

# DNA Mixture Statistics

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

2013 Spring Institute  
Commonwealth's Attorney's Services Council  
Richmond, Virginia  
March, 2013



Cybergenetics © 2003-2013

---

---

---

---

---

---

---

---

# Child molestation case

- June, 2011: Northern Virginia
  - daughter's birthday slumber party
  - 10 year old girls sleeping in basement
- 
- object sexual penetration
  - aggravated sexual battery (2 counts)
- Prosecutor: CDCA Nicole Wittmann

---

---

---

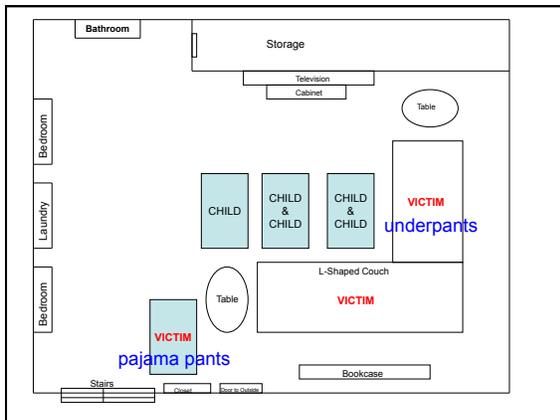
---

---

---

---

---



---

---

---

---

---

---

---

---

## DNA mixture statistics

Human review (using thresholds)

- **underpants**  
original = 10 million  
modified = 1 million
- **pajama pants**  
original = 2 million  
modified = 4

Computer interpretation requested

---

---

---

---

---

---

---

---

## Prosecutor question



What is the true match information of the evidence to the suspect?

---

---

---

---

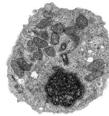
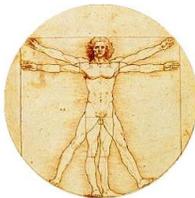
---

