



Interdisciplinary Symposium - #S01
*To See or Not to See:
 Unbiased Answers to Forensic Questions*



Eliminating Bias in Forensic Algorithms and Statistics



Tuesday, February 16, 2021
 1:35 - 2:30 pm Central Time

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




Cybergenetics © 2003-2021

1

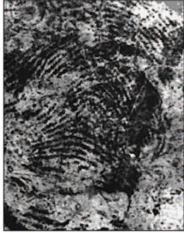
1. Targeted approaches in Forensic Science




Cybergenetics

2

Fingerprint pattern evidence



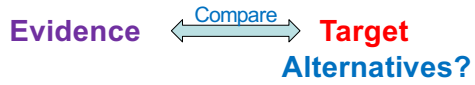
Latent fingerprint
 Madrid bombing



Exemplar fingerprint
 Brandon Mayfield

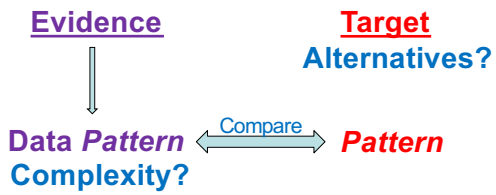
3

Making a comparison: issues



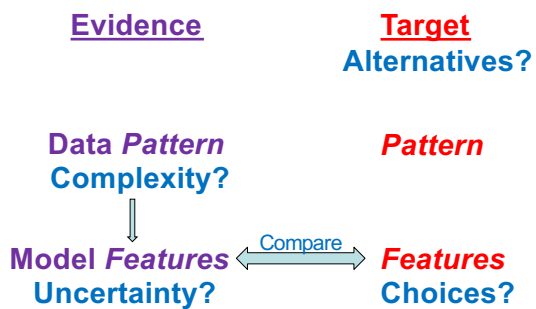
4

Making a comparison: issues



5

Making a comparison: issues



6

Likelihood assesses choices

Probability(Data|Hypothesis)

“Conditional probability of the **data**, given the **hypothesis**”

Data (D) is the evidence pattern

Hypothesis (H) chooses target **t**

$$Pr(D|H)$$

Likelihood measures how well **hypothesis** explains the **data**

A better probability model gives more accurate likelihood

Centering on target **t**, how well does choice explain the evidence?

7



2. Human bias in targeted approaches



Cybergenetics

8

Making choices

- Choosing evidence items
- Choosing data features
- Choosing target person
- Choosing decision criteria

Restricts possibilities & causes error

- Overstate conclusions
- Ignore other answers

Better methods do not limit choices
Instead, they consider all alternatives

9

Confirmation bias

With choices comes the risk of confirmation bias

“unwitting selectivity in the acquisition and use of evidence”

Dr. Raymond Nickerson
Tufts University, Psychology Department
“Confirmation bias: a ubiquitous phenomenon in many guises”
Review of General Psychology, 1998; 2(2):175-220

10

Hypothesis-driven thought

Hypothesis-Determined Information Seeking and Interpretation

- Restriction of attention to a favored hypothesis
- Preferential treatment of evidence supporting existing beliefs
- Looking only or primarily for positive cases
- Overweighting positive confirmatory instances
- Seeing what one is looking for

