Fighting for DNA Justice: Genotyping Software in the Hillary Acquittal

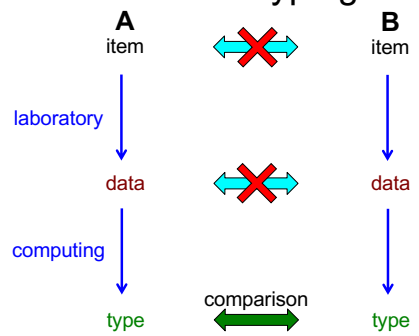
Questioning Forensics: Inside the Black Box
The Legal Aid Society of New York City
October, 2016
New York, NY

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

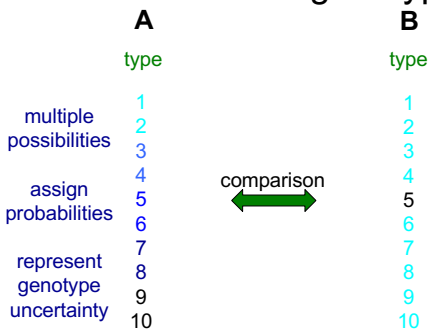
Cybergenetics

Cybergenetics © 2003-2016

Genotyping



"Probabilistic" genotyping

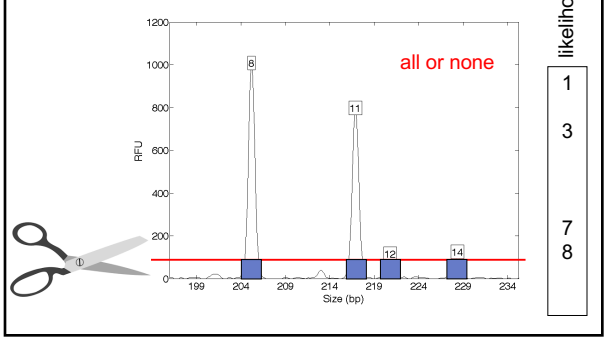


Bayes rule

A				B
prior type	likelihood of data	posterior type		type
1	1	1		1
2	2	2		2
3	3	3		3
4	4	4	comparison ⟺	4
5	5	5		5
6	6	6		6
7	7	7		7
8	8	8		8
9	9	9		9
10	10	10		10

X =

Threshold & CPI



likelihood

1

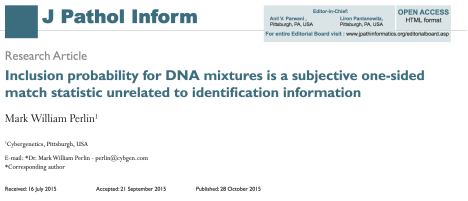
3

7

8

CPI lacks scientific basis

Misled for 15 years on millions of DNA mixtures



J Pathol Inform

Research Article

Inclusion probability for DNA mixtures is a subjective one-sided match statistic unrelated to identification information

Mark William Perlin[1]

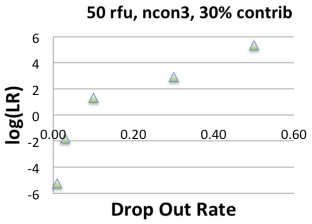
[1]Cybergenetics, Pittsburgh, USA

E-mail: *Dr. Mark William Perlin - perlin@cybgen.com
*Corresponding author

Received: 16 July 2015 Accepted: 21 September 2015 Published: 28 October 2015

Dropout & FST

Nature vs. parameter



FST precluded

Judge Dwyer Issues Written Decision in Landmark DNA Case Won by Legal Aid's DNA Unit

WEDNESDAY, JULY 15, 2015

After two and a half years of litigation, including live testimony from 11 scientists, Acting Brooklyn Supreme Court Justice Mark Dwyer decided not to admit so-called "lower copy number" or "high-sensitivity" analysis. Lawyers in Legal Aid's DNA Unit litigated the admissibility of LCN and FST evidence in a Frye hearing.

