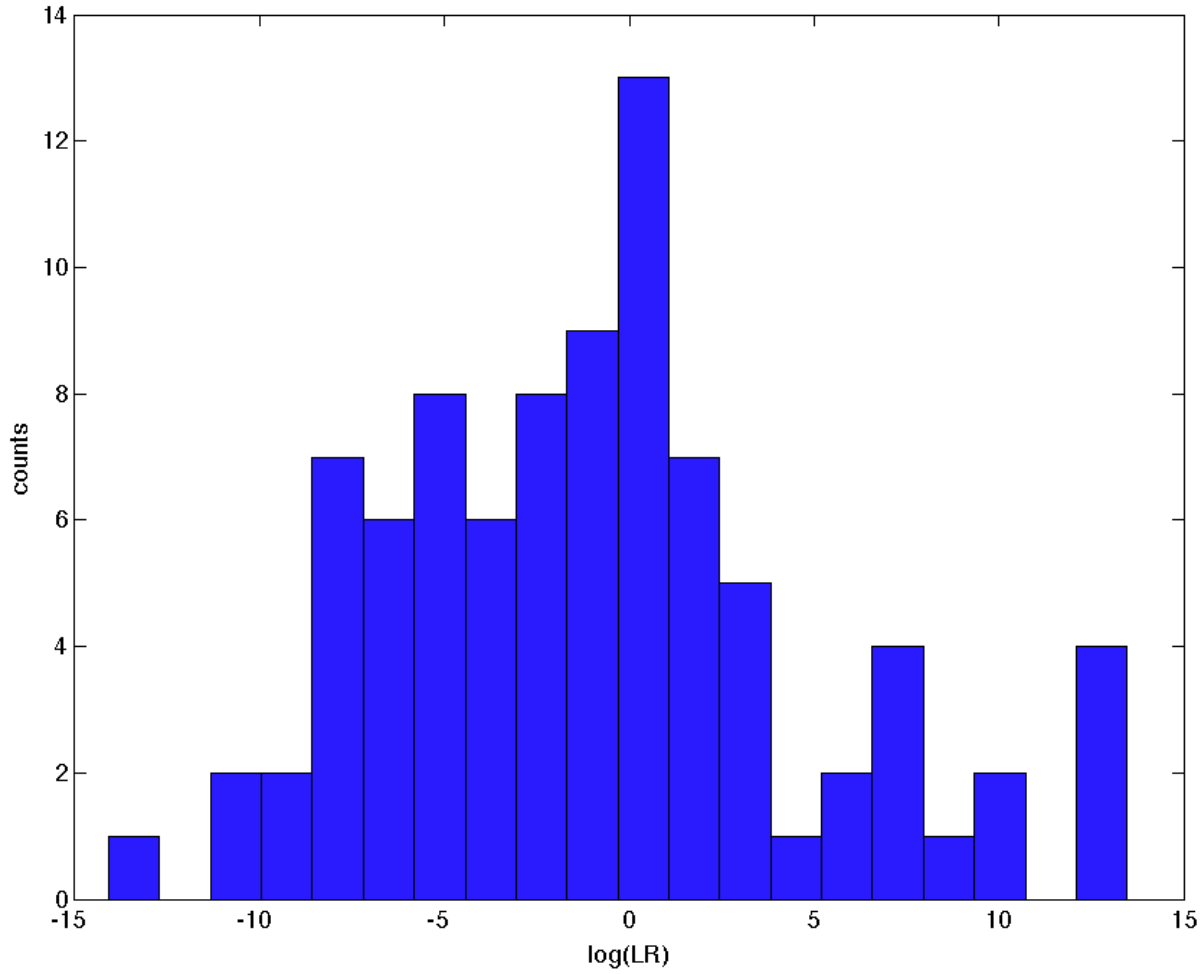


Figure 3. A histogram of log(LR) information for within-case evidence to evidence genotype comparisons.

