

# TrueAllele Speed for Grand Jury Need: Same Day Reporting of Complex Mixtures

Mark W. Perlin, PhD, MD, PhD & William P. Allan, MS, Cybergenetics, Pittsburgh, PA USA

## ABSTRACT

Current DNA mixture interpretation is plagued by a lack of speed and accuracy. Crime laboratory match statistics can vary a billion-fold on an uncomplicated two-person mixture, with many mixtures never reported. Months may pass between generating electronic DNA data and producing a report. Yet there is a clear societal need for more speed and information when reporting mixtures in important cases.

Fast, parallel computing by a validated genotype probability modeling system can overcome this mixture interpretation bottleneck, providing rapid turnaround time and preserving identification information. Cybergenetics TrueAllele® Casework is regularly used to rapidly solve DNA mixtures that have three, four or more unknown contributors. By successively "peeling" away genotype layers, TrueAllele can dissect complex mixtures, often of related individuals. Out of twenty TrueAllele validation studies, seven are published peer-reviewed papers.

At 2 pm on Wednesday, December 11, 2013, a New York State district attorney contacted Cybergenetics. His rape case was going to a Grand Jury the next day. The crime lab had done STR analysis on the inside of a glove. One swab was reported as "consistent with" the victim (major) and two male donors (minors), but excluded the suspect. Another swab contained at least three contributors, but due to "the complexity of the genetic information" no comparisons were made. New DNA results by morning could postpone the case a week.

At 4:30 pm, the lab emailed their .fsa data files to Cybergenetics. By 5 pm, the peak height data were analyzed and uploaded to a TrueAllele server. The computer was asked twelve separate questions: consider both evidence items, in duplicate computer runs, and assume three contributors under different scenarios – all unknown, victim known, and victim and elimination both known. At 5:30 pm, twelve computer cores set to work on their assigned mixture questions. Parallel processing of all twelve 10,000 cycle Markov chains was completed by 11 pm.

Within hours Cybergenetics emailed the DA a preliminary match report on the suspect, a 15% contributor to each three-person mixture. TrueAllele match statistics for the two evidence items were around fifty thousand and five hundred thousand, respectively. A phone call to the DA the next morning explained the results, in time for the Grand Jury.

The following Tuesday Cybergenetics received the lab's case reports, and that day completed a draft TrueAllele report, along with PowerPoint visual aids sent to the DA. On Thursday, December 19, a Cybergenetics scientist appeared in NYS before the Grand Jury to present the DNA match results. The defendant was indicted, arrested, and charged with multiple counts of burglary and rape.

Time and information are important for public safety and criminal justice. DNA delayed is DNA denied. When informative evidence is wrongly deemed "inconclusive" or exclusionary, violent criminals can remain at large. Rapid and reliable parallel computer processing can meet society's demands for timely DNA information. TrueAllele routinely provides this capability.

## TIMELINE

December 11, 2013 (Wednesday)

- 2 pm. Cybergenetics contacted by New York State Chemung County District Attorney
- 4:30. Crime lab emails data files to Cybergenetics
- 5:00. Cybergenetics analyzes and uploads data We ask TrueAllele twelve questions
- 5:30. TrueAllele server computers start solving, with twelve processors working concurrently
- 11:00. TrueAllele finishes computing genotypes. We have the computer compare genotypes, and calculate match statistics.

December 12, 2013 (Thursday)

- 2 am. Cybergenetics sends preliminary report
- 11 am. We speak with the District Attorney, before his Grand Jury meets

December 17, 2013 (Tuesday)

- Cybergenetics receives crime lab case reports
- Issues TrueAllele report and PowerPoint aids

December 19, 2013 (Thursday)

- Cybergenetics scientist presents TrueAllele DNA match results to Grand Jury

## COMPUTING

### Requests

Number of contributors: 3  
Markov chain cycles: 10,000

Items: 2 swabs from interior of glove  
Knowns: none, victim, victim + elimination  
Replicates: 2 independent computer runs

