

Real-Time DNA Investigation TrueAllele® System 3

16th International Symposium on Human Identification
Sponsored by Promega Corporation
September, 2005

Mark W Perlin, PhD, MD, PhD
Pittsburgh, PA USA
perlin@cybgen.com



Cybergenetics

Cybergenetics © 2003-2005

DNA Match Prevents Crime

UK: 45% hit rate
property crimes match offenders
VA: 40% hit rate
sexual assaults match nonviolent

Fast, complete DNA processing:
prevent 300 stranger rapes/day

DNA Match Scenarios



TrueAllele® History

1994 - Automated STR Genotyping

- AJHG paper published (stutter deconvolution)
- patents filed (6 issued) – licensing
- NIH grant funded

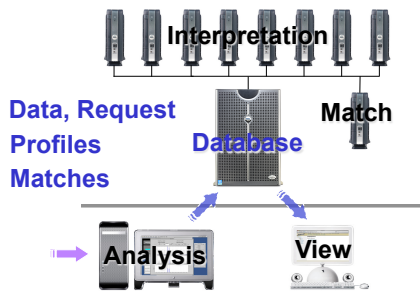
2000 - FSS System 2 Deployment

- databank: reduced time, error, people
- > 1,000,000 databank & property crime samples

2000 - TrueAllele System 3

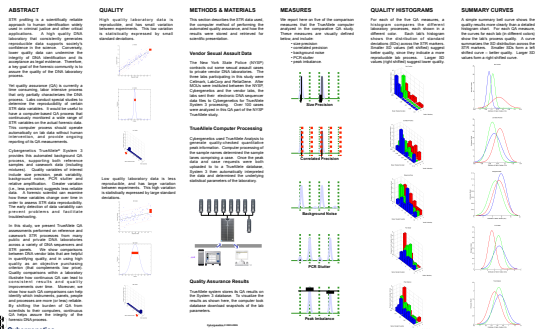
- casework applications (mixture deconvolution)
- property crime, sexual assault, mass disaster
- goal: *faster, better, cheaper* than people

TrueAllele® System 3




#8 Automated Quality Assurance for the Forensic DNA Laboratory A Cybergenetics TrueAllele® System 3 Study

Mark W. Perlin, PhD, MD, PhD Pittsburgh, PA USA



Property Crime


unknown profiles



Crime Scene

MATCH

reference profiles




Offenders

New York State Police; example process

Time?
Cost?

Cybergenetics trueallele.net

time
08:57 pm
08:57 pm
08:58 pm
08:59 pm
08:58 pm
08:59 pm
09:00 pm
09:00 pm
09:00 pm
09:01 pm
09:01 pm
09:02 pm
09:02 pm
09:02 pm
09:03 pm
09:03 pm
09:03 pm
09:04 pm
09:04 pm
09:05 pm
09:05 pm
09:05 pm
09:05 pm



Welcome to trueallele.net, where automated casework interpretation provides rapid DNA response.

4 interpret computers

64 cases

23 minutes

3 cases/minute
(or 1,000,000 cases/year)

Copyright Cybergenetics 2003-2005. All Rights Reserved.

User
Login
Register

Report
New

Technology
TrueAllele
Processing
Reporting
Patents

Contact
Company
Sales
Support

Productivity

1,000,000 cases

x (2 reviews/case)

x (1 analyst/200 reviews)

= 10,000 analysts 1,000x faster

x (\$100,000/analyst)

= \$1,000,000,000 1,000x cheaper

Sexual Assault

unknown profiles



Crime Scene

reference profiles



Offenders

MATCH

Allegheny County; screening process

Orchid Cellmark; vendor process

Information?

NIJ Mixture Study Design

Mixture ratio	ng amplified			
	1	0.5	0.25	0.125
10:0	A1	A2	A3	A4
9:1	B1	B2	B3	B4
7:3	C1	C2	C3	C4
5:5	D1	D2	D3	D4
3:7	E1	E2	E3	E4
1:9	F1	F2	F3	F4
0:10	G1	G2	G3	G4

1, 1/2, 1/4, 1/8 ng

Mixture ratio	ng amplified			
	1	0.5	0.25	0.125
10:0	H1	H2	H3	H4
9:1	I1	I2	I3	I4
7:3	J1	J2	J3	J4
5:5	K1	K2	K3	K4
3:7	L1	L2	L3	L4
1:9	M1	M2	M3	M4
0:10	N1	N2	N3	N4