11

Focus on target: $\Pr(D|H)$

Restriction of attention to a favored hypothesis



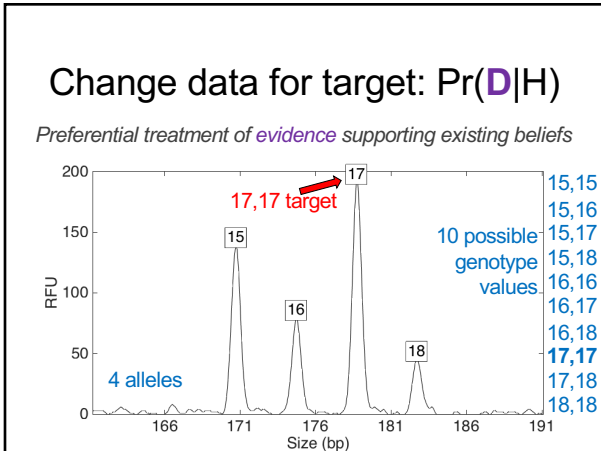
12

Focus on target

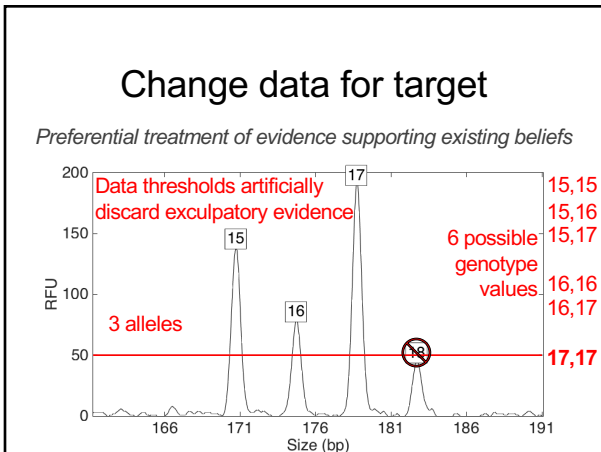
Restriction of attention to a favored hypothesis

What does **non-target** hair look like?

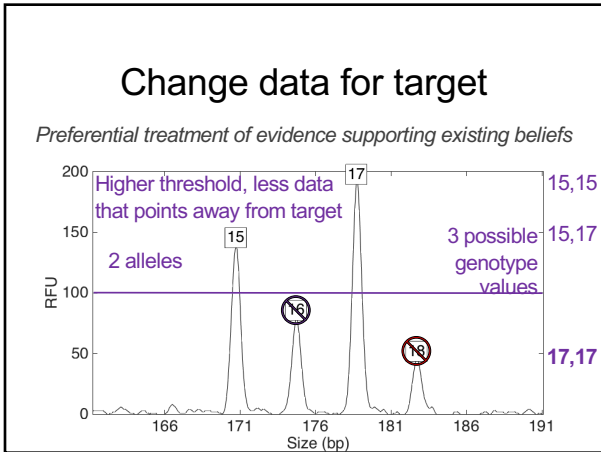
13



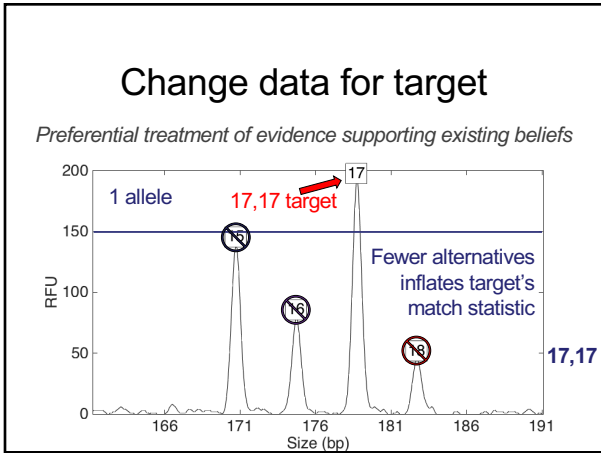
14



15



16



17

Only the positive: $\Pr(D+|H)$

Looking only or primarily for positive cases

Weakly positive?

The image shows a photograph of a P30 test kit on the left and a diagram of the test results on the right. The diagram shows two test strips: 'Negative Result' and 'Positive Result'. Each strip has a 'Control Test' window with 'C' and 'T' labels, and a 'Result' window with 'S' and 'T' labels. In the 'Negative Result' strip, the 'C' window shows a red line and the 'S' window shows a red line. In the 'Positive Result' strip, the 'C' window shows a red line and the 'T' window shows a red line.