2 x 3 x 2 = 12 questions

### Peeling

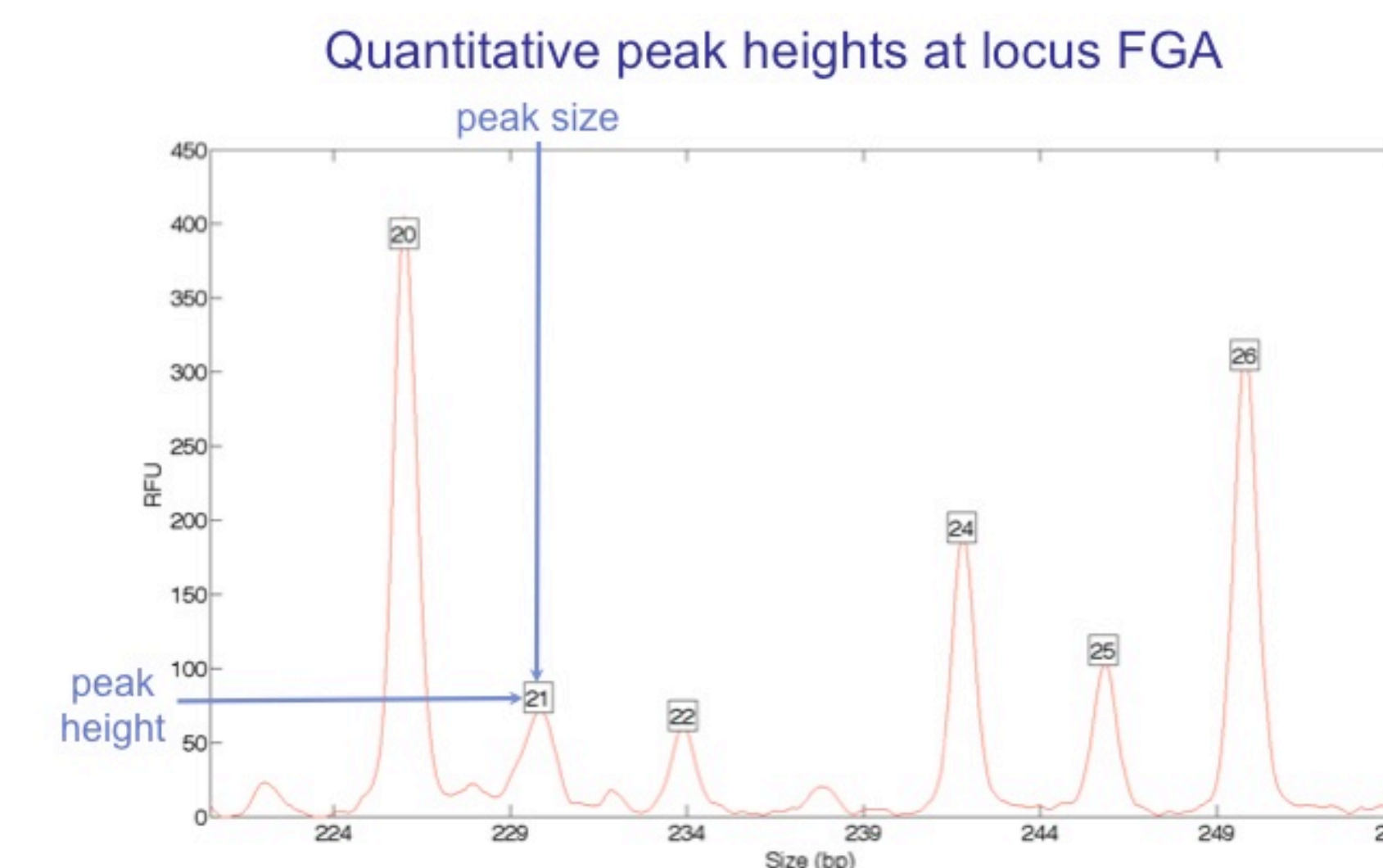
1. Assume nothing, identify major contributor
2. Assume major, identify 1<sup>st</sup> minor contributor
3. Assume major and 1<sup>st</sup> minor, identify 2<sup>nd</sup> minor

