Creating informative DNA libraries using computer reinterpretation of existing data

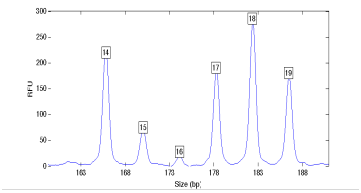
Northeastern Association of Forensic Scientists
November, 2011
Newport, RI

Ria David, PhD
Mark W Perlin, PhD, MD, PhD
Cybergenetics, Pittsburgh, PA

Cybergenetics

Crime scene biological evidence

Challenging DNA data



Human review uninformative

SWGDAM 2010 triage

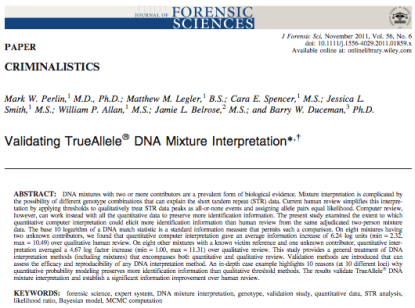
Interpret DNA evidence

resolvable → human
stochastic
threshold

inconclusive → computer
probabilistic
genotype
(3.2.2)

MW Perlin, MM Legler, CE Spencer, JL Smith, WP Allan, JL Belrose, BW Duceman.
Journal of Forensic Sciences, November 2011, Volume 56, Issue 6, Pages 1430-1447.

Validated probabilistic genotyping



J Forensic Sci, November 2011, Vol. 56, No. 6
doi: 10.1111/j.1556-4029.2011.01859.x
Available online at: onlinelibrary.wiley.com

PAPER

CRIMINALISTICS

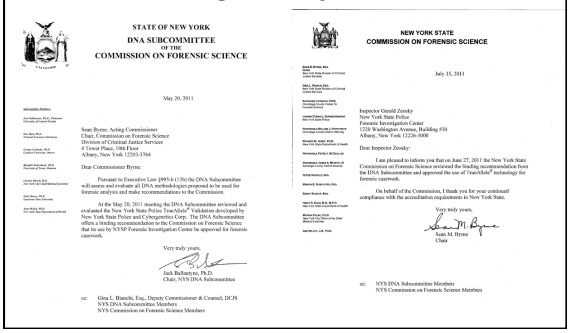
Mark W. Perlin,[1] M.D., Ph.D.; Matthew M. Legler,[1] B.S.; Cara E. Spencer,[1] M.S.; Jessica L. Smith,[1] M.S.; William P. Allan,[1] M.S.; Jamie L. Belrose,[2] M.S.; and Barry W. Duceman,[3] Ph.D.

Validating TrueAllele® DNA Mixture Interpretation*,†

ABSTRACT: DNA mixtures with two or more contributors are a prevalent form of biological evidence. Mixture interpretation is complicated by the possibility of different genotype combinations that can explain the short tandem repeat (STR) data. Current human review simplifies this interpretation by applying thresholds to qualitatively treat STR data peaks as all-or-none events and assigning allele pairs equal likelihood. Computer review, however, can work instead with all the quantitative data to preserve more identification information. The present study examined the extent to which quantitative computer interpretation could elicit more identification information than human review from the same adjudicated two-person mixture data. The base 10 logarithm of a DNA match statistic is a standard information measure that permits such a comparison. On eight mixtures having two unknown contributors, we found that quantitative computer interpretation gave an average information increase of 6.24 log units (min = 2.32, max = 10.49) over qualitative human review. On eight other mixtures with a known victim reference and one unknown contributor, quantitative interpretation averaged a 4.87 log factor increase (min = 1.00, max = 11.31) over qualitative review. This study provides a general treatment of DNA interpretation methods (including mixtures) that encompasses both quantitative and qualitative review. Validation methods are introduced that can assess the efficacy and reproducibility of any DNA interpretation method. An in-depth case example highlights 10 reasons (at 10 different loci) why quantitative probability modeling preserves more identification information than qualitative threshold methods. The results validate TrueAllele® DNA mixture interpretation and establish a significant information improvement over human review.