First set of tests p30 test for semen

18

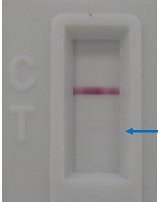
Only the positive

Looking only or primarily for positive cases

Weakly positive?



First set of tests




See the band?
Faintly?
Not at all?

19

Only the positive


Looking only or primarily for positive cases

Weakly positive?



First set of tests
Hospital

Entirely negative

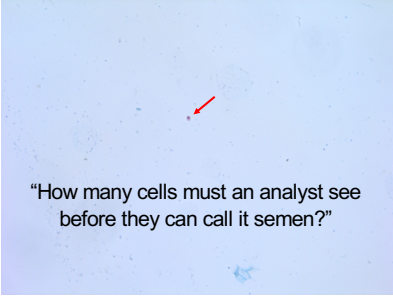


Second set of tests
Autopsy

20

Positive over negative: $Pr(D|H)$

Overweighting positive confirmatory instances



"How many cells must an analyst see before they can call it semen?"

21

Seeing the target: Pr(D|H)

Seeing what one is looking for

DNA from target defendant	DNA from victim's body	Prosecution theory
<p>No</p> <p>No</p> <p>Yes</p> <p>Yes</p>	<p>Face</p> <p>Scrotum</p> <p>Rectum</p> <p>Fingernails</p>	<p>Facial bruising</p> <p>Scrotal bruising</p> <p>Anal-rectal injuries</p> <p>Victim fought back</p>

22

Seeing the target

Seeing what one is looking for

DNA from target defendant	DNA from victim's body	DNA from unknown person
<p>No</p> <p>No</p> <p>Yes</p> <p>Yes</p>	<p>Face</p> <p>Scrotum</p> <p>Rectum</p> <p>Fingernails</p>	<p>Yes</p> <p>Yes</p> <p>Yes</p> <p>Yes</p>

Non-target alternative better explains the theory

23




3. Non-targeted approaches to Forensic Science



Cybergenetics

24

Consider all alternatives

Probability(Data|Alternative)

"Probability of the data given alternative hypothesis"

Alternative hypothesis (~H)

What if it's someone else?

$$Pr(D|\sim H)$$

Considering the non-targets,
how well do they explain the evidence?

25

Likelihood ratio

$$LR = \frac{Pr(D|H)}{Pr(D|\sim H)}$$



Evaluate the evidence data under both target hypothesis H and non-targeted complement alternative hypothesis ~H.

H & ~H are mutually exclusive and exhaustive

"an observation gives one little evidence about the probability of the truth of a hypothesis unless the probability of that observation, given that the hypothesis is true, is either substantially larger or substantially smaller than the probability of that observation, given that the hypothesis is false." – Nickerson

26

Likelihood ratio forms

Likelihoods are computed from forensic model variables. A variable V is an uncertain quantity, taking multiple values.

Probability form of the LR (from Bayes Theorem):

Posterior probability q of target t, after examining evidence D.

$$LR = \frac{q(t)}{p(t)} = \frac{Pr(V = t | D)}{Pr(V = t)}$$

Prior probability p of target t, before seeing evidence.

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



27

Non-targeted approach


1. Without looking at the **target t**, find:
 - Prior probability p(v)**, for all values v in V, **before** looking at the **evidence D**.
From population prevalence studies.
 - Posterior probability q(v)**, for all values v in V, **after** analyzing the **evidence D**.
Usually involves statistical computing.
2. Afterwards, plug in the **target value t**:

$$LR(t) = \frac{q(t)}{p(t)}$$

28

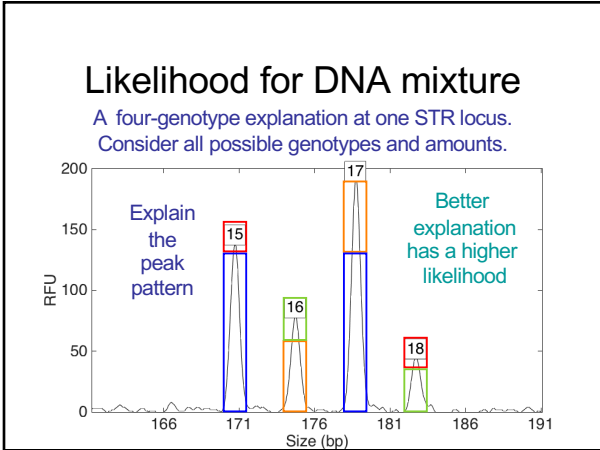



4. Decoupling data analysis and comparison

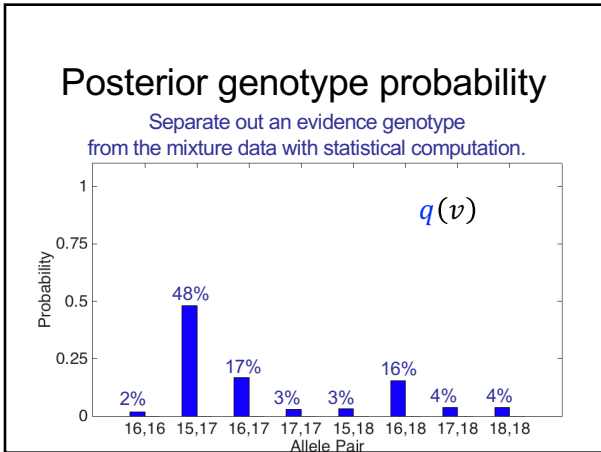


Cybergentics

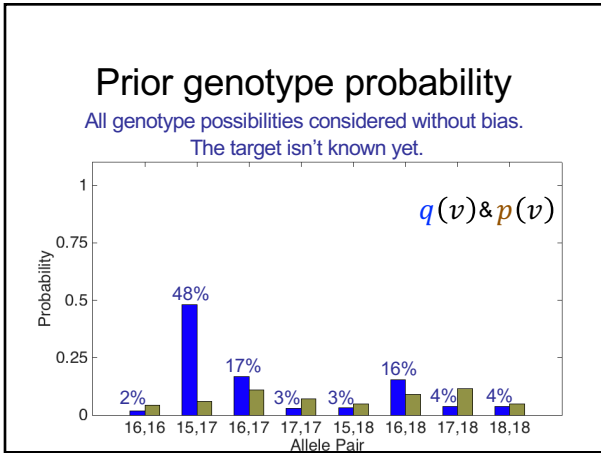
29



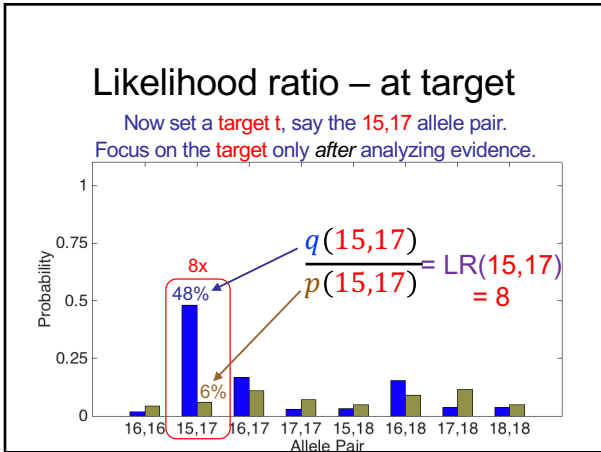
30



31



32



33

Objective LR: two steps



- Objective, unbiased two-step approach
- Split LR calculation into two parts
- Eliminate direct pattern comparison
- Work with data features instead
- Use all the STR genotyping data
- Consider all alternative genotypes

Non-targeted probabilistic approach:
Target is unknown when examining evidence data


Approach can eliminate bias in other forensic subdisciplines

Perlin MW, Kadane JB, Cotton RW.
"Match likelihood ratio for uncertain genotypes."
Law, Probability and Risk. 2009;8(3):289-302.