---

---

---

## Biology



1 trillion cells

---

---

---

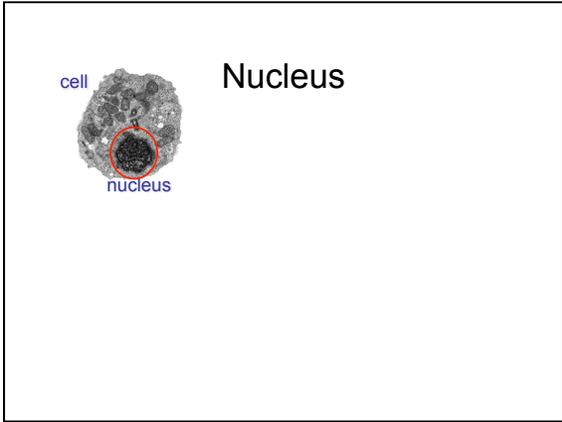
---

---

---

---

---



---

---

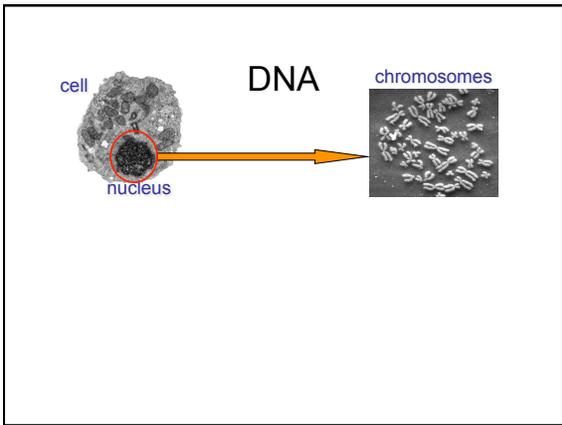
---

---

---

---

---



---

---

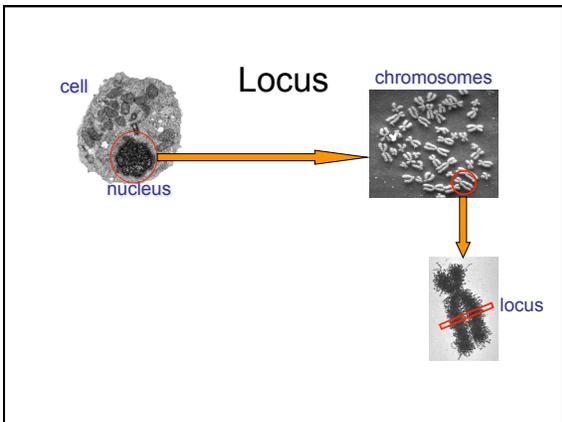
---

---

---

---

---



---

---

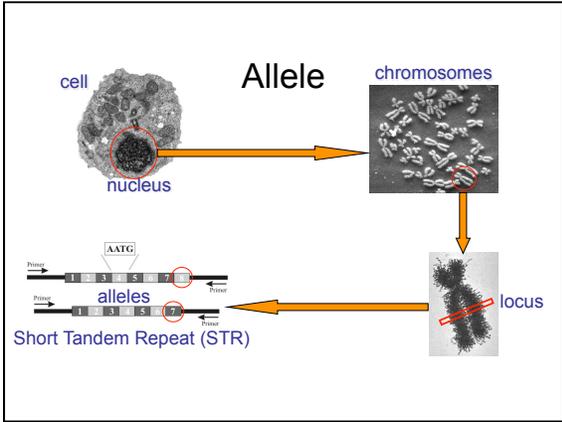
---

---

---

---

---




---

---

---

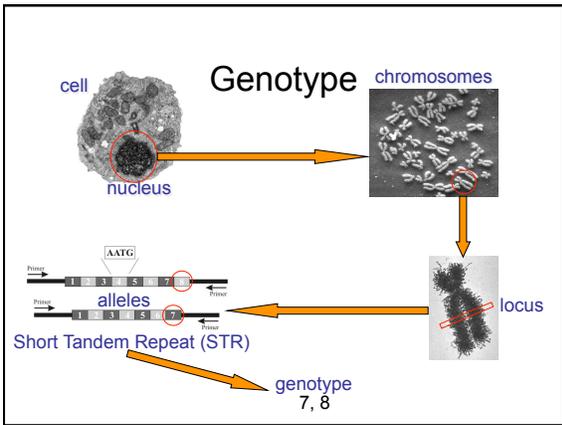
---

---

---

---

---




---

---

---

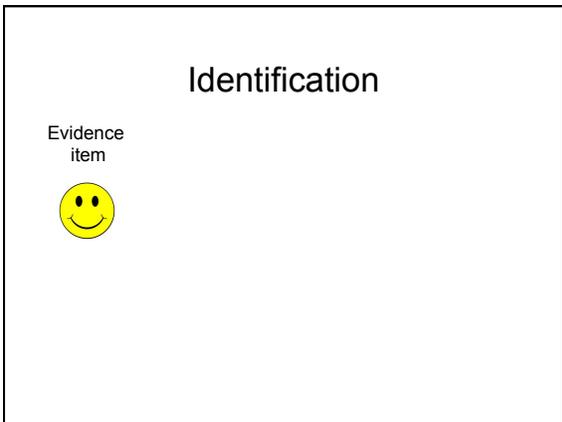
---

---

---

---

---




---

---

---

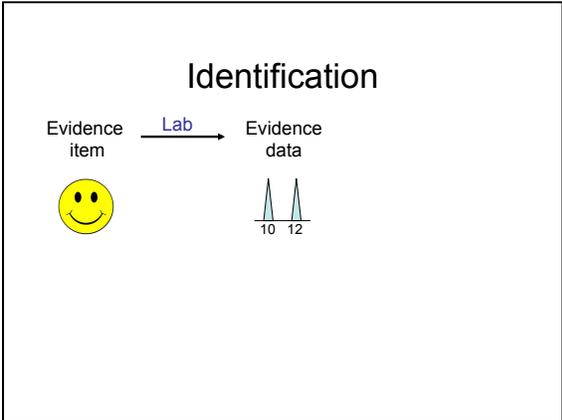
---

---

---

---

---




---

---

---

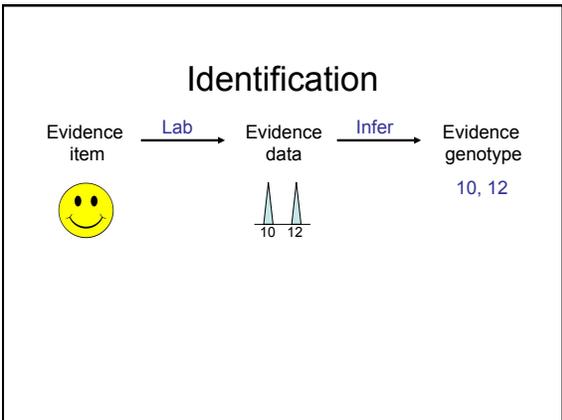
---

---

---

---

---




---

---

---

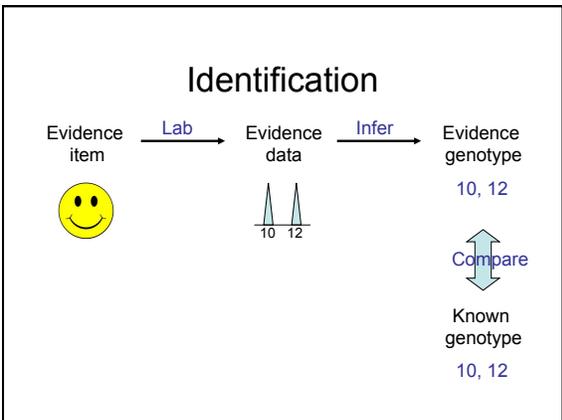
---

---

---

---

---




---

---

---

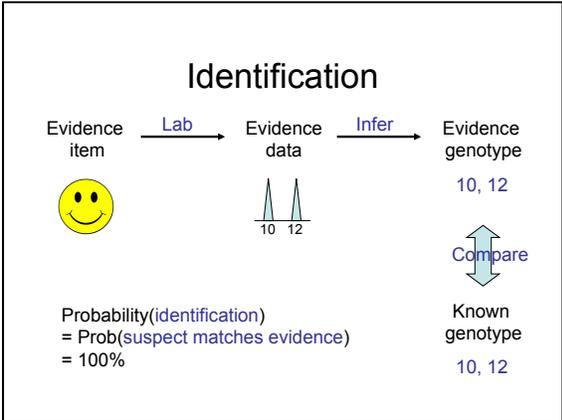
---

---

---

---

---




---

---

---

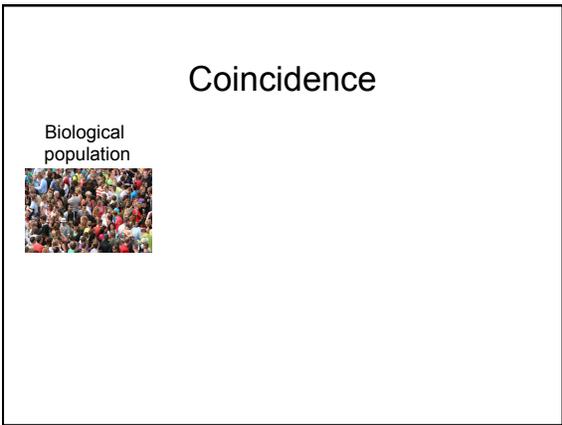
---

---

---

---

---




---

---

---

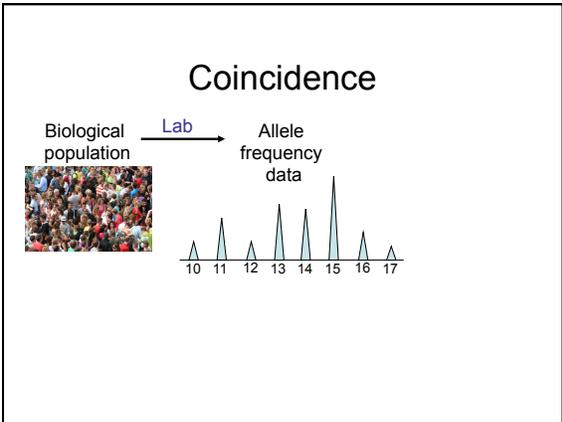
---

---

---

---

---




---

---

---

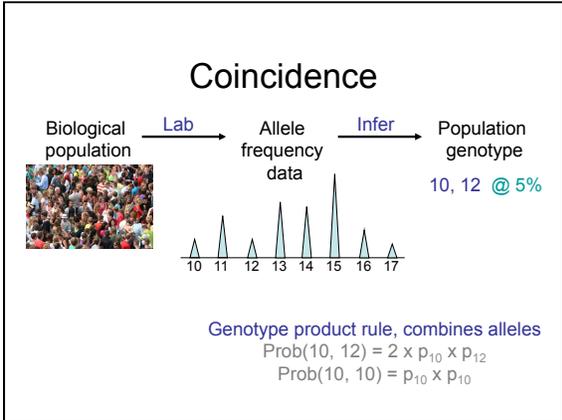
---

---

---

---

---




---

---

---

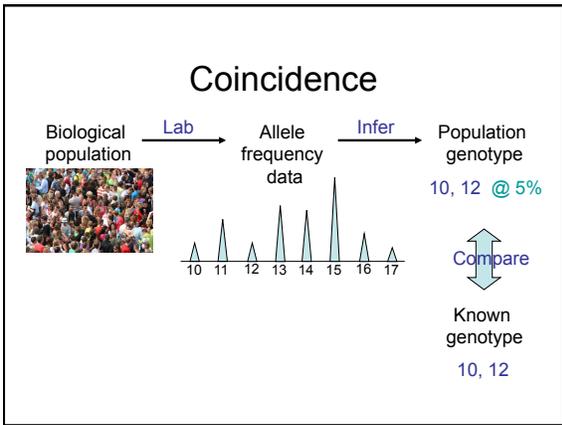
---

---

---

---

---




---

---

---

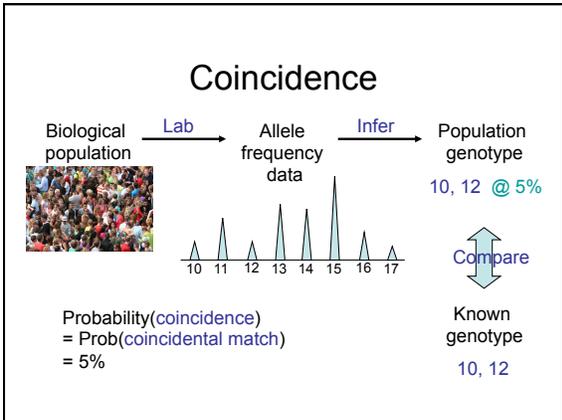
---

---

---

---

---




---

---

---

---

---

---

---

---

### Identification information

At the suspect's genotype,  
identification vs. coincidence?

after  
(evidence)

↑ data

before  
(population)

$$\frac{\text{Prob}(\text{identification})}{\text{Prob}(\text{coincidence})}$$

Evidence changes our belief

---

---

---

---

---

---

---

---

