

# TrueAllele® Computational Forensics for Solving “Uninterpretable” DNA Mixtures

Mark W Perlin, PhD, MD, PhD  
Cybergenetics  
Dharwad, Karnataka, India

International Conference on  
Recent Trends in DNA Forensics  
November 2023



Cybergenetics © 2003-2023

---

---

---

---

---

---

---

---

## Cybergenetics history

Innovating the interpretation of forensic DNA evidence

- 1994. Founded in Pittsburgh, Pennsylvania.
- 1995. Solved the PCR stutter problem.
- 2000. Eliminated UK DNA database backlog.
- 2001. Solved the DNA mixture problem – TrueAllele.
- 2005. Identified World Trade Center remains.
- 2009. Testified in world's first prob genotyping case.
- 2016. First prob genotyping DNA exoneration.
- 2020. Validated on 10-unknown mixtures.
- 2022. Reported on over 1,000 DNA cases.
- 2023. Used in over 90% of the 50 USA states.



---

---

---

---

---

---

---

---

## Computational Genomics



Genomics  
Volume 28, Issue 2, 20 July 1995, Pages 315-327



Regular Article  
Rapid Construction of Integrated Maps  
Using Inner Product Mapping: YAC Coverage  
of Human Chromosome 11

Mark W. Perlin, David J. Duggan, Kotyo Davis, Joan E. Faur, Robert B. Findler, Michael J. Higgins,  
Norma J. Howes, Glen A. Evans, Shizhen Qin, Jialu Zhang, Thomas B. Shows, Michael B. James,  
Charles W. Richard III

Show more

+ Add to Mendeley Share Cite

<https://doi.org/10.1006/jgeno.1995.1148>

Get rights and content

---

---

---

---

---

---

---

---

# STR genotyping bottleneck

1995. Cybergenetics solves the DNA analysis problem.  
Replaces ineffective human review with efficient computers.

*Am. J. Hum. Genet.* 57:1199–1210, 1995

## Toward Fully Automated Genotyping: Genotyping Microsatellite Markers by Deconvolution

Mark W. Perlin,<sup>1</sup> Giuseppe Lancia,<sup>2</sup> and See-Kiong Ng<sup>1</sup>

<sup>1</sup> Computer Science Department and <sup>2</sup> Graduate School of Industrial Administration, Carnegie Mellon University, Pittsburgh

**United States Patent** (19) (11) Patent Number: 5,541,067  
 Perlin (42) Date of Patent: Jul. 30, 1996

(54) METHOD AND SYSTEM FOR GENOTYPING

(56) References Cited

(\*) Notice: The portion of this patent abstract that appears in this journal has been disclosed.

(21) Appl. No. 314,060  
 (22) Filed: Sep. 20, 1994

OTHER PUBLICATIONS

Perlin et al. "Toward Fully Automated Genotyping: Genotyping Microsatellite Markers By Deconvolution", *Am. J. Hum. Genet.* 57:1199-1210 (1995).

Schwartz et al. "Fluorescent multiplex linkage analysis and other methods for nonradioisotopic molecular biology" *Am. J. Hum. Genet.* 51: 721-729, 1992.

Primary Examiner: W. Gary Jones  
 Attorney: Lawrence J. Zeman, Esq.  
 Attorney Agent or Firm: Anand M. Schwab

---

---

---

---

---

---

---

---

---

---

# DNA database backlog

Year 2000. Britain. The FSS.  
The first national DNA database.  
Human review of DNA data fails.  
Backlog of 350,000 swabs.



Enter computer automation.  
Cybergenetics **TrueAllele** clears the backlog.  
Every year, does another 350,000 samples.

**Accurate:** eliminates human error (1 in 2,000).  
**Fast:** "swab to DB" turnaround time in 5 days.  
**Labor-saving:** from 100 people down to 6.  
**Inexpensive:** 2-3 computers, not 100 people.

---

---

---

---

---

---

---

---

---

---

# DNA mixtures



eye of newt



toe of frog



Double, double toil and trouble

---

---

---

---

---

---

---

---

---

---

### DNA reporting language

“Due to the data being **uninterpretable**, no comparison can be made to the reference sample.”

“A *match* between the evidence and the suspect is a trillion times **more probable** than coincidence.”