STR Data Generation

- premixed DNA templates: NIST
- lab protocols: Cybergenetics
- data generation (ten DNA labs)
Florida, New York, Ohio, Pennsylvania,
Virginia, Cellmark, UK FSS, Cybergenetics
- DNA sequencers:
FMBio/II, 377, 310, 3100, 3700
- STR panels: PowerPlex (1, 2, 16),
ProfilerPlus, Cofiler, SGMplus, Identifier

No Suspect Test Case

1 ng DNA, PowerPlex16, ABI/310

Contributors

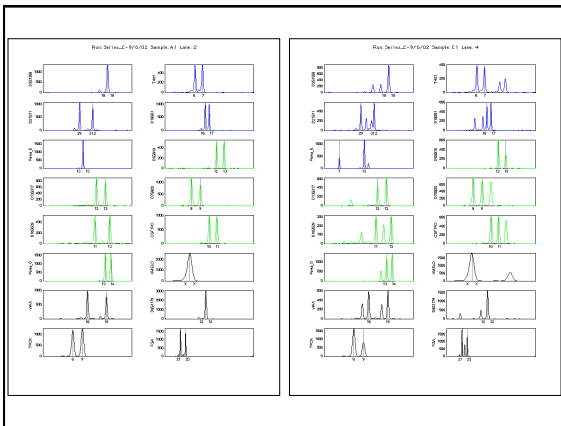
A: Victim

G: Unknown Suspect

Samples

A1 (Victim)

C1 (Mixture) 70% A + 30% G



Cybergenetics trueAllele.net User: markperlin

1 report 2 search 3 select 4 review

Study: N/A Part: CYB310 Client: CYB Request: A1C1 Process: onunknown

profile power match strength mature data

peak height cutoff: 100 (update)

select template

specimen	reference	lab	cutting	extract	prep
A1	yes	CYB	1	standard	1
C1	no	CYB	1	standard	1

Technology TrueAllele Processing Reporting Patents

Contact Company Sales Support

lane tag	laboratory	STR	PCR	amp	sequencer	run	lane
A	Cybergenetics	PowerPlex 16	Standard	1	ABI310	Series_C-9/6/02	4

new window | download

lane tag	locus	peak	length	desig	height	area
A	AMELO	5	106.0	1.0	1854	929
A	AMELO	7	112.0	2.0	580	275
A	CSF1PO	5	337.0	10.0	653	397
A	CSF1PO	6	341.1	11.0	666	399
A	CSF1PO	7	345.0	12.0	543	281
A	D13S317	3	177.1	9.0	144	72
A	D13S317	9	188.0	12.0	484	197
A	D13S317	10	192.0	13.0	622	272
A	D16S539	5	280.0	9.0	135	62
A	D16S539	8	287.9	11.0	327	150
A	D16S539	9	292.0	12.0	221	101
A	D16S539	10	296.0	13.0	318	145
A	D18S51	3	310.0	13.0	266	160
A	D18S51	6	318.0	15.0	303	183

User: markperlin

[1 report](#) [2 search](#) [3 select](#) [4 review](#)

Study: NI/mixture Part: CV9310 Client: CYB Request: A1G1 Process: oneunknown

[profile](#) [power](#) [match](#) [strength](#) [mixture](#) [data](#)

User
[Logout](#)
[Register](#)

Report
[New](#)

Technology
[TrueAllele](#)
[Processing](#)
[Reporting](#)
[Patents](#)

Contact
[Company](#)
[Sales](#)
[Support](#)

	Template	
	A1	C1
1	1.000	0.680
2	0.000	0.320

Copyright Cybergenetics 2003-2005. All Rights Reserved.

Mark W. Perlin,¹ Ph.D., M.D., Ph.D., and Beata Szatholy,¹ Ph.D.

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

BYERENTZ, P. *Ph.D., M.D., Ph.D., and Beata Szatholy,¹ Ph.D.*

ABSTRACT: With the advent of the TrueAllele DNA profile processing algorithm, the need to resolve mixed DNA samples (MDS) has become a critical forensic issue. This paper introduces a novel mathematical approach to resolving MDS, based on the use of a Bayesian network model. The model is able to resolve MDS with a high level of accuracy and is able to provide a measure of uncertainty for each profile. This approach is a novel and powerful method for resolving MDS, and is able to provide a measure of uncertainty for each profile.

KEYWORDS: TrueAllele, Bayesian network, MDS, forensic DNA, profile processing, uncertainty, probability, forensic DNA, profile processing, uncertainty, probability, forensic DNA, profile processing, uncertainty, probability.

