Exploring Forensic Scenarios

with TrueAllele® Mixture Automation

59th Annual Meeting
American Academy of Forensic Sciences
February, 2007
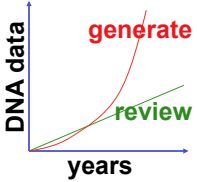
Mark W Perlin, PhD, MD, PhD
Pittsburgh, PA USA

Employee, Shareholder, and Discussion
of Commercial Products or Services

Cybergenetics

So many cases, so little time

DNA data

generate

review

years

meeting the data challenge

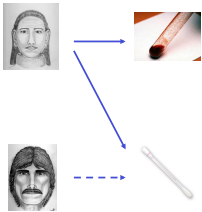---

Do the *best* job possible

- **Look at all the good data**
- **Ignore the bad data**
- **Consider every possibility**
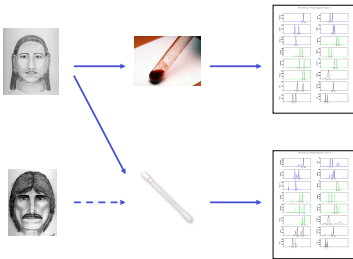- **Obtain the most match information**

DNA evidence

person

DNA evidence

person   specimen

DNA evidence

person   specimen   STR data

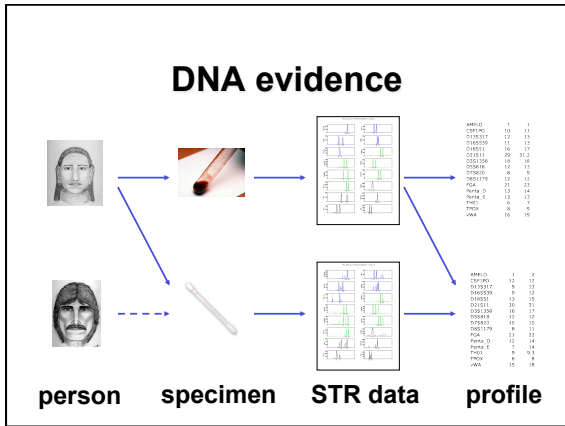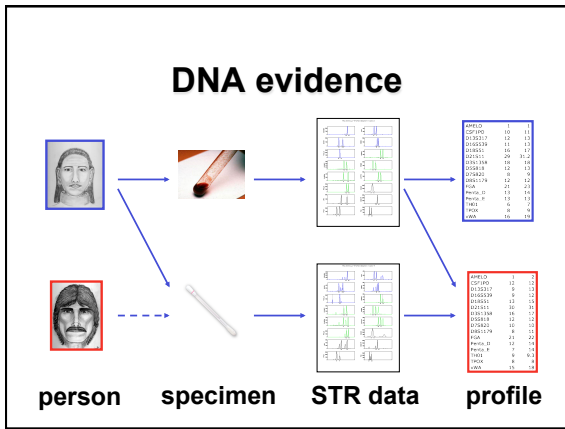**DNA evidence**

person　　specimen　　STR data　　profile

DNA evidence

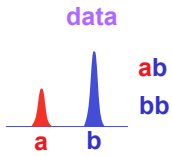person　　specimen　　STR data　　profile

Consider every possibility

- every allele call
- all mixing weights
- stochastic effects
- stutter artifact
- peak imbalance
- specimen combinations

DNA profile ambiguity

data

ab
bb

a   b

Use wild card
[* b]

List possibilities
[a b], [b b]

Attach probability
0.6   0.4

Match DNA profiles

$$\text{Match Information} = \frac{\text{Prob(specific)}}{\text{Prob(random)}}$$

Match Information logarithm –
Very large number:
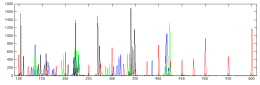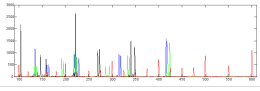"billion billion" or 10^{18}
Use the exponent:
18

Comparing interpretation via match information

	Information
Wildcards	11.9
Listing	15.9
Probability	17.2

Scientific calculator

• interpret DNA
• match DNA
• database engine

Linear Mixture Analysis
J Forensic Sci
2001
46 (6) 1372-1378

"What if ...?" scenarios

• process additional samples?
• which samples should be used?
• is there too much data?
• discard low quality data?
• combine low signal data?
• solve serial crimes?
• how many mixture contributors?