---

---

---

---

---

---

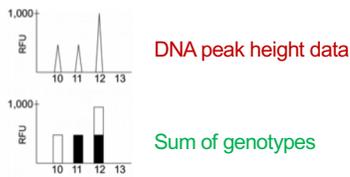
---

---

### DNA mixtures “uninterpretable”?<sup>8</sup>

JOURNAL OF FORENSIC SCIENCES  
2001 Mark W. Perlin,<sup>1</sup> Ph.D., M.D., Ph.D. and Beata Szabady,<sup>1</sup> Ph.D.

Linear Mixture Analysis: A Mathematical Approach to Resolving Mixed DNA Samples



---

---

---

---

---

---

---

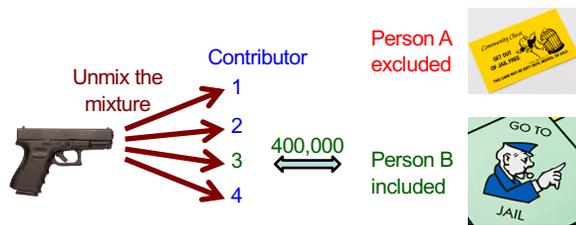
---

### Informative genotyping

(Inconclusive or “uninterpretable” DNA mixture)

Cybergenetics TrueAllele analysis **unmixes mixtures**

Match statistics provide **information**



---

---

---

---

---

---

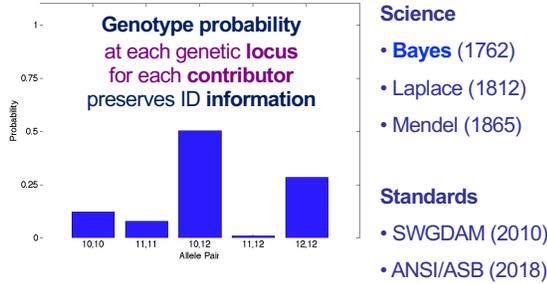
---

---

MW Perlin, MM Legler, CE Spencer, JL Smith, WP Allan, JL Belrose, BW Duceman<sup>10</sup>  
 Validating TrueAllele<sup>®</sup> DNA mixture interpretation. Journal of Forensic Sciences, 2011.

## Separated genotypes

Genotype uncertainty expressed through probability




---

---

---

---

---

---

---

---

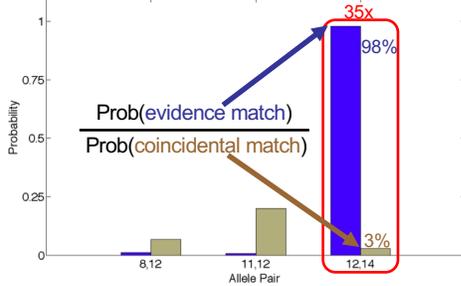
---

---

11

## Quantify DNA match information

How much more does the suspect match the evidence than a random person? – the likelihood ratio




---

---

---

---

---

---

---

---

---

---

12

## World Trade Center attack

September 11, 2001 in New York City



18,500 remains



2,700 victims

---

---

---

---

---

---

---

---

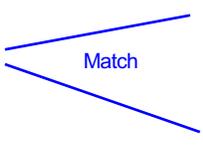
---

---

### Identify victim remains

Victim Remains

Missing People



---

---

---

---

---

---

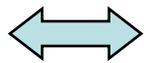
---

---

### Automated TrueAllele® system



VUler™ Software



Parallel Processing & Database Servers

- Fully automated
- Uses all the data
- No thresholds or choices
- No calibration – case data

---

---

---

---

---

---

---

---

### Bus crash mass disaster



- 2008 bus crash in Komatipoort, South Africa
- Police recover burned victim remains
- Relatives submit DNA to help identify remains
- Lab was unable to identify victim remains

---

---

---

---

---

---

---

---

## Bus crash in South Africa TrueAllele DVI using relatives

16

From "insufficient information" to identified victims

Victim Remains	Sample	Relation
AO0553E	AP2438C	Grandfather
AO0554E	AP2439C	Daughter
AO0555E	AP2440C	Son
AO0556E	AP2441C	Mother
AO0557E	AP2442C	Son
AO0558E	AP2443C	Sister
AO0559E	AP2444C	Father
AO0560E	AP2445C	Son
AO0561E	AP2446C	Unknown
AO0562E	AP2447C	Brother
AO0563E	AP2448C	Brother
AO0564E	AP2449C	Sister
AO0565E	AP2450C	Mother
AO0566E	AP2451C	Son
AO0567E	AP2452C	Father
	AP2455C	Son