34

5. DNA case example: targeted vs. non-targeted



Cybergenetics

35

California v. Lopez

Man accused of rape and murder of girlfriend's toddler son

- Facing the death penalty, or life in prison.
- The child was 2 years and 10 months old.
- There were bruises to his face, genitals, and rectum.
- An autopsy showed brain swelling, skull fracture, cheek bruises, and asphyxia.
- A rectal swab from the boy showed semen.
- The swab matched the defendant's DNA.

36

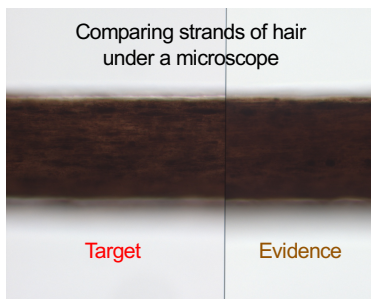
Two different views

Prosecution. The defendant raped and killed a two-year old boy who lived in his house.

Defense. An abused toddler died. The defendant had nothing to do with his death.

37

Defendant's hair on clothes bag?



38

Hair match statistics (DNA,PG)



H1
hair root found on the bag

		Person in	
2 REF Sister	X	21A EC on 17 items	X
3 REF Brother	X	21D SP on 7 items	X
23 Mother	X	7B-4 SP on 8 items	X
24 Defendant	X	8C SP on 8 items	X
25 Brother	X		
35 Victim	X		

39

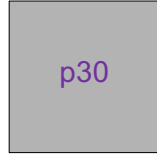
Def's semen in victim's rectum?

Weakly positive?



First in hospital

Entirely negative



Later at autopsy

40

Rectal swabs at hospital (DNA)

16A/B SP

first set of rectal swabs sperm fraction

		Person in	
2 REF Sister	X	21A EC on 17 items	X
3 REF Brother	X	21D SP on 7 items	X
23 Mother	X	7B-4 SP on 8 items	X
24 Defendant	3.81	8C SP on 8 items	X
25 Brother	X		
35 Victim	X		

41

Rectal swabs at autopsy (DNA)

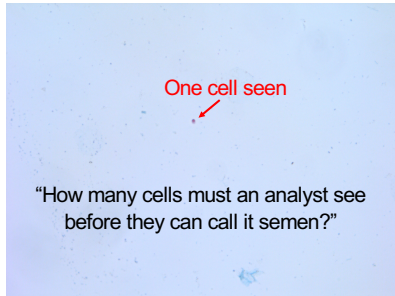
39A/B SP

second set of rectal swabs sperm fraction

		Person in	
2 REF Sister	X	21A EC on 17 items	X
3 REF Brother	X	21D SP on 7 items	X
23 Mother	X	7B-4 SP on 8 items	X
24 Defendant	X	8C SP on 8 items	X
25 Brother	X		
35 Victim	2.79		

42

Def's sperm on victim's penis?



43

Penile swabs (STR)

38A/B SP

second set of penile swabs sperm fraction

		Person in	
2 REF Sister	X	21A EC on 17 items	X
3 REF Brother	X	21D SP on 7 items	X
23 Mother	X	7B-4 SP on 8 items	X
24 Defendant	X	8C SP on 8 items	X
25 Brother	X	Y-STR positive	
35 Victim	7.57		

44

Forensic DNA evidence

Crime Laboratory

STR analyzed 97 evidence items
Reported 43 matches
Discovered 1 unknown person

Cybergenetics

Processed 77 items using TrueAllele®
Reported 138 matches
Discovered 5 unknown people

45

Targeted DNA – manual review

CPI/RMNE just counts how many loci an analyst reported

J Pathol Inform

Editor-in-Chief: **Alan K. Pearce**, Pittsburgh, PA, USA
 Editor-in-Chief: **Lisa Passaniti**, Pittsburgh, PA, USA
 For entire Editorial Board visit: www.pathinformatics.org/edboard.asp