New York v. Nick Hillary

Garrett Phillips (12)
Died from strangulation
October 24, 2011

Oral "Nick" Hillary
Arrested for murder
May 15, 2013

DNA evidence

Clothes 22	shorts long sleeve shirt black long sleeve shirt
Tile 3	material from tile crack hair small black fuzz
Window 3	window and screen wood sill stone sill
Victim 7	oral swab right hand left hand left thumb neck

Cybergenetics

- Pittsburgh, Pennsylvania
- Company founded in 1994
- Extracts information from genetic data
- Medical diagnostics, mass disasters
- DNA mixture analysis

June, 2013
New York State Police sent Cybergenetics
Hillary DNA evidence for free screening
Focus on victim's left hand fingernails

Cybergenetics

TrueAllele® technology

1994. Solved the PCR stutter problem
1999. Solved the DNA mixture problem
2009. 25th version used in criminal case

- Accurate. 34 validation studies, 7 published
- Objective. Workflow removes human bias
- Accepted. 10 Daubert/Frye challenges
- Transparent. Give math, software (4GB DVD)
- Neutral. Can statistically include or exclude

Objective workflow

(1) Enter all data

(2) Calculate statistic

(3) Math decides

- Keep people out of the process
- Because software is robust
- And eliminate human bias

TrueAllele findings in Hillary

2013. 26 Identifiler tests on left fingernails
Mixture of 95% victim + 5% other
No statistical connection to Hillary
Advised Minifiler for degraded DNA

2014. More lab data on left fingernails
9 tests using new kit & machine
NYSP requested TrueAllele analysis
Again, no connection to Hillary

STRmix™ software

- uses thresholds, dropout
- adds peak heights and some modeling
- limited likelihood explanation of data
- requires calibration of key variables

- solves easy mixtures quite well
- good CPI replacement, better statistics
- retains conventional workflow

It's been a very successful year for ESR in many ways. ... ESR's financial results improved markedly, with a record profit achieved. We also made good progress in executing our strategy, growing international markets for our science services and developing new products. – ESR, 2016 Annual Report, New Zealand

Conventional DNA workflow

(1) Choose data

(3) Person decides

(2) Calculate statistic

- Put people in the process
- To overcome software limits
- And introduce human bias

January. NZ results

STRmix software developers consult for prosecution using version 2.3

Find a match statistic of ten million, connecting Phillips' fingernails to Hillary's DNA

Their match statistic lowers to ten thousand, when they remove forward stutter peaks

Contributor under 0.5%, 1:200

Legal Aid Society of NYC

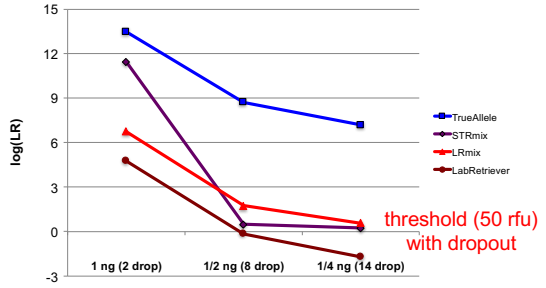
- Contacted Cybergenetics
- Sent over STRmix declaration
- Initial discovery not helpful

- Requested additional discovery
- Computer results file important
- Choose data to get result

An investigation of software programs using "semi-continuous" and "continuous" methods for complex DNA mixture interpretation.
Coble M, Myers S, Klaver J, Kloosterman A, Leiden University, The Netherlands, 9th International Conference on Forensic Inference and Statistics, 2014.

STRmix threshold impact

Simple two person mixture, 10% minor contributor

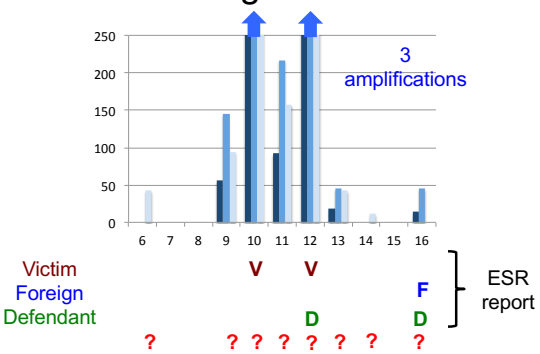


Varying fingernail threshold

RFU	ESR choices	All stutters
80	0	9
70	30	51
60	250	1,660
50	15,500,000	69,200
40	0	0
30	0	0