### Results

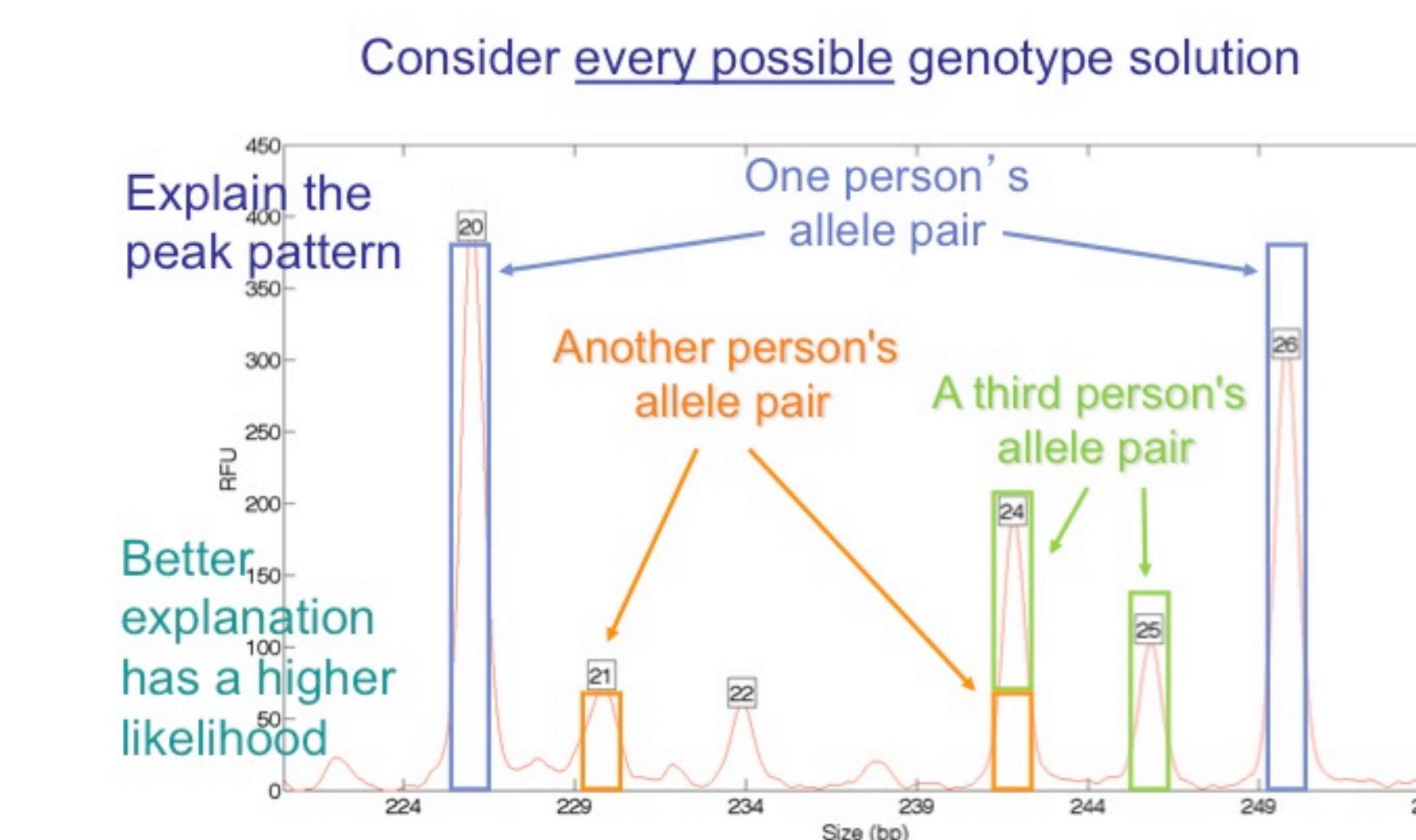
Contributor	Weight	Assumed Knowns	log(LR) match statistics
			Victim Elimination Suspect
victim	80%		17
elimination	5%	Victim	3
suspect	15%	Victim, Elimination	6