OPEN ACCESS
 HTML format

Research Article

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

Mark William Perlin¹

¹Cybergenetics, Pittsburgh, USA
 E-mail: *Dr. Mark William Perlin - perlin@cygn.com
 *Corresponding author

Received 18 July 2015 Accepted 21 September 2015 Published 28 October 2015

46

JOURNAL OF FORENSIC SCIENCES

Untargeted DNA computer review

Mark W. Perlin,¹ Ph.D., M.D.; Jennifer M. Hornyak,¹ M.S.; Garrett Sugimoto,² M.S.; and Kevin W.P. Miller,² Ph.D.

TrueAllele® Genotype Identification on DNA Mixtures Containing up to Five Unknown Contributors*

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







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

J Forensic Sci. July 2015, Vol. 60, No. 4
 doi: 10.1111/1556-4029.12788
 Available online at: onlinelibrary.wiley.com

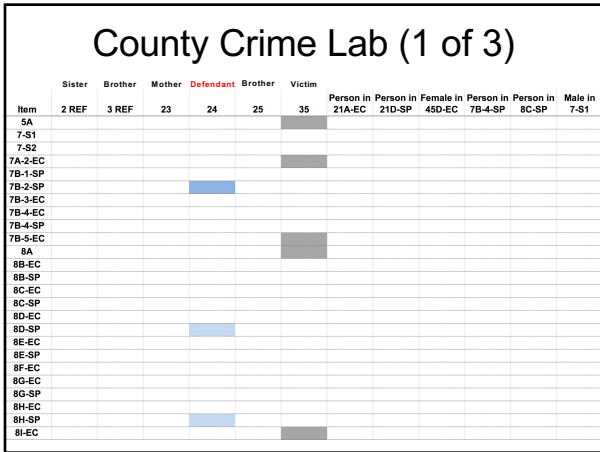
J Forensic Sci. 2019
 doi: 10.1111/1556-4029.14204
 Available online at: onlinelibrary.wiley.com

47

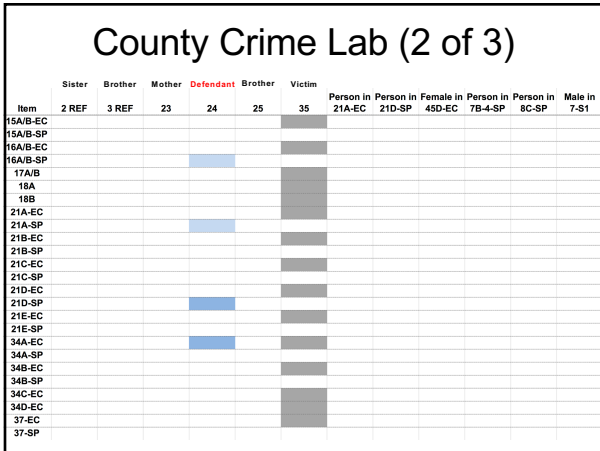
Crime lab vs. TrueAllele information comparison

	Log(LR)	
First 3 charts County Crime Lab Manual review	25 to 30	
	19 to 24	
	13 to 18	
Second 3 charts Cybergenetics TrueAllele PG	7 to 12	
	1 to 6	
	Inclusion	