STRmix calculations done independently

Choosing data at D8



STRmix dropout impact

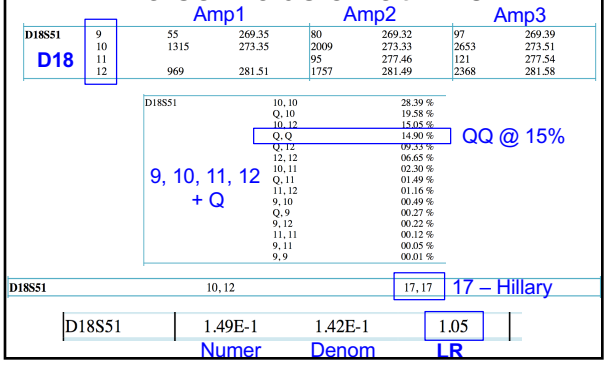
Getting nowhere fast

Alleles (15):
10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

Allele pairs (100):
10,10 10,11 10,12 … 10,25
11,11 11,12 ... 11,25
…
25,25

Too much work and time for the computer:
look at just a few allele data peaks
and merge the rest into "Q"
VIOLATES BAYES RULE

False inclusion at D18



Software comparison

CPI, FST, STRmix, TrueAllele

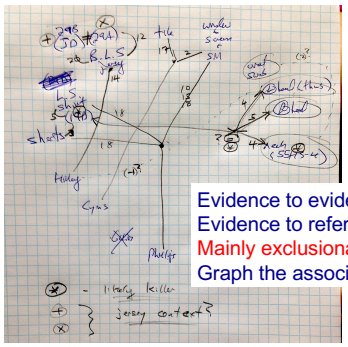
Data
Model
Match
Human
Process
System

Looked at over
50 attributes

Human involvement

Feature	Description	CPI	FST	STRmix	TrueAllele
bias	Cognitive bias?	yes	yes	yes	no
bias	Contextual bias?	yes	yes	yes	no
prepare	People prepare input data?	yes	yes	yes	no
filtering	People can discard input data?	yes	yes	yes	no
intervention	People can adjust input data?	yes	yes	yes	no
subjective	Human decision making?	yes	yes	yes	no
limits	Follow human limitations?	yes	yes	yes	no
time	Labor intensive process?	yes	yes	yes	no
error	Human processing error?	yes	yes	yes	no

March. TrueAllele clusters



Evidence to evidence
Evidence to reference
Mainly exclusionary
Graph the associations

April. NZ report

New software, version 2.4

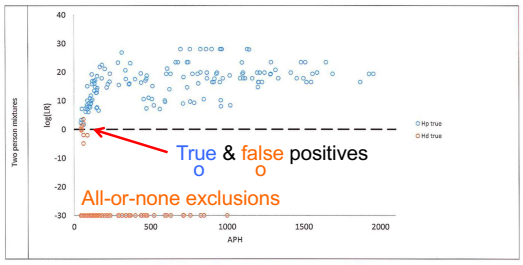
After-stutter modeling
LR = hundred thousand

But …

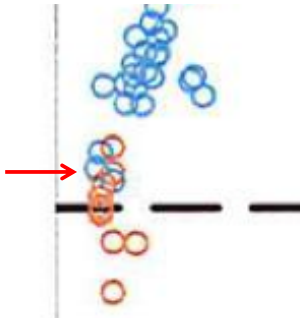
Threshold still at 50 rfu
Higher dropout rate applied

Low-level mixture validation

District of Columbia Validation Study, December 2015



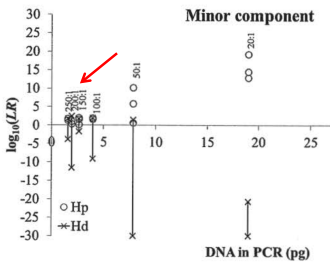
Positive log(LR) inconclusive



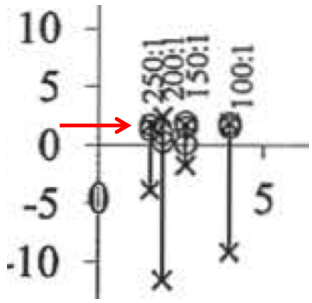
June. NZ validation

True & false positives indistinguishable at low levels

o x



Positive log(LR) inconclusive



July. Frye hearing

Hughes: And subsequent to that, you had no curiosity about what the value would be if you ran it at 30 relative fluorescence units?