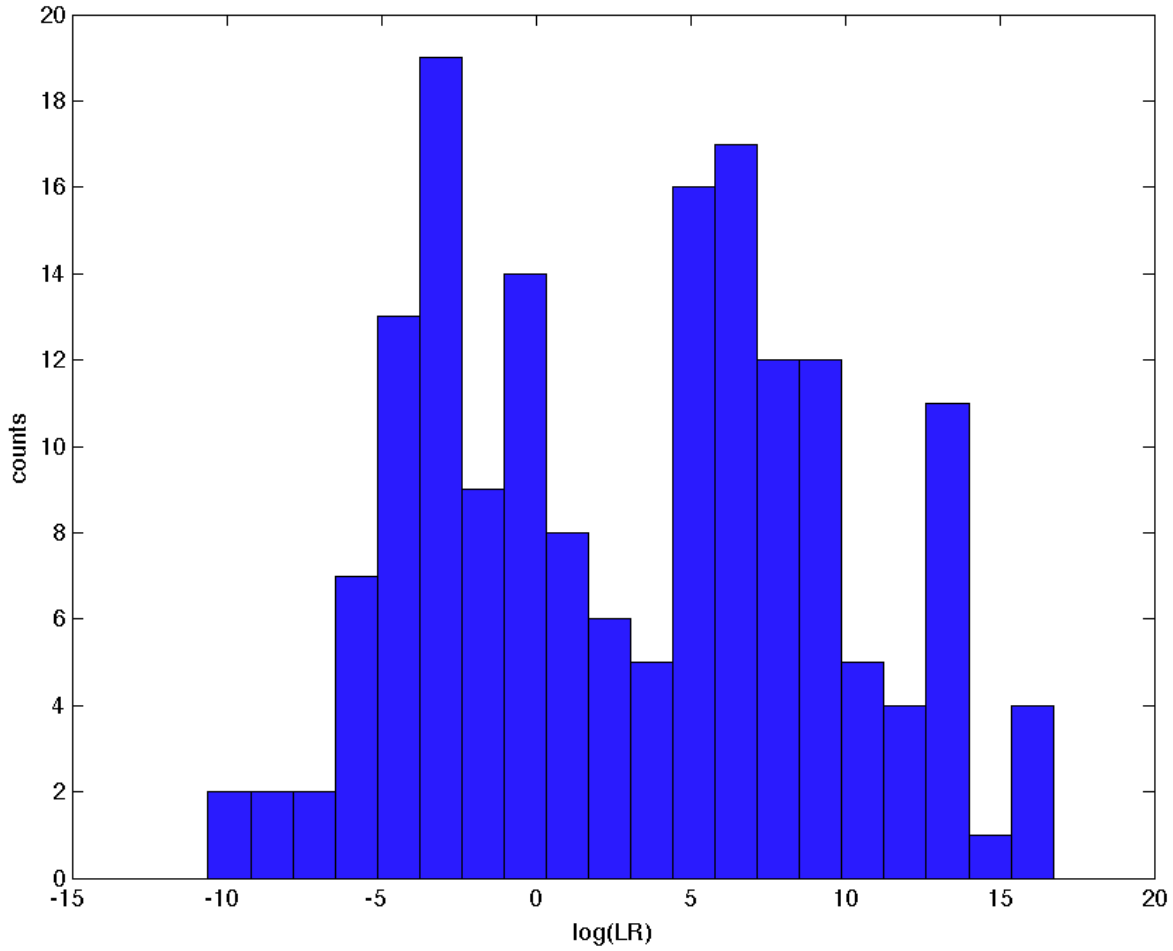


Figure 4. A histogram of log(LR) information for between-case evidence to reference genotype comparisons.