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

### Match statistic

At the suspect's genotype,  
identification vs. coincidence?

after  
(evidence)

↑ data

before  
(population)

$$\frac{\text{Prob}(\text{evidence matches suspect})}{\text{Prob}(\text{coincidental match})}$$

---

---

---

---

---

---

---

---

### Match statistic

At the suspect's genotype,  
identification vs. coincidence?

after  
(evidence)

↑ data

before  
(population)

$$\frac{\text{Prob}(\text{evidence matches suspect})}{\text{Prob}(\text{coincidental match})} = \frac{100\%}{5\%}$$

$$= 20$$

---

---

---

---

---

---

---

---

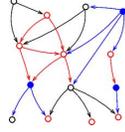
## Bayes theorem



Rev Bayes, 1763

Calculate probability

Belief in hypothesis after having seen data is proportional to how well hypothesis explains the data times our initial belief.



Computers, 1985

All hypotheses must be considered.  
Need computers to do this properly.

Hypothesis: Defendant contributed to DNA evidence

---

---

---

---

---

---

---

---

---

---

---

---

## Mixture interpretation varies

National Institute of Standards and Technology  
Two Contributor Mixture Data, Known Victim

Some Differences in Reporting Statistics

LabID	Kits Used	Cases		
		Caucasians	African-Americans	Hispanics
30	ProPlus/Collier	1.18E+15	2.13E+14	3.09E+15
34	ProPlus/Collier	2.40E+11	7.66E+09	3.93E+10
33	ProPlus/Collier	2.94E+08	1.12E+08	1.74E+09
6	ProPlus/Collier	40,000,000	3,500,000	260,000,000
9	ProPlus/Collier	1.14E+07	1.97E+07	1.54E+08
79	ProPlus/Collier	930,000	47,000	1,350,000
16	ProPlus/Collier	434,620	31,710	399,100

213 trillion (14)

31 thousand (4)

