

# Getting Past First Bayes with DNA Mixtures

American Academy of Forensic Sciences  
February, 2014  
Seattle, WA

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



Cybergenetics

Cybergenetics © 2003-2014

---

---

---

---

---

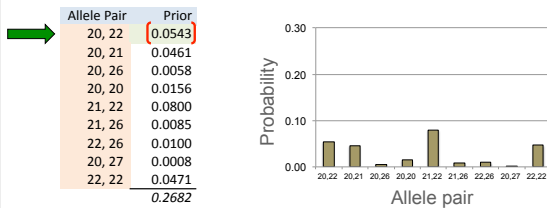
---

---

---

## Prior genotype probability

Genotype of contributor at locus: random variable



---

---

---

---

---

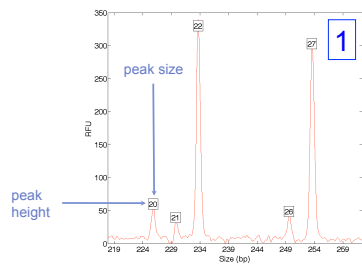
---

---

---

## DNA mixture data

Use all the data



---

---

---

---

---

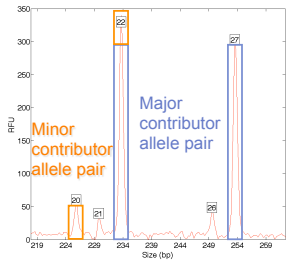
---

---

---

## Genotype likelihood

Higher with good explanation




---

---

---

---

---

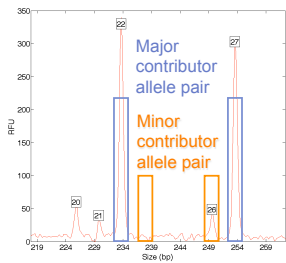
---

---

---

## Genotype likelihood

Lower with poor explanation




---

---

---

---

---

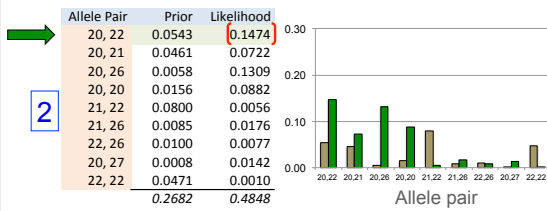
---

---

---

## Genotype likelihood

Consider all genotype hypotheses




---

---

---

---

---

---

---

---

## Posterior probability



The Reverend Thomas Bayes

$$posterior_i = \frac{prior_i \times likelihood_i}{\sum_j prior_j \times likelihood_j}$$

Update beliefs by combining  
 prior belief with likelihood  
 to get posterior belief

Divide by total to sum to 1

---

---

---

---

---

---

---

---

## Posterior genotype probability

**3** Don't confuse likelihood with probability

Allele Pair	Prior	Likelihood	Prior*Like	Posterior
20, 22	0.0543	0.1474	0.008004	0.5636
20, 21	0.0461	0.0722	0.003328	0.2344
20, 26	0.0058	0.1309	0.000759	0.0535
20, 20	0.0156	0.0882	0.001376	0.0969
21, 22	0.0800	0.0056	0.000448	0.0315
21, 26	0.0085	0.0176	0.000150	0.0105
22, 26	0.0100	0.0077	0.000077	0.0054
20, 27	0.0008	0.0142	0.000011	0.0008
22, 22	0.0471	0.0010	0.000047	0.0033
	0.2682	0.4848	0.014200	1.0000