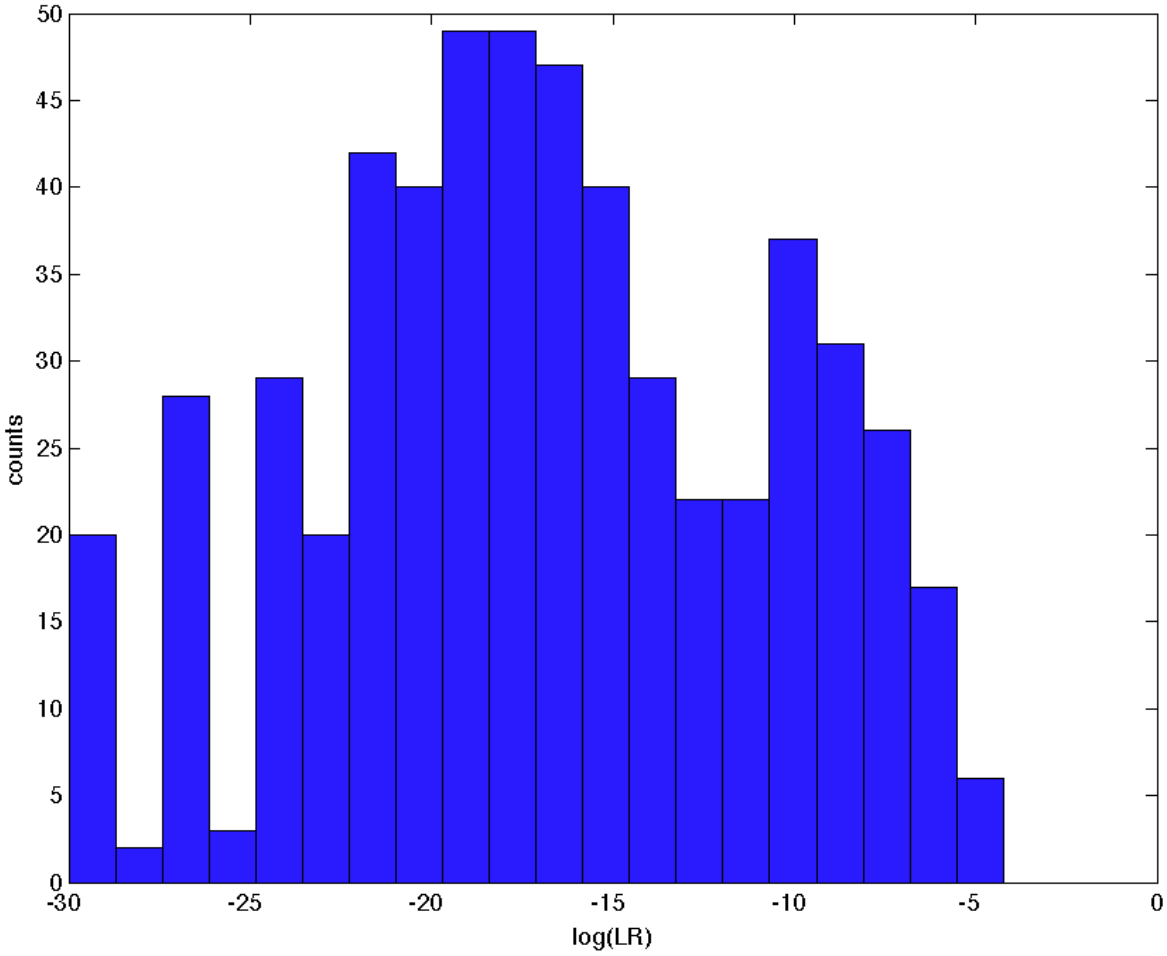


Figure 5. A histogram of log(LR) information for between-case evidence to evidence genotype comparisons.