Remember that these labs are interpreting the same MIX05 electropherograms

---

---

---

---

---

---

---

---

---

---

---

---

## DNA mixture

Evidence item




---

---

---

---

---

---

---

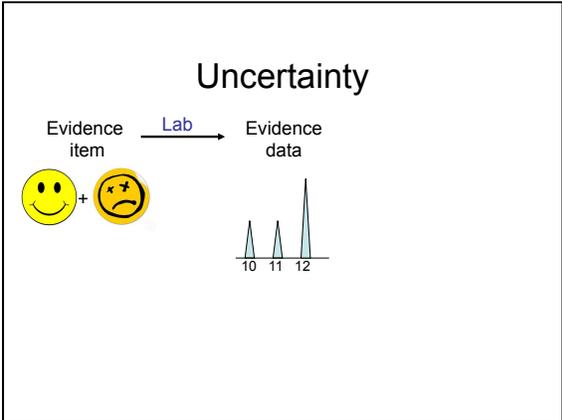
---

---

---

---

---




---

---

---

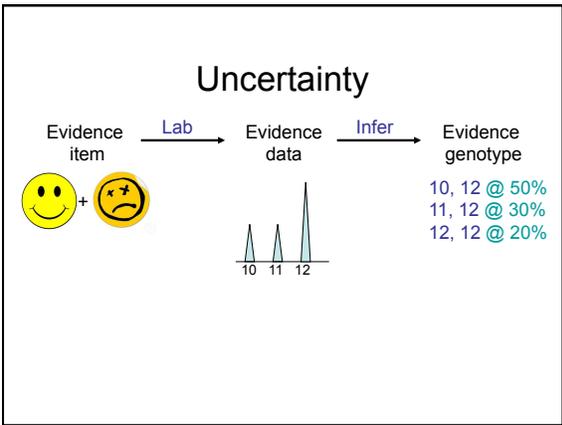
---

---

---

---

---




---

---

---

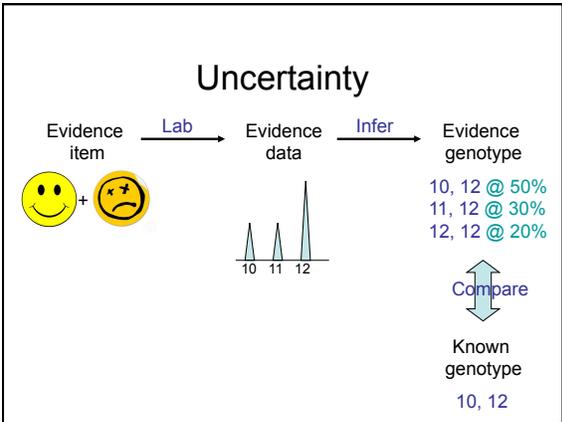
---

---

---

---

---




---

---

---

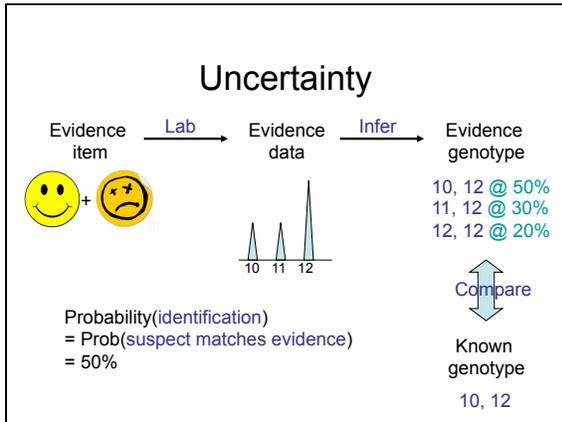
---

---

---

---

---




---

---

---

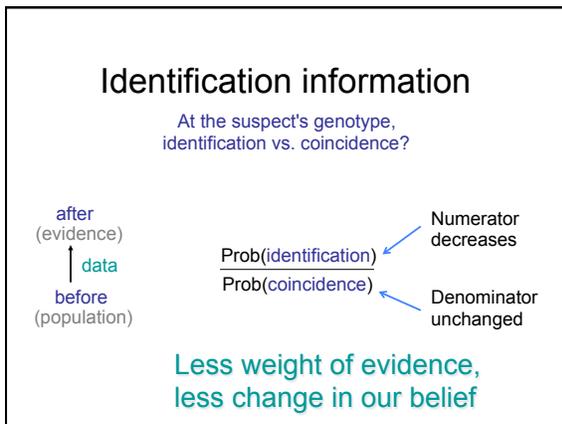
---

---

---

---

---




---

---

---

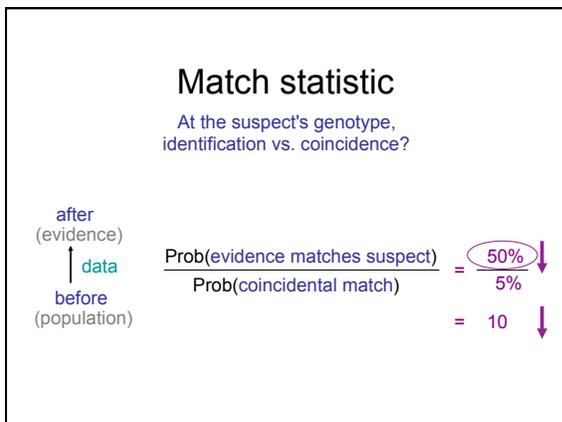
---

---

---

---

---




---

---

---

---

---

---

---

---

## TrueAllele operator

### STR evidence data

.fsa genetic analyzer files

- Replicate computer runs for each item
- 2 or 3 unknown mixture contributors
- Victim genotype was considered