In this paper, we introduce the Linear Mixture Analysis (LMA) model, and then provide a detailed description of the model. We describe several practical applications, each one based on a particular subset of data available in the mixture. We then focus on laboratory data and provide some important mixture problems before reviewing the method in other studies. We conclude with some observations on the practical application of LMA.

Linear Mixture Analysis (LMA) is a novel and powerful method for resolving MDS, based on the use of a Bayesian network model. The model is able to resolve MDS with a high level of accuracy and is able to provide a measure of uncertainty for each profile.

In the PCR amplification of a mixture, the amount of each PCR product varies in rough proportion to relative weights of each

DNA Profile Uncertainty

Preserve match information

genetic

data

Use probability

User: markperlin

[1 report](#) [2 search](#) [3 select](#) [4 review](#)

Study: NI/mixture Part: CV9310 Client: CYB Request: A1G1 Process: oneunknown

[profile](#) [power](#) [match](#) [strength](#) [mixture](#) [data](#)

User
[Logout](#)
[Register](#)

Report
[New](#)

Technology
[TrueAllele](#)
[Processing](#)
[Reporting](#)
[Patents](#)

Contact
[Company](#)
[Sales](#)
[Support](#)

allele call confidence level: 0.999 [\(update\)](#)

select contributor

contrib	power	template
1	16.397	A1
2	17.301	A1

new window | download

locus	desig 1	desig 2	probability	frequency
AMELO	1.0	2.0	1.000	0.4950
CSF1PO	12.0	12.0	1.000	0.0894
D13S317	8.0	13.0	1.000	0.0159
D16S539	9.0	12.0	1.000	0.0832
D18S51	13.0	15.0	1.000	0.0288
D21S11	30.0	31.0	1.000	0.0345
D3S1358	16.0	17.0	1.000	0.1128
D5S818	12.0	12.0	1.000	0.1181
D7S820	10.0	10.0	1.000	0.0921
DES1179	8.0	11.0	1.000	0.0112
FGA	21.0	22.0	1.000	0.0512
Penta_D	12.0	14.0	1.000	0.0107
Penta_E	7.0	14.0	1.000	0.0168
TH01	9.0	9.3	1.000	0.0586
TPOX	8.0	8.0	1.000	0.1748
vWA	15.0	18.0	1.000	0.0549

Copyright Cybergenetics 2003-2005. All Rights Reserved.

Cybergenetics © 2003-2005

6

DNA Information

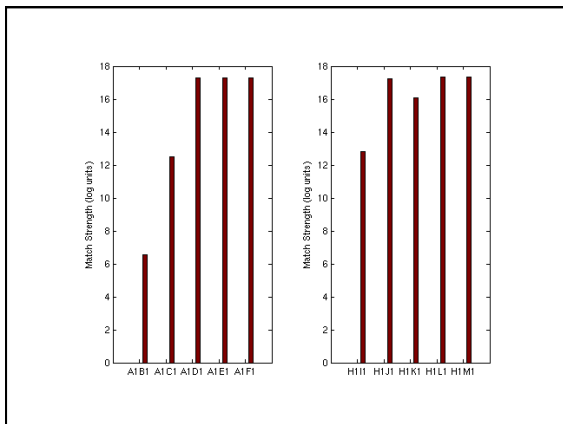
$$\text{Match Strength} = \frac{\text{Prob}(\text{match})}{\text{Prob}(\text{random})}$$

Match Strength logarithm –

Very large number:
“billion billion” or 10¹⁸

Use the exponent:
18

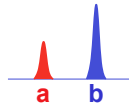
The screenshot shows the TrueAllele web interface. At the top, it says 'Cybergenetics trueallele.net User: markparin'. Below that are navigation tabs: '1 report', '2 search', '3 select', '4 review'. A sub-menu includes 'profile', 'power', 'match', 'strength', 'mixture', and 'data'. The 'match' tab is active, showing a 'match strength cutoff: 6' and an 'Update' button. Below this is a table titled 'select matching profile' with columns: source, contrib, target, case, label, strength, hit, miss. The table shows one entry: 'CV6 2 NIST G 17.501 13 0'. Below the table is a 'new window | download' link and a larger table with columns: locus, design 1, design 2, prob A, prob B, probAAB, and freq. The table lists various loci such as AMELO, CSF1PO, D13S317, etc., with their respective design numbers and probabilities.