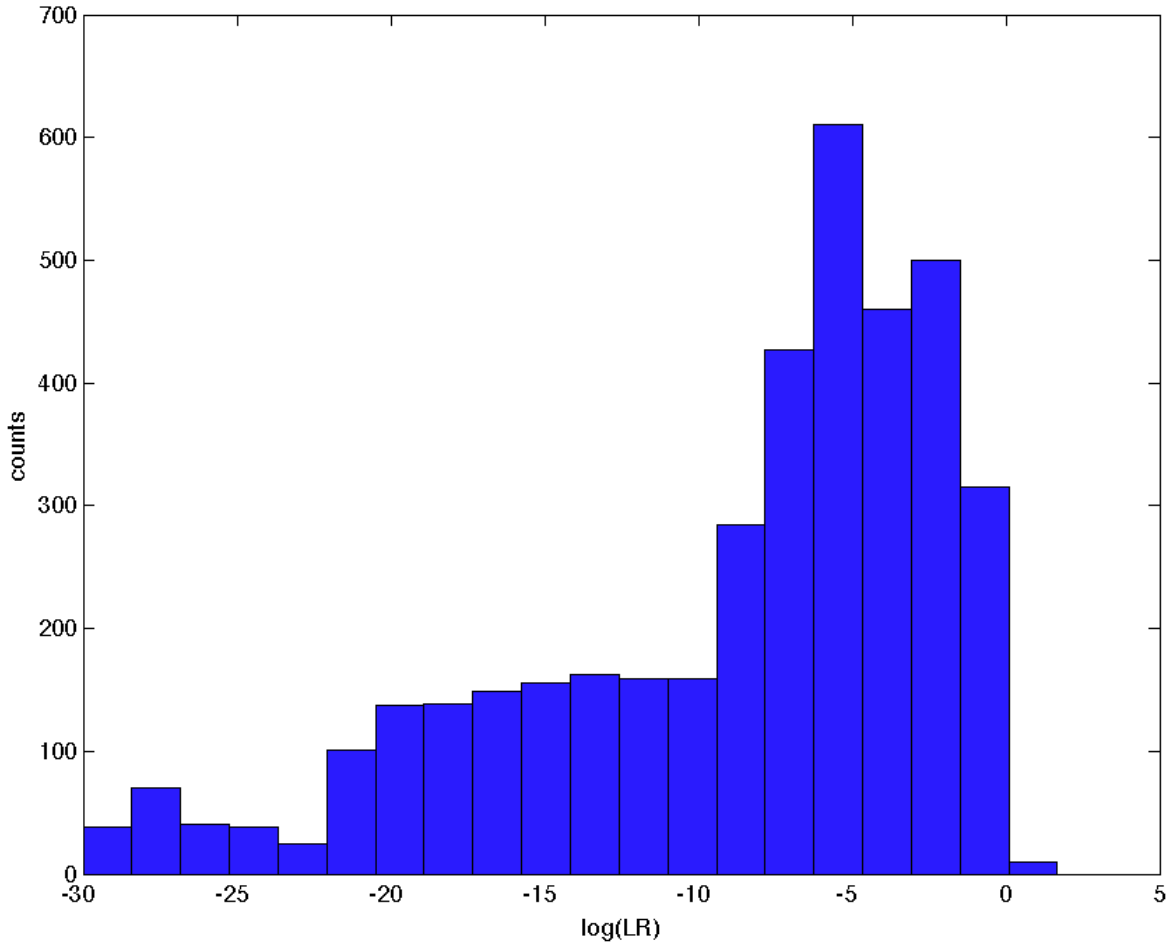


Figure 6. In the previously reported 2010 mixture study, we compared the information efficacy of quantitative (blue) and qualitative (gray, green, orange) interpretation of 86 case mixture items, conducted on the same data. The y-axis log(LR) shows identification information.