KEYWORDS: forensic science, expert system, DNA mixture interpretation, genotype, validation study, quantitative data, STR analysis, likelihood ratio, Bayesian model, MCMC computation

NYS Regulatory Approval



STATE OF NEW YORK
DNA SUBCOMMITTEE
OF THE
COMMISSION ON FORENSIC SCIENCE

May 20, 2011

Sean Byrne, Acting Commissioner
Chair, Commission on Forensic Science
Division of Criminal Justice Services
4 Tower Place, 10th Floor
Albany, New York 12203-3764

Dear Commissioner Byrne:

Pursuant to Executive Law §995-b (13b) the DNA Subcommittee will assess and evaluate all DNA methodologies proposed to be used for forensic analysis and make recommendations to the Commission.

At the May 20, 2011 meeting the DNA Subcommittee reviewed and evaluated the New York State Police TrueAllele® Validation developed by New York State Police and Cybergenetics Corp. The DNA Subcommittee offers a binding recommendation to the Commission on Forensic Science that its use by NYSP Forensic Investigation Center be approved for forensic casework.

Very truly yours,

Jack Ballantyne, Ph.D.
Chair, NYS DNA Subcommittee

cc: Gina L. Bianchi, Esq., Deputy Commissioner & Counsel, DCJS
NYS DNA Subcommittee Members
NYS Commission on Forensic Science Members

NEW YORK STATE
COMMISSION ON FORENSIC SCIENCE

July 13, 2011

Inspector Gerald Zoosky
New York State Police
Forensic Investigation Center
1220 Washington Avenue, Building #30
Albany, New York 12226-3000

Dear Inspector Zoosky:

I am pleased to inform you that on June 27, 2011 the New York State Commission on Forensic Science reviewed the binding recommendation from the DNA Subcommittee and approved the use of TrueAllele® technology for forensic casework.

On behalf of the Commission, I thank you for your continued compliance with the accreditation requirements in New York State.

Very truly yours,

Sean M. Byrne
Chair

cc: NYS DNA Subcommittee Members
NYS Commission on Forensic Science Members

Key TrueAllele® features

- objective interpretation never sees suspect

I. E. Dror, G. Hampikian. "Subjectivity and bias in forensic DNA mixture interpretation." *Science & Justice*. 2011

- thorough consideration of all genotype possibilities

J.-A. Bright, P. Gill, J. Buckleton. "Composite profiles in DNA analysis." *Forensic Science International: Genetics*. 2011

- challenging DNA data: mixed, low-level, degraded

H. Kelly, J. Bright, J. Curran, J. Buckleton. "The interpretation of low level DNA mixtures." *Forensic Science International: Genetics*. 2011

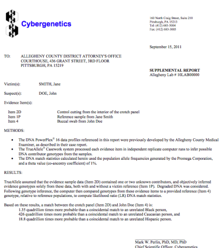
TrueAllele investigation

- lab and prosecutor select case
- lab sends Cybergenetics data
- TrueAllele solves the problem with parallel statistical search
- Cybergenetics reports initial match results to prosecutor

Case library of informative DNA match results

Prosecutor assesses strength of DNA evidence, offers plea bargain, and decides how to try case.

On-demand case report



Cybergenetics

160 North Craig Street, Suite 210
Pittsburgh, PA 15213
Tel: (412) 683-3004
Fax: (412) 683-3005

September 15, 2011

TO: ALLEGHENY COUNTY DISTRICT ATTORNEY'S OFFICE
COURTHOUSE, 436 GRANT STREET, 3RD FLOOR
PITTSBURGH, PA 15219

SUPPLEMENTAL REPORT
Allegheny Lab # 10LAB00000

Victim(s): SMITH, Jane

Suspect(s): DOE, John

Evidence Item(s):

Item 2D Control cutting from the interior of the crotch panel
Item 1P Reference sample from Jane Smith
Item 4 Buccal swab from John Doe

METHODS:

- The DNA PowerPlex® 16 data profiles referenced in this report were previously developed by the Allegheny County Medical Examiner, as described in their case report.
- The TrueAllele® Casework system processed each evidence item in independent replicate computer runs to infer possible DNA contributor genotypes from the samples.
- The DNA match statistics calculated herein used the population allele frequencies generated by the Promega Corporation, and a theta value (co-ancestry coefficient) of 1%.