Conservative Human Review

Avoid overcalling the results

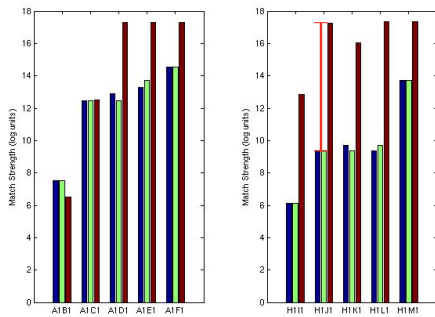
uncertain data

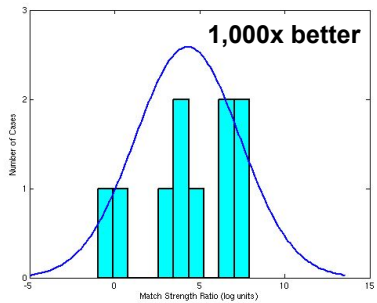


Allele 1: b

Allele 2: anything

Report 0, 1 or 2 alleles

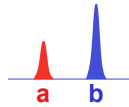




Aggressive Human Review

Try ruling out unlikely combinations

uncertain data

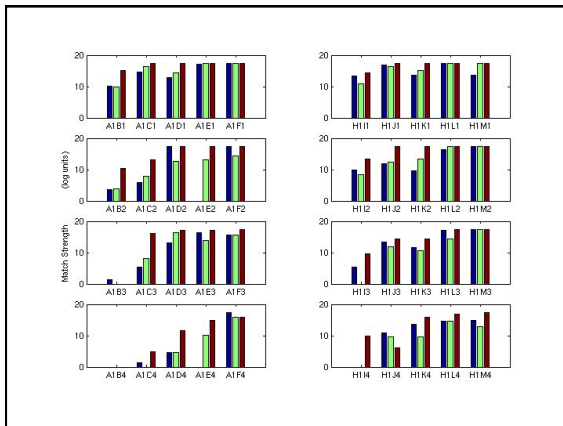


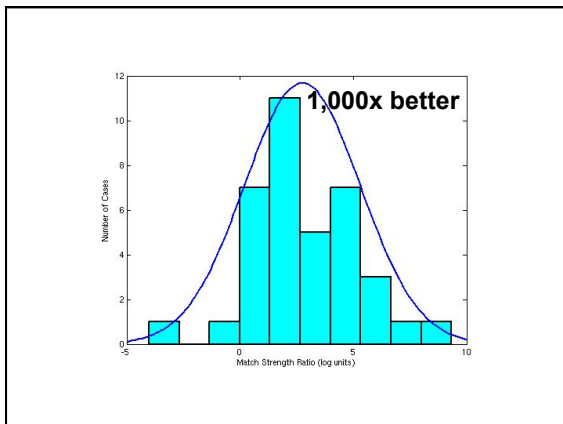
Allele calls:

1. a b

2. b b

Report list of allele calls







Mass Disaster

unknown profiles reference profiles

Personal Effects


World Trade Center — **MATCH** — 

Victim Remains — **MATCH** — 

Family References


Status?

Real-Time TrueAllele Investigation



1,000x faster
1,000x better
1,000x cheaper

Rapid DNA policing prevents crime

 **Cybergenetics** www.trueallele.net

<p>Scientific Collaborators</p> <p>Jeff Ban, DFS Robin Cotton, Cellmark Cecelia Crouse, PSBO Barry Duceman, NSYP Trevor Howett, FSS Jay Kadane, CMU Bob Shaler, OCME</p> <p>Participating Labs</p> <p>Albany, NY Allegheny County, PA Cellmark, MD FSS, UK Miami Valley, OH New York, NY NIST, MD Palm Beach, FL Richmond, VA</p> <p> Cybergenetics</p>	<p>Cellmark Case Review</p> <p>Christine Baer Jason Befus Julie Black Paula Clifton Kathryn Colombo Lisa Grossweiler Juliet Harris Jeff Hickey Jacki Higgins Christopher Knickerbocker Jason Kokoszka Lewis Maddox Jennifer Reynolds Leslie Rosier Ryan Satcher Alissa Shofkom Margaret Terrill Melissa Thompson Charlotte Word Cindy Zimmerman</p>	<p>NYSP Case Review</p> <p>John Brenner Russell Gettig Melissa Lee Maria Mick Shannon Morris Urfan Mukhtar Michael Portzer Laura Post Diana Seaburg</p> <p>Cybergenetics Staff</p> <p>Bill Allan Meredith Clarke Matt Legler Donna Scheuble Alex Sineinikov</p> <p>NIJ Grant Award #2001-IJ-CX-K003</p>
---	--	--

Real-Time TrueAllele Investigation



1,000x faster
1,000x better
1,000x cheaper

Rapid DNA policing prevents crime



Cybergenetics

www.trueallele.net