---

---

---

---

---

---

---

---

---

---

## 2022 Chandigarh workshop

17

- 1 TrueAllele database server
- 12 interpretation processors
- 20 students (3<sup>rd</sup> training day)

### Task

Have each student use TrueAllele to identify the victim remains

---

---

---

---

---

---

---

---

---

---

## Student success!

18

15 victim items	→	15 victim genotypes
16 family items	→	16 family genotypes
16 family references	→	16 kinship genotypes
DNA data		Genotypes

15 victim genotypes x 16 kinship genotypes = 240 genotype comparisons

240 comparisons x 20 students = 4,800 total comparisons

Before lunch:	Students uploaded the genotype data
During lunch:	TrueAllele solved all the problems
After lunch:	Students reviewed identification results

---

---

---

---

---

---

---

---

---

---

## Manual mixture interpretation

- Inaccurate. Disagrees with true information
- Subjective. Workflow introduces human bias
- Widespread. Millions of case items
- Opaque. Choices use only some of the data
- Biased. Can only include – or give no answer

**Inconclusive (case stays open)**

---

---

---

---

---

---

---

---

## TrueAllele® computer technology

- Accurate. 42 validation studies, 8 published
- Objective. Workflow removes human bias
- Accepted. Used in 90% of states, and by labs
- Transparent. Give math, software (4GB DVD)
- Neutral. Statistics inclusionary or exclusionary

**Informative (case is closed)**

---

---

---

---

---

---

---

---

## How is TrueAllele used?

- Prosecution
- Defense
- Investigation
- Post-conviction
- Mass disaster
- Touch DNA
- Complex mixtures
- Kinship, paternity
- DNA database
- Familial search
- Preventing crime

---

---

---

---

---

---

---

---

22

Homicide with 7% DNA under victim's fingernails

## Commonwealth vs. Foley

Apr 2006: Blairsville Dentist John Yelenic murdered  
 Nov 2007: Trooper Kevin Foley charged with crime



Feb 2008: Defense questions 13,000 DNA match score

---

---

---

---

---

---

---

---

23

## Three DNA match scores

<u>Score</u>	<u>Method</u>
13 thousand	FBI Lab
23 million	Dr. Cotton
189 billion	TrueAllele

Same DNA data, better data analysis

---

---

---

---

---

---

---

---

24

## Virginia v. David Black (Kinship homicide)

Item	Description	David Black	Bonnie Black	BettyAnn Armstrong	Craig Black	Eleanora Black
08	Baseball hat velcro strap	32.5 quintillion	16.1 billion		1/1.83 thousand	1/62.6
94	Master bedroom light switch	364 million	8.14 million			
95	Master bathroom light switch	1/19.5	554 million			3.63 million

---

---

---

---

---

---

---

---

## Pennsylvania v. Akaninyene Akan (Serial rape)



Akaninyene Eftong Akan

- Female students raped in their Pittsburgh apartments
- Akan accused of serial crimes; tossed cigarette, victim's clothing
- TrueAllele found a DNA match statistic of a **quadrillion**
- Sentenced to 32 years in prison

---

---

---

---

---

---

---

---

## Pennsylvania v. Terry Lipinsky (Incest rape)

From "uninterpretable DNA mixture" to closed case

Item	Description	10 – Victim	10 – Suspect
		Daughter	Terry Lipinski
3A	Underwear	17.25	0.03
6A	Boxer shorts – crotch area	0.75	1.60
6B	Boxer shorts – interior rear	-24.17	16.31
7A	T-shirt – stain area 1	16.02	15.66
7B	T-shirt – stain area 6	5.83	12.56
7C	T-shirt – stain area 7	11.07	12.89
12A	Bra	8.75	6.70