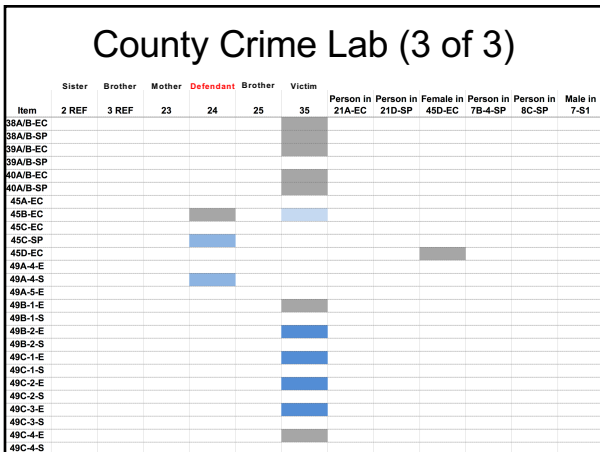
48



49



50



51

The defense sees no criminal

Lots of people's DNA in a messy dirty spermy house



58



Where's Mom's DNA?

Q Did you and the defendant have sex that night?

Mother Just a blow job.

Q Did he ejaculate inside your mouth?

A Yeah.

Q What did you do after that?

A Went and cleaned myself off.

Q What did you clean yourself up with?

A Baby wipes.

Q And then what did you do with the baby wipes afterwards?

A Threw them away.

Q Where did you throw the baby wipes away?

A I don't remember if I threw them in the trash can in my bedroom or if I threw them in the restroom trash.

59



Mother's DNA masked

Her hands were covered in defendant's semen

The child was in toilet training.

The child's primary caregiver had the defendant's semen on her hands.

Swamping her own DNA as she attended to her child.

Which is why we couldn't detect her DNA.

The mother was spreading the defendant's semen:

from her hands, to baby wipes, to garbage bags, and whatever she touched – like to her child.

There was no probative value in this expected DNA.

60



Rectal DNA conflict

Q And in this particular investigation at the hospital, did you collect DNA swabbing from the victim's anus?

Pathologist Yes, sir.

Q And how did you do that?

A The same way we did for his genitals. Swabs were broken from a sterile package, sterile water was applied. I would place the swabs around the skin of anus in a circular manner, insert an inch to two inches, pull them out, and then they would have been handed off to law enforcement.

Q Okay. So you never stuck the swab into the rectum or into the anus more than two inches?

A Right. The beginning part of the anus and rectum is sufficient.

61



Rectal/anal cleaning swab

At the hospital. The mother had transferred the defendant's semen from her hands to the toddler's bottom. The first pathologist swabbed his anus, cleaning the external (anal) semen onto the swab. Before the swab was rectally inserted, the so-called "rectal swab" already contained external (anal) semen.

At the autopsy. The toddler's bottom had been cleaned by the first hospital swabbing. No more external (anal) semen; and there never was any internal (rectal) semen. So the second "rectal swab" was devoid of DNA. No external semen to collect.

Misleading terminology: "**anal/rectal**" swab, not "rectal" swab.

62

Final verdict

The prosecution was target-driven.

The defense was nontarget-driven.

Forensic experts educated the jury.


The nontargeted scenario better explained the evidence.
The jury acquitted the defendant of all charges.

The county no longer seeks the death penalty.

63




6. Impact of non-targeted methods on justice



Cybergenetics

64

Eliminate human bias

Confirmation bias is real, prevalent, and dangerous
 Targeted forensics can introduce confirmation bias
 Likelihood ratios consider both targets and nontargets

The LR can be calculated without targeting suspects
 Split the calculation into two parts:


1. Examine the evidence for probabilities p & q for all possible untargeted values without ever seeing the target
2. After step 1, plug in the target value

65

Unbiased forensic science

Limited DNA analysis targeted Lopez
 The forensics led to his prosecution
 He faced the death penalty
 Better untargeted methods saved him

Lopez is the tip of the forensic iceberg
 Innocent people need better forensics
 Victims of crime need better forensics
 Past, present & future cases need accurate, objective, automated solution



Cybergenetics

66

Thank you

American Academy
Christopher Milroy
Agnes Winokur

Contact information
perlin@cybgen.com
www.cybgen.com

Cybergenetics
William Allan
Jennifer Bracamontes

Public Defender
Kelley Kulick
Michael Ogul



Cybergenetics