Obtain DNA data

100% |
70% | 30%

Infer DNA profile

Match DNA profile

17.3

How many contributors?

"What if" scenarios

Number	Match Information
1	0.0
2	17.3
3	9.7

Discard bad data
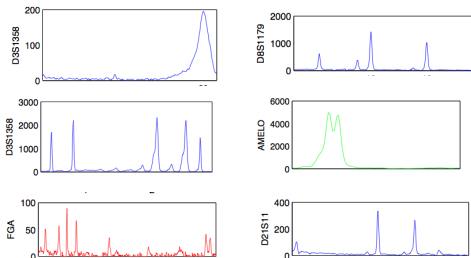
Low-level informative data

1 ng known

125 pg mixture

Using informative data

125 pg
30%
mixture

7.9

Using informative data

125 pg
30%
mixture

7.9

125 pg
50%
mixture

11.4
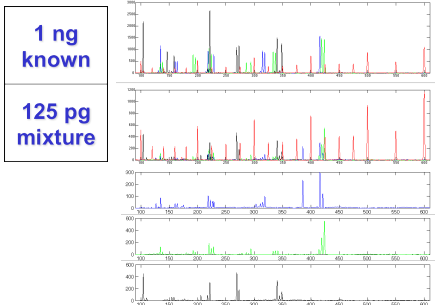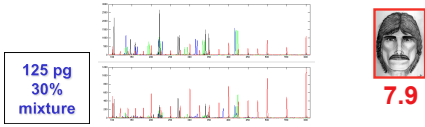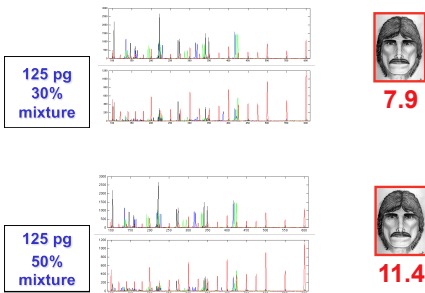
Combine informative data

125 pg
30%
mixture

125 pg
50%
mixture

12.9
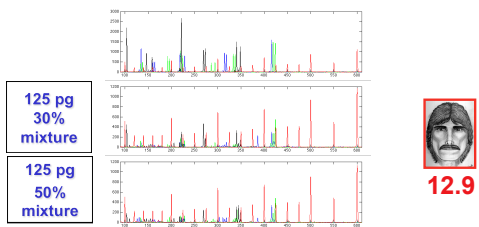
Serial crime investigation

12.9
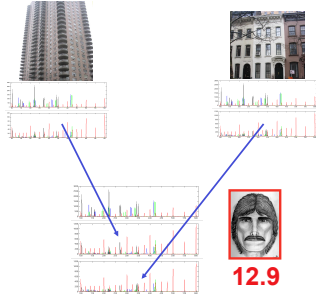
Many sample combinations

Which subset of DNA samples
gives the most informative result?

seven groupings

125 pg 10% mixture	X		X		X	X
125 pg 30% mixture		X	X	X		X
125 pg 50% mixture			X	X	X	X
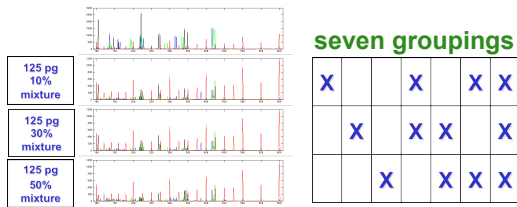
TrueAllele scientific calculator

- 20 interpret CPUs
- 1 match computer
- central database

interpret
match
database
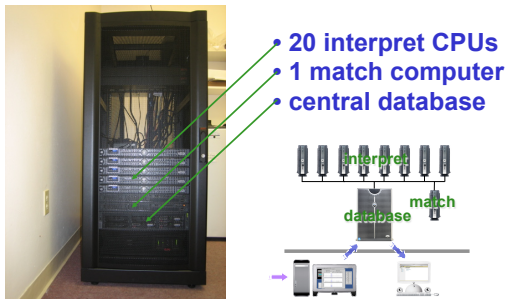
Combining low level data

seven groupings

125 pg 10% mixture	X			X		X	X
125 pg 30% mixture		X		X	X		X
125 pg 50% mixture			X		X	X	X
	0.0	7.9	11.4	7.7	12.9	10.8	12.7

Do the *best* job possible

- Look at all the good data
- Ignore the bad data
- Consider every possibility
- Obtain the most match information