---

---

---

---

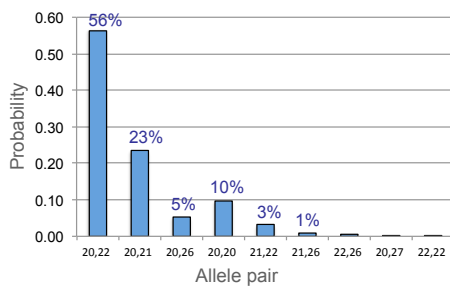
---

---

---

---

## Posterior genotype probability




---

---

---

---

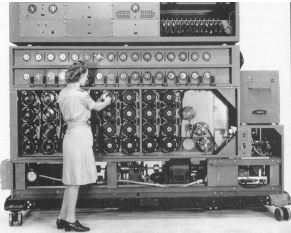
---

---

---

---

## Match information



Dr. Alan Turing  
Decoded the WWII  
Enigma Machine

Likelihood ratio (LR)  
Bayes update for  
hypothesis and its  
alternative

Measures how data  
changes our belief

---

---

---

---

---

---

---

---

## Genotype match information

Use simple ratios, not complex formulas **4**

Allele Pair	Prior	Likelihood	Prior*Like	Posterior	Post/Prior
20, 22	0.0543	0.1474	0.008004	0.5636	10.3800
20, 21	0.0461	0.0722	0.003328	0.2344	5.0843
20, 26	0.0058	0.1309	0.000759	0.0535	9.2180
20, 20	0.0156	0.0882	0.001376	0.0969	6.2111
21, 22	0.0800	0.0056	0.000448	0.0315	0.3944
21, 26	0.0085	0.0176	0.000150	0.0105	1.2394
22, 26	0.0100	0.0077	0.000077	0.0054	0.5422
20, 27	0.0008	0.0142	0.000011	0.0008	1.0000
22, 22	0.0471	0.0010	0.000047	0.0033	0.0704
	0.2682	0.4848	0.014200	1.0000	

---

---

---

---

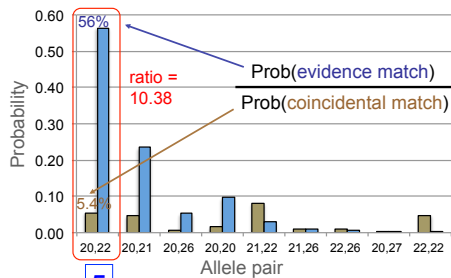
---

---

---

---

## Genotype match information



**5** Make meaningful comparisons

---

---

---

---

---

---

---

---

## NIST teaches Bayes

FGA	Inferred	Prob.	True	PT-True
20,22	0.1474	0.0543	0.008	
20,21	0.0722	0.0461	0.0033	
20,26	0.1309	0.0058	0.0008	
20,20	0.0882	0.0156	0.0014	
21,22	0.0096	0.186	0.0004	
21,26	0.0176	0.0055	0.0001	
22,26	0.0077	0.01	0.0001	
20,27	0.0142	0.0008	0	
22,22	0.001	0.0471	0	
		$\Sigma$ 0.0143		

Statistical Calculation

$H_D$

$$LR = \frac{0.1474}{0.0143}$$

$$LR = 10.33$$

April 2013, NIST Webcast  
DNA Analyst Training on  
Mixture Interpretation

Google: nist trueallele corrections

<http://www.cybgen.com/information/newsroom/2013/apr/Corrections-to-NIST-mixture-interpretation-webcast-misstatements-about-TrueAllele-Casework.shtml>

---

---

---

---

---

---

---

---

---

---

---

---

## Bayes for beginners

1. Use all the data
2. Consider all genotype hypotheses
3. Don't confuse likelihood with probability
4. Use simple ratios, not complex formulas
5. Make meaningful comparisons

---

---

---

---

---

---

---

---

---

---

---

---

## Bringing probabilistic genotyping into the lab & courtroom



- Understand Bayesian update
- Unlearn qualitative peak thresholds
- Learn quantitative genotype modeling
- Train to use the software & certification
- Validate the genotyping system
- Develop standard operating procedures



- Succeed in admissibility hearings
- Write case reports on match results
- Prepare necessary trial materials
- Educate trial attorneys for court
- Explain the LR to jurors and judges
- Answer cross-examination questions

---

---

---

---

---

---

---

---

---

---

---

---

## Genotyping experience

Devise the math & algorithms	20 years
Develop computer systems	15 years
Support users and workflow	10 laboratories
Validate system reliability	20 studies
Educate the community	50 talks
Train & certify analysts	200 students
Go to court for admissibility	5 hearings
Testify about LR results	20 trials
Educate lawyers and laymen	1,000 people
Make the ideas understandable	150 reports

---

---

---

---

---

---

---

---

## Learning about genotyping

<http://www.cybgen.com/information>



- Courses
- Newsletters
- Newsroom
- Presentations
- Publications

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



Cybergenetics



[perlin@cybgen.com](mailto:perlin@cybgen.com)

---

---

---

---

---

---

---

---