RESULTS:

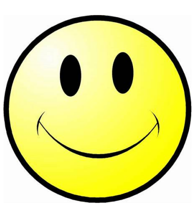
TrueAllele assumed that the evidence sample data (Item 2D) contained one or two unknown contributors, and objectively inferred evidence genotypes solely from these data, both with and without a victim reference (Item 1P). Degraded DNA was considered. Following genotype inference, the computer then compared genotypes from these evidence items to a provided reference (Item 4) genotype, relative to reference populations, to compute likelihood ratio (LR) DNA match statistics.

Based on these results, a match between the crotch panel (Item 2D) and John Doe (Item 4) is:
1.35 quadrillion times more probable than a coincidental match to an unrelated Black person,
426 quadrillion times more probable than a coincidental match to an unrelated Caucasian person, and
18.8 quadrillion times more probable than a coincidental match to an unrelated Hispanic person.

Mark W. Perlin, PhD, MD, PhD
Chief Scientific Officer, Cybergenetics

- prosecutor sets trial date
- Cybergenetics lets TrueAllele further explore the evidence
- Cybergenetics sends the case report to the prosecutor

TrueAllele evidence



- objective computer results
- thoroughly explored data
- addressed data challenges
- likelihood ratio match statistic
- usually a million times more DNA identification information
- sometimes determines no statistical support for a match

DNA justice is served

Role of Cybergenetics scientist

- confer with prosecutor
- prepare exhibits for court
- appear in court
- direct and cross examination
- follow up on case

Case referral from lab analyst

I think this is a perfect example to demonstrate the power of True Allele vs. conventional methods.

The questioned sample is a control area from clothing of the victim. The sample is a mixture with the major contributor matching the victim.

The suspect could not be excluded as a minor contributor. Since we don't resolve the minor contributor's genotypes, a CPI was used.

Also, two loci were excluded from the calculations because two of the minor alleles fell within stutter positions and were probably filtered out by our conventional software's stutter filters.

TrueAllele investigation

Received lab data: Friday afternoon, 3:00 pm

Preliminary report: Monday morning, 8:30 am

Email to the prosecutor:

The interior crotch panel is a mixture that has a 15% minor contributor that reproducibly matches the suspect.

Statistically, a match between the suspect and the evidence is about a quadrillion (15 zeros) times more probable than coincidence.

TrueAllele evidence

TrueAllele assumed that the evidence sample data (Item 2D) contained one or two unknown contributors, and objectively inferred evidence genotypes solely from these data, both with and without a victim reference (Item 1P). Degraded DNA was considered. Following genotype inference, the computer then compared genotypes from these evidence items to a provided reference (Item 4) genotype, relative to reference populations, to compute likelihood ratio (LR) DNA match statistics.

Based on these results, a match between the crotch panel (Item 2D) and the suspect (Item 4) is:

- 1.35 quadrillion times more probable than a coincidental match to an unrelated Black person,
- 426 quadrillion times more probable than a coincidental match to an unrelated Caucasian person, and
- 18.8 quadrillion times more probable than a coincidental match to an unrelated Hispanic person.

Conclusions

- TrueAllele computing expands human capability
- Scientifically validated and peer reviewed
- Satisfies SWGDAM and regulatory guidelines
- Resolves "inconclusive" DNA evidence
- Have issued about 75 case reports so far
- Most for prosecution, some for defense
- Public-private partnership complements DNA lab
- Cybergenetics offers both products and services

Learning More

The science of quantitative DNA mixture interpretation

www.cybgen.com/information

- Newsletters
  gentle introduction to ideas
- Courses
  for scientists and lawyers
- Presentations
  talks, handouts, movies, transcripts
- Publications
  abstracts, manuscripts, papers