### Evidence genotypes

probability distributions



Cybergenetics

---

---

---

---

---

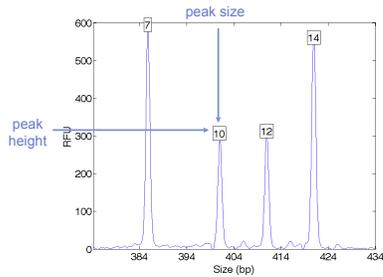
---

---

---

## DNA mixture data

Quantitative peak heights at a locus



---

---

---

---

---

---

---

---

## TrueAllele® Casework



ViewStation  
User Client

Visual User Interface  
VUIer™ Software



Database  
Server

Parallel Processing Computers



Interpret/Match  
Expansion

---

---

---

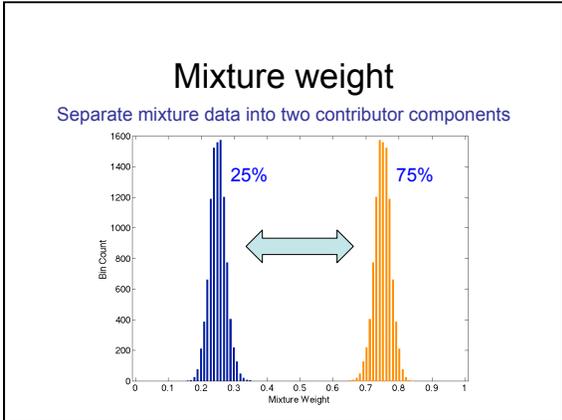
---

---

---

---

---




---

---

---

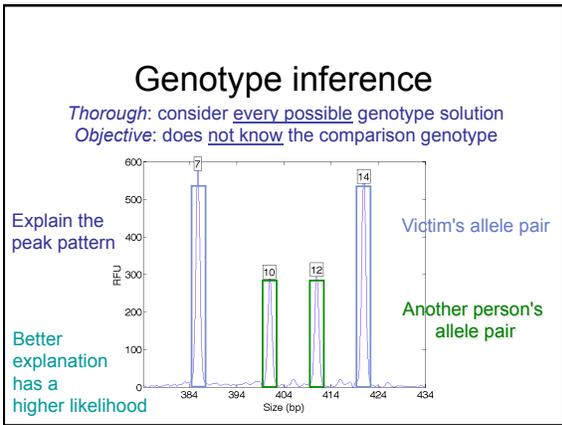
---

---

---

---

---




---

---

---

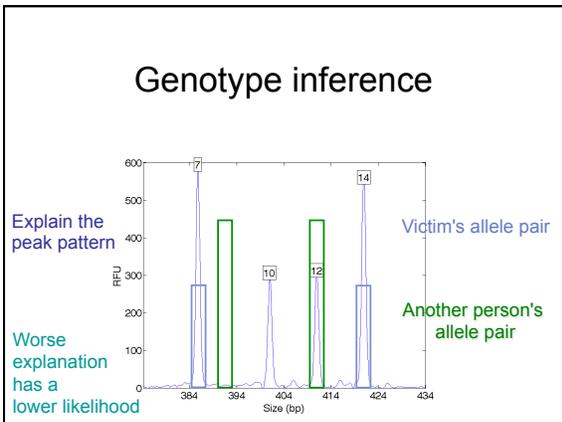
---

---

---

---

---




---

---

---

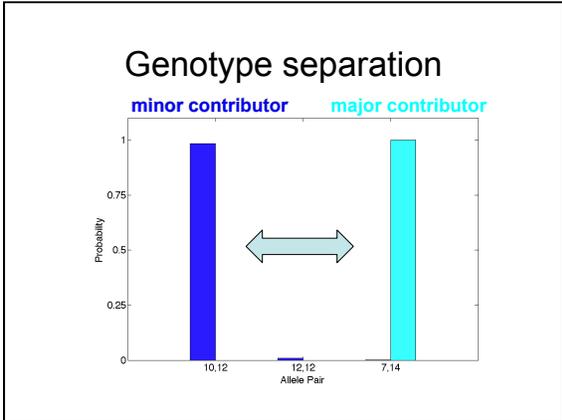
---

---

---

---

---




---

---

---

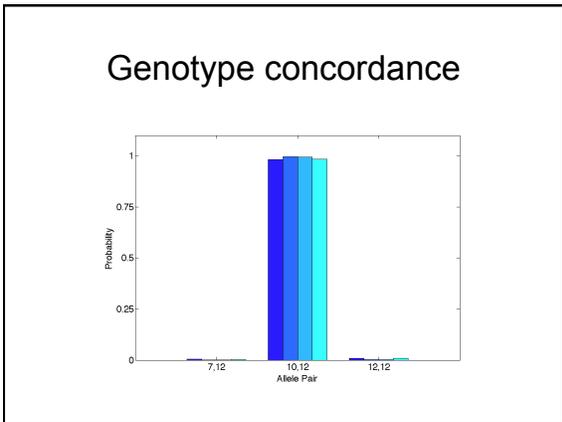
---

---

---

---

---




---

---

---

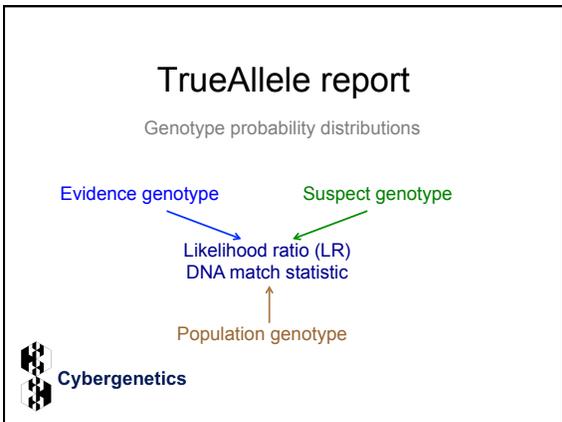
---

---

---

---

---




---

---

---

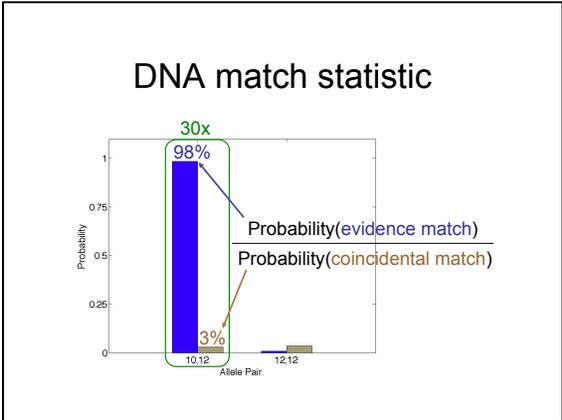
---

---

---

---

---




---

---

---

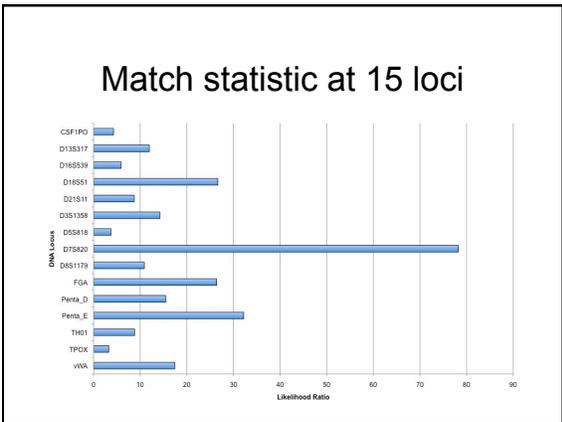
---

---

---

---

---




---

---

---

---

---

---

---

---

### TrueAllele DNA match

LR match to Defendant

Underpants		Pajama pants	
Black	36.6 quintillion	Black	319 thousand
Caucasian	20.7 quadrillion	Caucasian	3.86 thousand
Hispanic	212 quadrillion	Hispanic	32.9 thousand

Cybergenetics

---

---

---

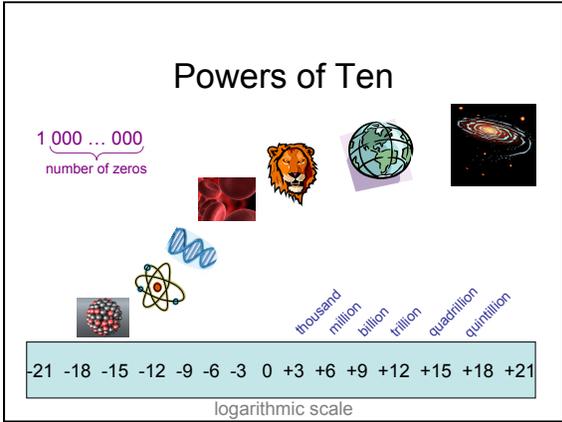
---

---

---

---

---




---

---

---

---

---

---

---

---

### Trial preparation

- discuss case report
- direct examination