## GRAND JURY TESTIMONY

### (1) Computers can use all the data

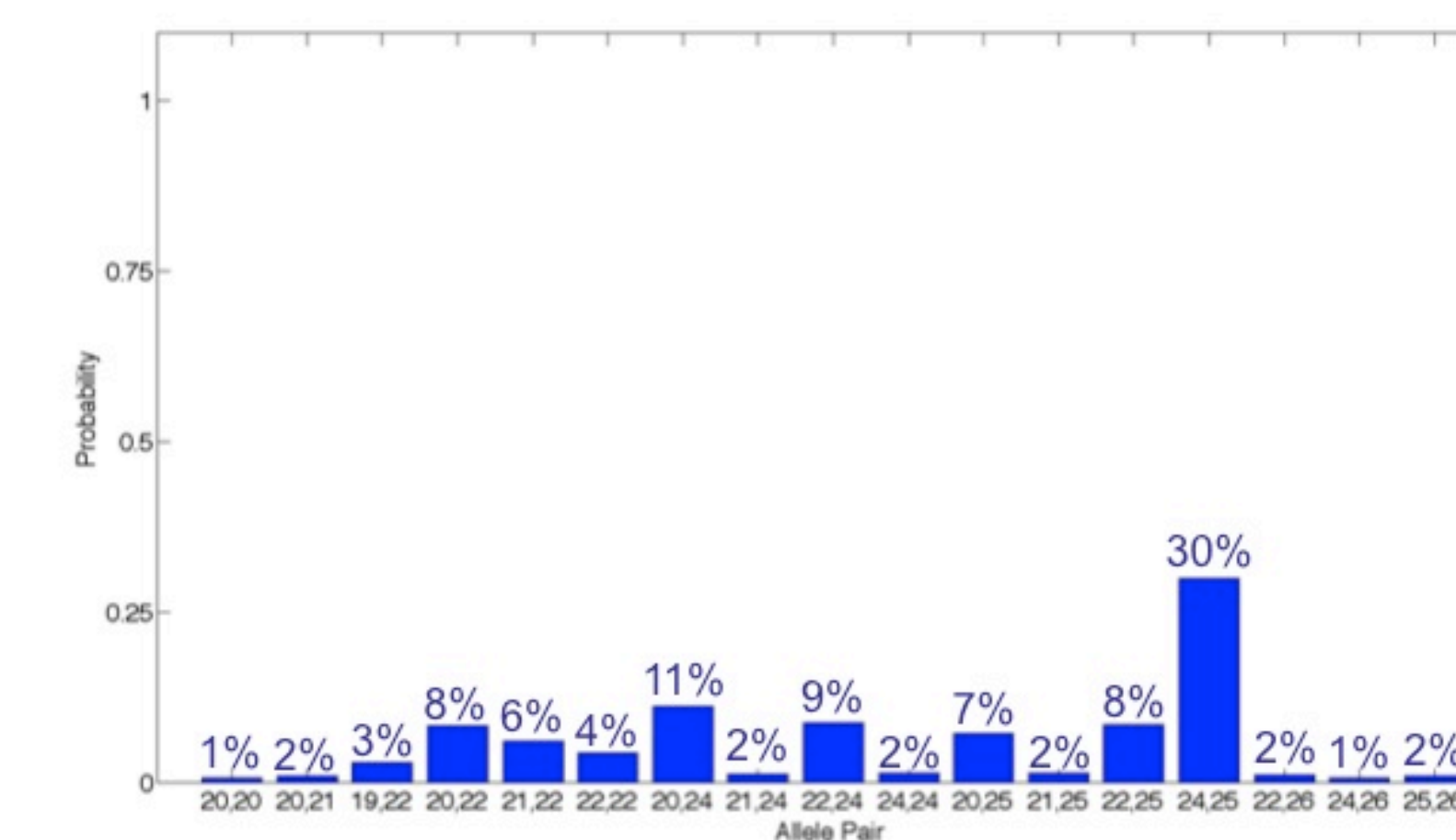


### (2) How the computer thinks