---

---

---

---

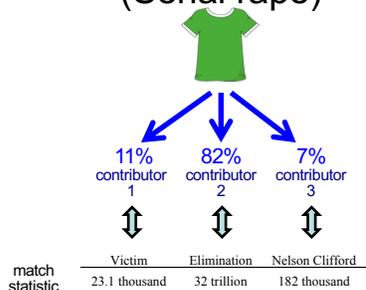
---

---

---

---

## Maryland v. Nelson Clifford (Serial rape)




---

---

---

---

---

---

---

---

# Indiana v. Darryl Pinkins (TrueAllele exoneration)




---

---

---

---

---

---

---

---

# Information and Automation

- 1. **Information**
  - a) No "inconclusive", close cases
  - b) Accurately weigh DNA evidence
- 2. **Automation**
  - a) Computer does all the work
  - b) Without thousands more people
- 3. **Easy to use**
  - a) Intuitive user interface
  - b) Ask questions in seconds




---

---

---

---

---

---

---

---

# Peer-reviewed validation studies<sup>30</sup>

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

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. *Science & Justice*. 2013;53(2):103-114.

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

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

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

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

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

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.

---

---

---

---

---

---

---

---

**7. TRUEALLELE VALIDATION** JOURNAL OF FORENSIC SCIENCES **31**  
*J Forensic Sci.* November 2011, Vol. 56, No. 6  
 doi: 10.1111/j.1556-4029.2011.01859.x  
 Available online at: [onlinelibrary.wiley.com](http://onlinelibrary.wiley.com)

**PAPER**

**CRIMINALISTICS**

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

Validating TrueAllele® DNA Mixture Interpretation\*<sup>†</sup>

---

---

---

---

---

---

---

---

---

---

---

---

JOURNAL OF FORENSIC SCIENCES **32**  
*J Forensic Sci.* November 2013, Vol. 58, No. 6  
 doi: 10.1111/1556-4029.12223  
 Available online at: [onlinelibrary.wiley.com](http://onlinelibrary.wiley.com)

**PAPER**

**CRIMINALISTICS**

Mark W. Perlin,<sup>1</sup> M.D., Ph.D.; Jamie L. Belrose,<sup>2</sup> M.S.; and Barry W. Duceman,<sup>3</sup> Ph.D.

New York State TrueAllele® Casework Validation Study\*

---

---

---

---

---

---

---

---

---

---

---

---

JOURNAL OF FORENSIC SCIENCES **33**  
*J Forensic Sci.* 2019  
 doi: 10.1111/1556-4029.14204  
 Available online at: [onlinelibrary.wiley.com](http://onlinelibrary.wiley.com)

**PAPER**

**CRIMINALISTICS**

David W. Bauer,<sup>1</sup> Ph.D.; Nasir Butt,<sup>2</sup> Ph.D.; Jennifer M. Hornyak,<sup>1</sup> M.S.; and Mark W. Perlin,<sup>1</sup> Ph.D., M.D., Ph.D.

Validating TrueAllele® Interpretation of DNA Mixtures Containing up to Ten Unknown Contributors\*

TABLE 4—Independent analysis.

	Operator Site	CCRFSL							
			13	7	K	K	K	K	K
			22	6	7	K	K	K	K
			12	5	4	5	K	K	K
Genotypes	78	78	16	4	4	5	6	K	K
Minimum	-5.16	-9.14	13	4	3	2	1	6	K
Mean	8.36	8.48	15	3	3	4	1	6	8
Median	5.98	5.61	2	1	1	1	1	3	3
Maximum	29.03	29.12	2	0	2	2	3	2	3
SD	8.37	8.54	4	0	1	1	1	2	0
$\sigma_{\alpha}$		0.70	1	0	0	0	0	-1	-1

TABLE 7—Peeling sensitivity.

Mixture weight (%)	Peeling Round						
	0	1	2	3	4	5	6
7	K	K	K	K	K	K	K
6	7	K	K	K	K	K	K
5	4	5	K	K	K	K	K
4	4	5	6	K	K	K	K
3	3	4	1	6	K	K	K
2	1	1	1	1	3	3	4
1	0	2	2	3	2	3	4
0	1	1	1	2	0	2	2
0	0	0	0	0	-1	-1	0

---

---

---

---

---

---

---

---

---

---

---

---

## More TrueAllele information

<http://www.cybgen.com/information>



- Cloud & demo
- Courses
- Newsroom
- Presentations
- Publications
- Webinars

<http://www.youtube.com/user/TrueAllele>  
TrueAllele YouTube channel



Cybergenetics



---

---

---

---

---

---

---

---