- curriculum vitae
- PowerPoint slides
- background reading
- answer questions

---

---

---

---

---

---

---

---

## Computer Interpretation of Quantitative DNA Evidence

Commonwealth v Defendant  
April, 2012  
Arlington, Virginia

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

Cybergenetics © 2003-2012

---

---

---

---

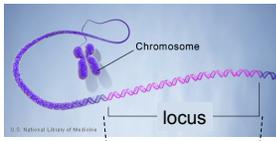
---

---

---

---

## DNA genotype



mother allele

ACGT repeated word

father allele

A genetic locus has two DNA sentences, one from each parent.

An **allele** is the number of repeated words.

A **genotype** at a locus is a pair of alleles.

**8, 9**

Many alleles allow for many many allele pairs. A person's genotype is relatively unique.

---

---

---

---

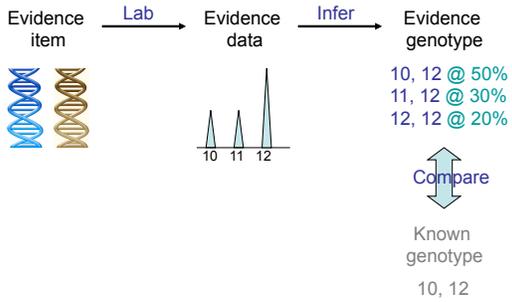
---

---

---

---

## DNA evidence interpretation




---

---

---

---

---

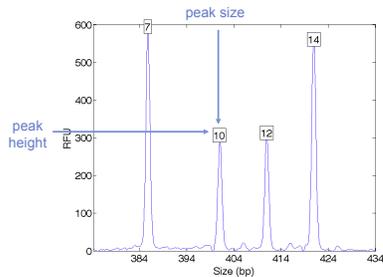
---

---

---

## Computers can use all the data

Quantitative peak heights at locus Penta E




---

---

---

---

---

---

---

---

## How the computer thinks

Consider every possible genotype solution

Explain the peak pattern

Better explanation has a higher likelihood



---

---

---

---

---

---

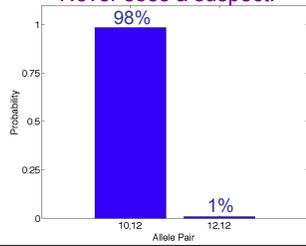
---

---

## Evidence genotype

Objective genotype determined solely from the DNA data.

Never sees a suspect.



---

---

---

---

---

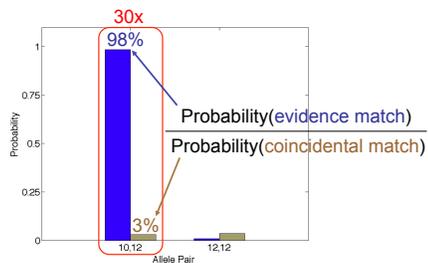
---

---

---

## DNA match information

How much more does the **suspect** match the evidence than a **random person**?