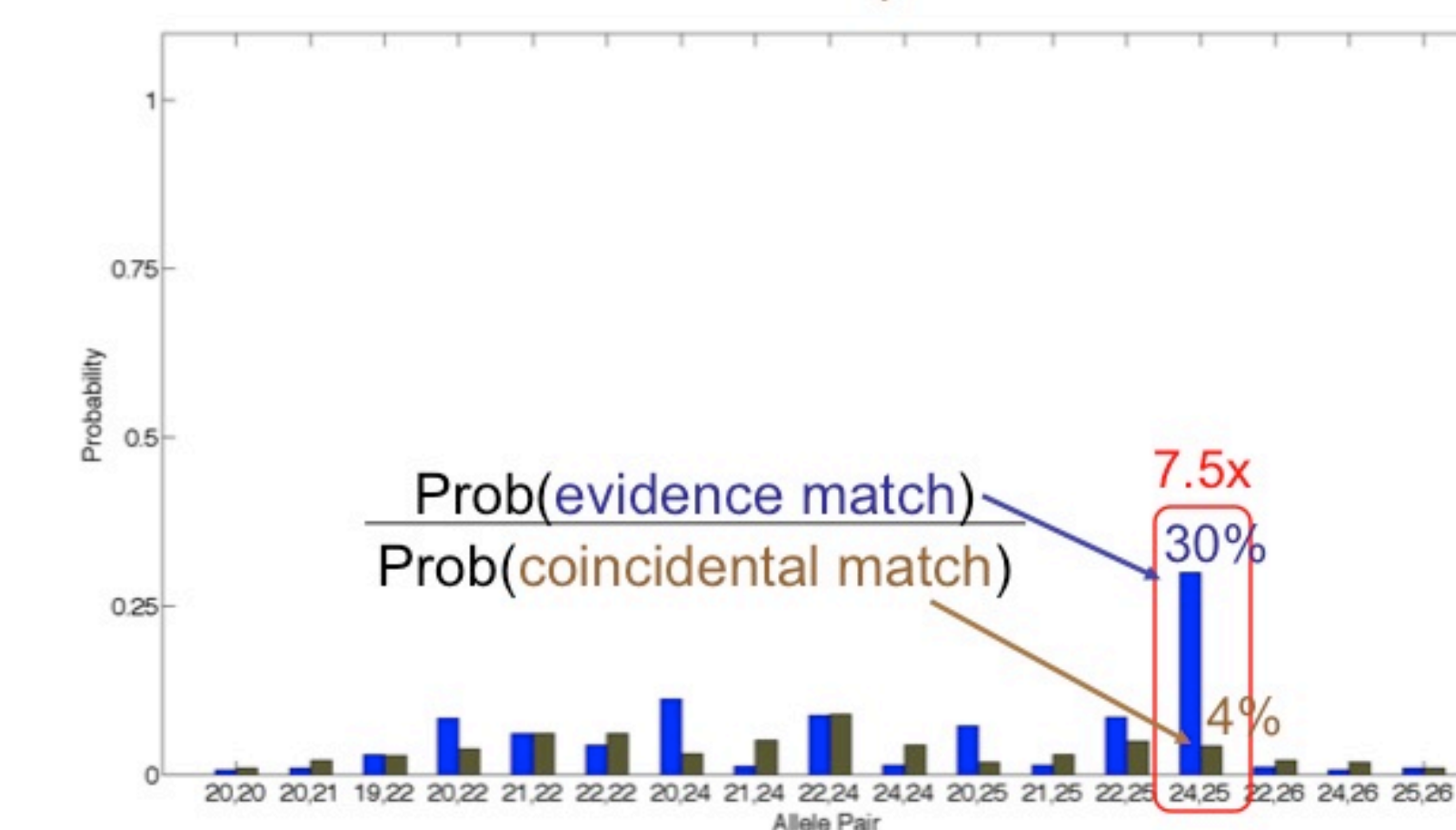
### (3) Evidence genotype

Objective genotype determined solely from the DNA data.  
Never sees a reference.