Buckleton: No point in running STRmix at 30. We have to go back to GeneMapper and reanalyze the software, the electropherogram, down to 30. And, yes, indeed, I am curious. In fact, I'd like to go to 10.

30 rfu analysis was never reported

Independent defense testing

Access to STRmix™ Software by Defense Legal Teams

- requested by Hillary defense
- agreed to terms & conditions
- software not obtained

The Legal Aid Society (LAS) provided the information presented on this slide, and says it is "100 percent correct." Their testifying expert "agreed" that he was willing to sign and abide by ESR's nondisclosure agreement (NDA), which was under legal review.

ESR says the LAS had not "agreed" because "the Hillary defense did not sign and return the two NDAs required" from the LAS defense experts.

These two views of the word "agreed" are noted on this updated slide. LAS and ESR concur that the NDA was not signed. Eleven days after ESR sent LAS the NDA document, however, software access became a moot issue.

August. STRmix precluded

Dr. Buckleton conceded at the hearing that no internal validation studies were performed by the New York State Police crime lab for the use of STRmix on casework samples developed at the lab.

As a result Dr. Buckleton was forced to **pick and choose** data from different "reliable sources" and input parameters into the program in such a way that he believed the system would tolerate.

ORDERED that the defendant's motion to preclude the prosecution from calling an expert witness to testify on their direct case regarding any conclusion reached by the use of STRmix is granted as the prosecution cannot lay a foundation for the introduction of evidence that had not been internally validated.

Out-of-court NZ statements

- Ran data at 30 rfu, but needed more contributors to make STRmix work and get inclusionary statistic
- Acknowledged STRmix need for thresholds, since it does not model baseline noise (3130 v. 3500)
- Pleased the judge found STRmix reliable, though evidence precluded for other case-specific reasons
- Suggested the main reason to ask for better data (Minifiler) is to support prosecutor's conclusion

September. Hillary acquitted

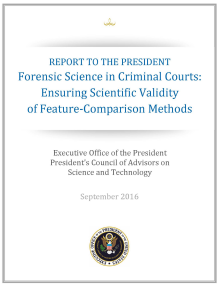
The New York Times

Oral Nicholas Hillary Acquitted in Potsdam Boy's Killing

Future. NZ roadmap

- Demand full discovery
- Test STRmix on case data
- Compare with accurate results
- Identify subjectivity and bias
- Challenge unreliable results
- Cross-examine on deficiencies
- Employ Hillary documents

What PCAST tells us


REPORT TO THE PRESIDENT
Forensic Science in Criminal Courts:
Ensuring Scientific Validity
of Feature-Comparison Methods

Executive Office of the President
President's Council of Advisors on
Science and Technology

September 2016

- 15 years of CPI failure
  never questioned
- 15 more years of nonscience
  time to question
- learn how to challenge
  legal education
- find truth in every case
  TrueAllele screening

Reading materials

TrueAllele Readings

Introductory book chapter
FoleyChapter2013.pdf

Defense perspective
Duffy2012.pdf

Admissibility ruling - TrueAllele admitted
Wakefield2015.pdf

TrueAllele Validation

New York State studies
JFS2011.pdf
JFS2013.pdf
Virginia study
PLoSONE2014.pdf
California study
JFS2015.pdf

Mixture Failure

How the CPI match statistic has failed
MixtureFailure2016.pdf

Why CPI is a random number
JPatholInform2015.pdf

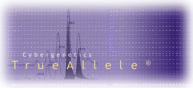
New York v. Nick Hillary

Admissibility ruling - STRmix precluded
HillaryDecision.pdf
HillaryExhibitB.pdf

Standard discovery for STRmix
STRmixDiscovery.docx

Resources

http://www.cybgen.com/information



- Courses
- Newsletters
- Newsroom
- Presentations
- Publications
- Webinars

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

Cybergenetics

perlin@cybgen.com