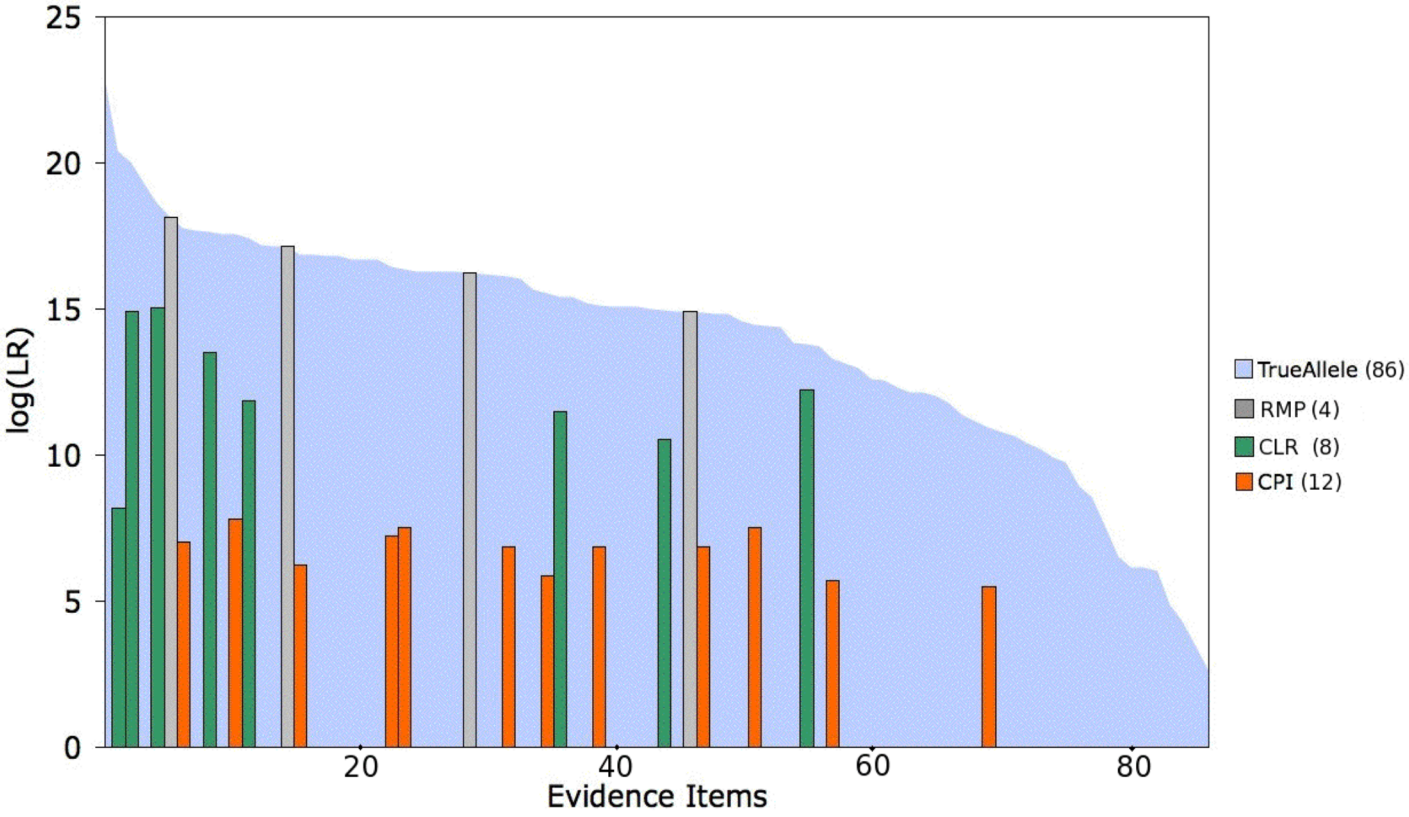
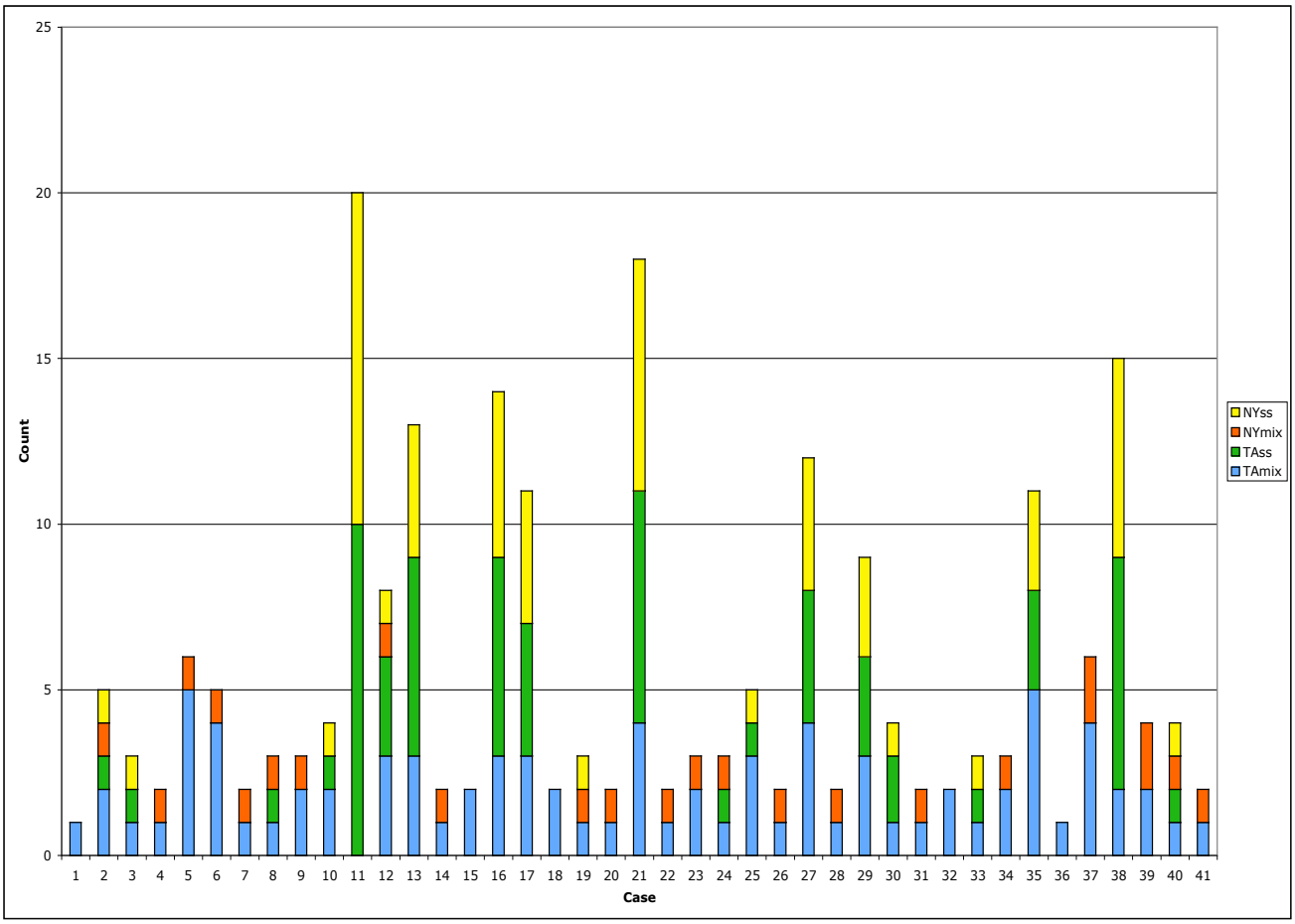