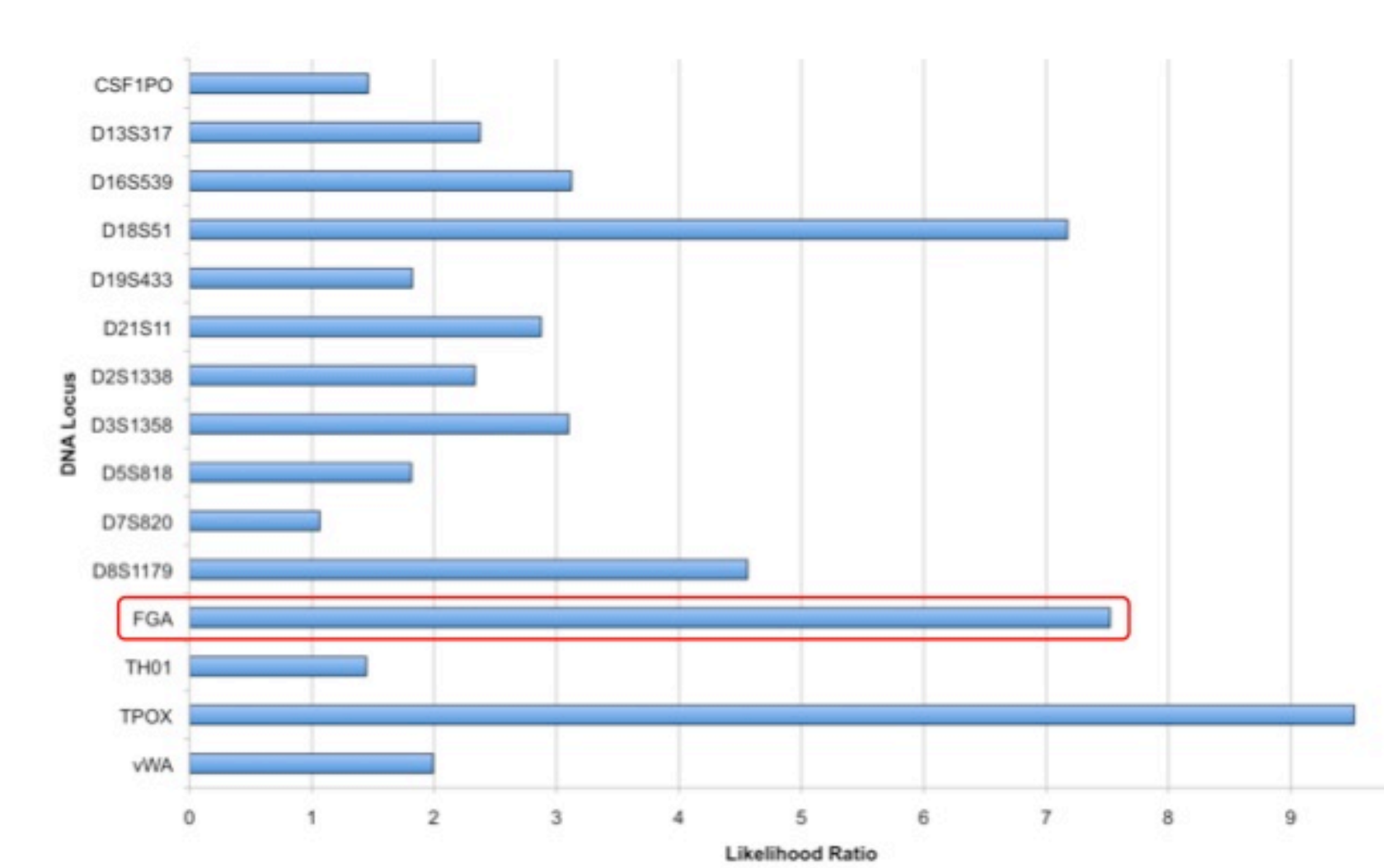


### (4) DNA match information

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



### (5) Match information at 15 loci



### (6) Is the suspect in the evidence?

A match between the glove and Casey Wilson is:

5.37 million times more probable than a coincidental match to an unrelated Black person

10.7 million times more probable than a coincidental match to an unrelated Caucasian person

14.7 million times more probable than a coincidental match to an unrelated Hispanic person

## VALIDATION PAPERS

Samples of known composition

Perlin MW, Sinelnikov 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): 103-14.

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;in press.

Greenspoon SA, Schiermeier-Wood L, Jenkins BC. Establishing the limits of TrueAllele® Casework: a validation study. *Journal of Forensic Sciences*. 2015;in press.

Samples from actual casework

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): 1458-66.

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.

## OUTCOME

Defendant Casey Wilson was tried in September of 2014 in Chemung County court (Elmira, NY). A Cybergenetics scientist testified at the trial on September 11 (Thursday) about the TrueAllele match results. The jury convicted Wilson in less than an hour the following day, finding him guilty of all charges. He faces 40 years in prison.

<http://www.cyngen.com/information/newsroom/2014/sep/Serial-rapist-found-guilty-in-first-New-York-trial-using-TrueAllele.shtml>