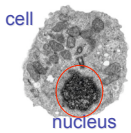
DNA Identification: Biology and Information

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

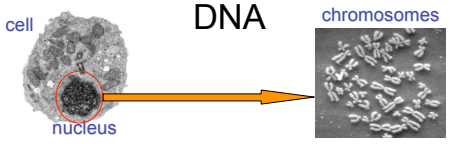
TrueAllele® Lectures
Fall, 2010

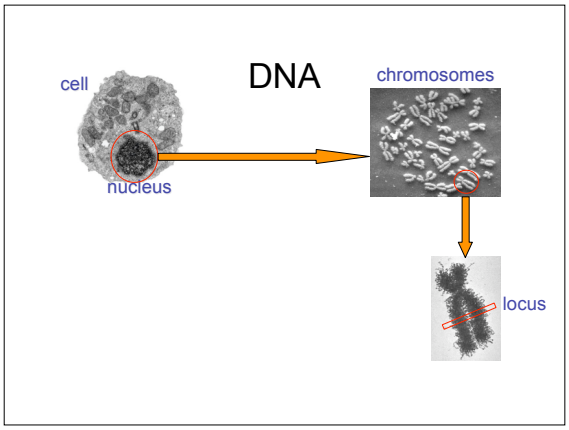
Cybergenetics

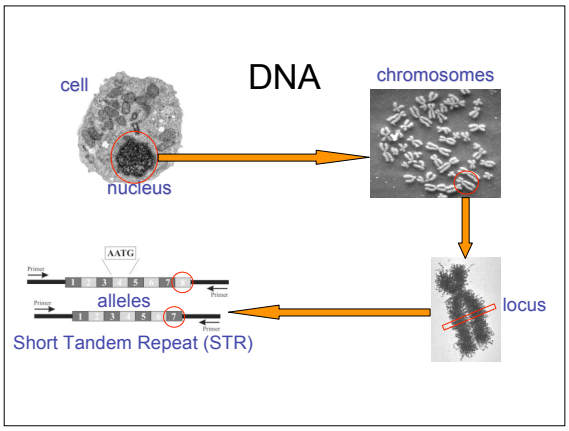
DNA

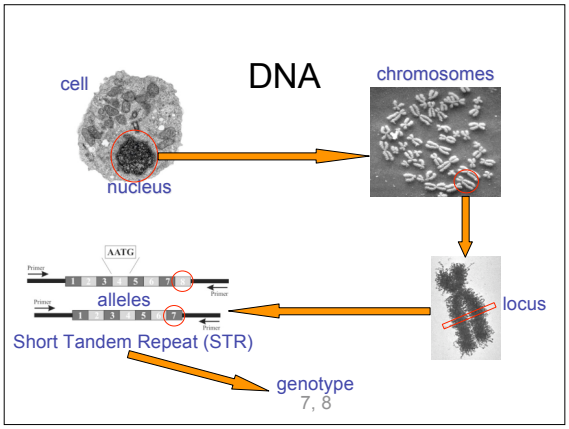


DNA




DNA
cell
nucleus
chromosomes
locus


DNA
cell
nucleus
chromosomes
AATG
alleles
locus
Short Tandem Repeat (STR)


DNA
cell
nucleus
chromosomes
AATG
alleles
locus
Short Tandem Repeat (STR)
genotype
7, 8

Identification

Biological evidence

Identification

Biological evidence → Lab → Questioned data

Identification

Biological evidence → Lab → Questioned data → Infer → Questioned genotype Q

10, 12

Identification

Biological evidence → Lab → Questioned data → Infer → Questioned genotype Q

10, 12

Compare

Suspect genotype S

10, 12

Identification

Biological evidence → Lab → Questioned data → Infer → Questioned genotype Q

10, 12

Compare

Prob(identification) is that of
Prob(suspect matches evidence) = 1

Suspect genotype S

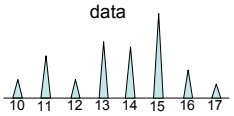
10, 12

Coincidence

Biological population

Coincidence

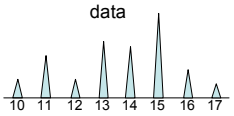
Biological population → Lab → Allele frequency data



Coincidence

Biological population → Lab → Allele frequency data → Infer → Population genotype R

10, 12 @ 5%



Genotype product rule, combines alleles

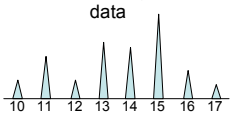
Prob(10, 12) = $2 \times p_{10} \times p_{12}$

Prob(10, 10) = $p_{10} \times p_{10}$

Coincidence

Biological population → Lab → Allele frequency data → Infer → Population genotype R

10, 12 @ 5%



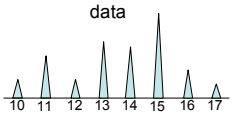
Compare

Suspect genotype S

10, 12

Coincidence

Biological population → Lab → Allele frequency data → Infer → Population genotype R



10, 12 @ 5%

Compare

Prob(coincidence) is that of
Prob(suspect matches population) = 5%

Suspect genotype S

10, 12

Information

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

after (evidence) ↑ data before (population)

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

Likelihood ratio

Information

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

after (evidence) ↑ data before (population)

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

Information

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

after (evidence)
↑ data
before (population)

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

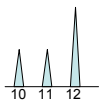
$$= 20$$

Uncertainty

Biological evidence

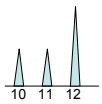
Uncertainty

Biological evidence → Lab → Questioned data



Uncertainty

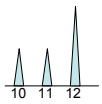
Biological evidence → Lab → Questioned data → Infer → Questioned genotype Q



10, 12 @ 50%
11, 12 @ 30%
12, 12 @ 20%

Uncertainty

Biological evidence → Lab → Questioned data → Infer → Questioned genotype Q



10, 12 @ 50%
11, 12 @ 30%
12, 12 @ 20%

Compare

Suspect genotype S

10, 12

Information

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

after (evidence)
↑ data
before (population)

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

Likelihood ratio

Information

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