---

---

---

---

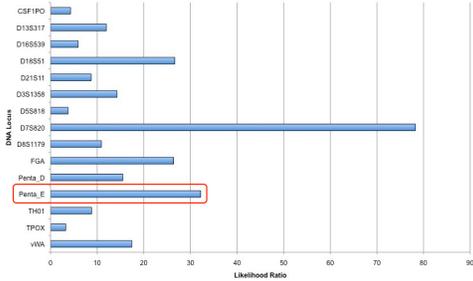
---

---

---

---

## Match information at 15 loci



---

---

---

---

---

---

---

---

## Is the suspect in the evidence?

A match between the underpants and Defendant is:

36.6 quintillion times more probable than a coincidental match to an unrelated Black person

20.7 quadrillion times more probable than a coincidental match to an unrelated Caucasian person

212 quadrillion times more probable than a coincidental match to an unrelated Hispanic person

---

---

---

---

---

---

---

---

## Is the suspect in the evidence?

A match between the pajama pants and Defendant is:

319 thousand times more probable than a coincidental match to an unrelated Black person

3.86 thousand times more probable than a coincidental match to an unrelated Caucasian person

32.9 thousand times more probable than a coincidental match to an unrelated Hispanic person

---

---

---

---

---

---

---

---

## Outcome

- Guilty
- object sexual penetration
  - two counts of aggravated sexual battery

- Sentence
- 22 years imprisonment

- Court of Appeals
- DNA chain of custody
  - appeal denied

---

---

---

---

---

---

---

---

## TrueAllele mixture validation: Virginia case study

Establish the reliability of TrueAllele mixture interpretation

Mark W Perlin, PhD, MD, PhD  
Kiersten Dörner, MS and Jennifer Hornyak, MS  
Cybergenetics, Pittsburgh, PA

Lisa Schiermeier-Wood, MS and Susan Greenspoon, PhD  
Department of Forensic Science, Richmond, VA

---

---

---

---

---

---

---

---

## Case composition

- 72 criminal cases
- 92 evidence items
- 111 genotype comparisons

- Criminal offense
- 18 homicide
  - 12 robbery
  - 6 sexual assault
  - 20 weapon

---

---

---

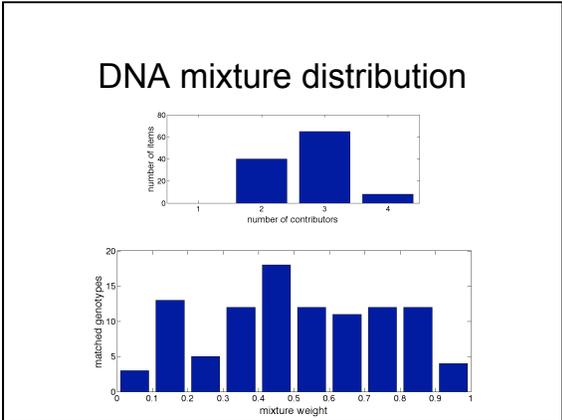
---

---

---

---

---




---

---

---

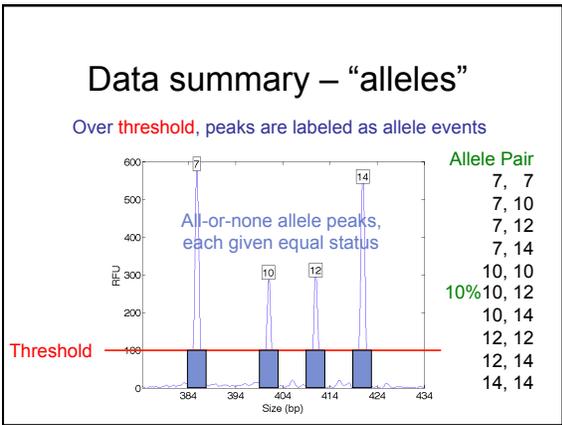
---

---

---

---

---




---

---

---

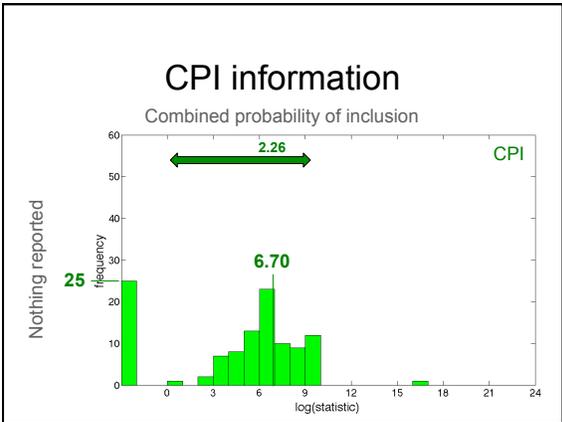
---

---

---

---

---




---

---

---

---

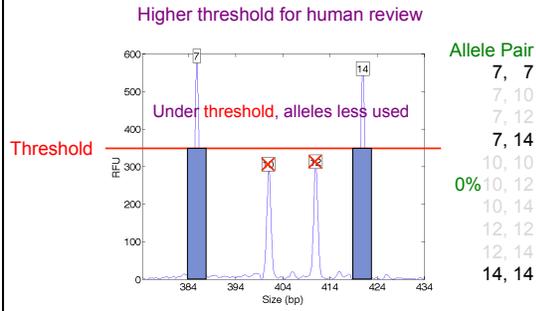
---

---

---

---

## SWGDM 2010 guidelines




---

---

---

---

---

---

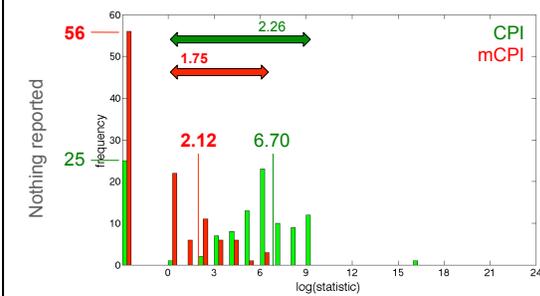
---

---

---

---

## Modified CPI information




---

---

---

---

---

---

---

---

---

---

## SWGDM 2010 guidelines

3.2.2. If a stochastic threshold based on peak height is not used in the evaluation of DNA typing results, the laboratory must establish alternative criteria (e.g., quantitation values or use of a probabilistic genotype approach) for addressing potential stochastic amplification. The criteria must be supported by empirical data and internal validation and must be documented in the standard operating procedures.