Figure 7. In the previously reported 2010 86 item mixture study, there were 41 cases (x-axis). For each case, the number of suspect matches found using quantitative TrueAllele interpretation from a mixture (blue) or single source (green) item is shown, along with the number found using qualitative manual interpretation from a mixture (red) or single source (yellow) item.



Tables

Table 1. Complex case resolution, initial TrueAllele computation results.

item	1F1	1G1	A1H1	A1I1	A1J1	A1K1
1A2	14.65					
1E1	14.65					
1E3	14.65					
1E5	14.64					
1E6				9.11		
1E7				4.11		
1H-I	6.55					
1J-L						
2A1A	14.65					
2E1A	14.65					
2E1C	14.65					
2E1F	14.65					
2E1H						
2E2A	14.65					
2E2B	14.65					
2E2C			12.40			
2E2D			6.98			
2E2E						
2F-H						
2I-J						
2K	14.65					
2L-N						
2O						
3D-E						
3F-H						
3I-K						
3L-N						
4D						
4E						
4F						

Table 2. Complex case resolution, adding Jane and John Doe references.

item	1F1	1G1	A1H1	A1I1	A1J1	A1K1	jane1	john1
1A2	14.65							
1E1	14.65							
1E3	14.65							
1E5	14.64						14.40	
1E6				9.11				
1E7				4.11				
1H-I	6.55						10.53	
1J-L								
2A1A	14.65							
2E1A	14.65							
2E1C	14.65							
2E1F	14.65							
2E1H								
2E2A	14.65							
2E2B	14.65							
2E2C			12.40					
2E2D			6.98					
2E2E								
2F-H								
2I-J								
2K	14.65							
2L-N								
2O								
3D-E								
3F-H								
3I-K							14.50	
3L-N							10.02	
4D								13.60
4E								13.23
4F								13.16

Table 3. Complex case resolution, joint interpretation results.

item	**1F1**	1G1	**A1H1**	A1I1	A1J1	A1K1	jane1	john1
1A2	14.65							
1E1	14.65							
1E3	14.65							
1E5	14.64						14.40	
1E6				9.11				
1E7				4.11				
1H-I	6.55						10.53	
1J-L								
2A1A	14.65							
2E1A	14.65							
2E1C	14.65							
2E1F	14.65							
2E1H			**6.44**					
2E2A	14.65							
2E2B	14.65							
2E2C			12.40					
2E2D			6.98					
2E2E			**6.44**					
2F-H								
2I-J								
2K	14.65							
2L-N								
2O								
3D-E								
3F-H								
3I-K							14.50	
3L-N							10.02	
4D								13.60
4E								13.23
4F								13.16

Table 4. Complex case resolution, joint interpretations found additional matches.

item	**1F1**	1G1	A1H1	A1I1	A1J1	A1K1	**jane1**	john1
1A2	14.65							
1E1	14.65							
1E3	14.65							
1E5	14.64						14.40	
1E6				9.11				
1E7				4.11				
1H-I	6.55						10.53	
1J-L								
2A1A	14.65							
2E1A	14.65							
2E1C	14.65							
2E1F	14.65							
2E1H			6.44					
2E2A	14.65							
2E2B	14.65							
2E2C			12.40					
2E2D			6.98					
2E2E			6.44					
2F-H								
2I-J								
2K	14.65							
2L-N								
2O								
3D-E								
3F-H								
3I-K	**5.34**						**13.77**	
3L-N	**5.34**						**13.77**	
4D								13.60
4E								13.23
4F								13.16