Use TrueAllele® Casework for DNA mixture statistics

---

---

---

---

---

---

---

---

---

---

## Validated genotyping method

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

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):1430-47.

Perlin MW, Belrose JL, Duceman BW. [New York State TrueAllele® Casework validation study](#). *Journal of Forensic Sciences*. 2013;58(6):in press.

---

---

---

---

---

---

---

---

## TrueAllele reinterpretation

**The Washington Post**

**Virginia reevaluates DNA evidence in 375 cases**

July 16, 2011

“Mixture cases are their own little nightmare,” says William Vosburgh, director of the D.C. police’s crime lab. “It gets really tricky in a hurry.”

“If you show 10 colleagues a mixture, you will probably end up with 10 different answers”  
Dr. Peter Gill, Human Identification E-Symposium, 2005

---

---

---

---

---

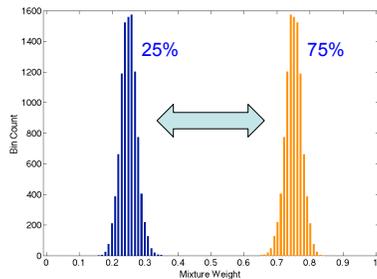
---

---

---

## Mixture weight

Separate mixture data into two contributor components



---

---

---

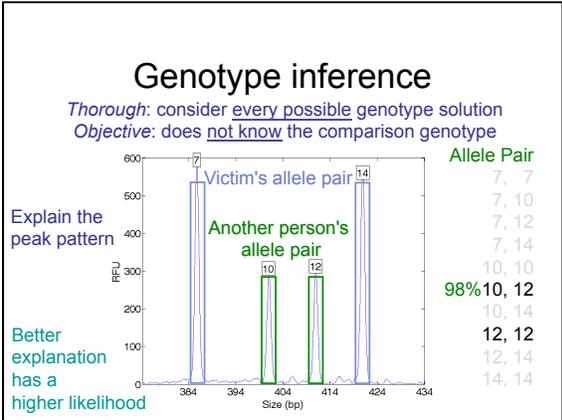
---

---

---

---

---




---

---

---

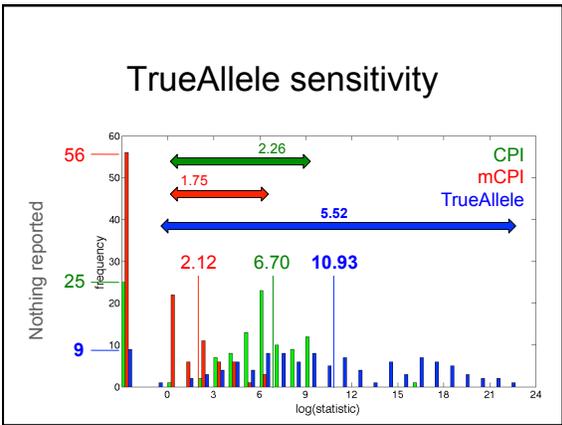
---

---

---

---

---




---

---

---

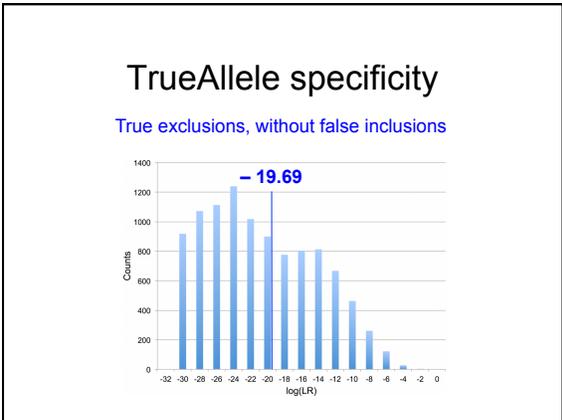
---

---

---

---

---




---

---

---

---

---

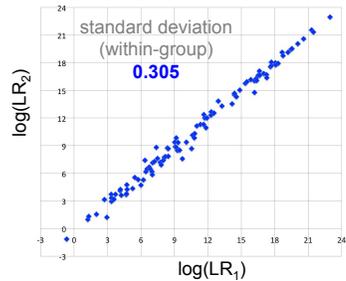
---

---

---

## TrueAllele reproducibility

Concordance in two independent computer runs



---

---

---

---

---

---

---

---

## Validation results

TrueAllele® Casework DNA mixture interpretation is:

A reliable method

- sensitive
- specific
- reproducible

TrueAllele computer genotyping is  
more effective than human review

---

---

---

---

---

---

---

---

## TrueAllele Virginia outcomes

144 cases analyzed  
72 case reports – 10 trials

City	Court	Charge	Sentence
Richmond	Federal	Weapon	50 years
Alexandria	Federal	Bank robbery	90 years
Quantico	Military	Rape	3 years
Chesapeake	State	Robbery	26 years
Arlington	State	Molestation	22 years
Richmond	State	Homicide	35 years
Fairfax	State	Abduction	33 years
Norfolk	State	Homicide	8 years
Charlottesville	State	Homicide	15 years
Hampton	State	Home invasion	5 years

---

---

---

---

---

---

---

---

