Identifying Victim Remains from Uncertain Data

American Academy of Forensic Sciences
February, 2009

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

Cybergenetics

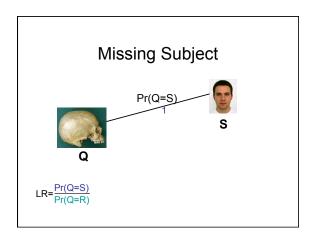Cybergenetics © 2003-2009

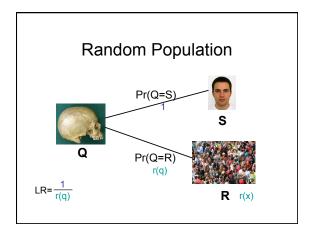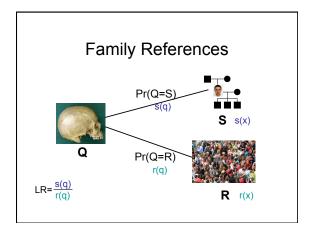
Introduction

- two step kinship approach:
    1. infer genotype from family references
    2. match genotype to human remains
- uncertain genotype has a probability function
- match genotypes using likelihood ratio (LR)
- store and match genotypes on a database
- visual user interface for analyst
- mass disaster applications

Questioned Remains

Q

Missing Subject

$$Pr(Q=S)$$
1

S

Q

$$LR = \frac{Pr(Q=S)}{Pr(Q=R)}$$

Random Population

$$Pr(Q=S)$$
1

S

Q

$$Pr(Q=R)$$
r(q)

$$LR = \frac{1}{r(q)}$$

R r(x)

Family References

$$Pr(Q=S)$$
s(q)

S s(x)

Q

$$Pr(Q=R)$$
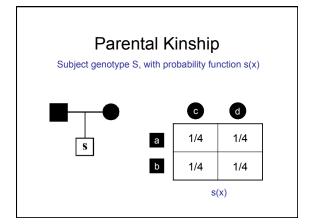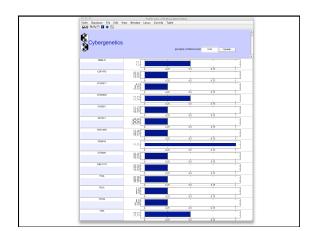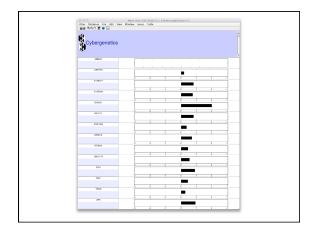r(q)

$$LR = \frac{s(q)}{r(q)}$$

R r(x)

Infer Kinship Genotype
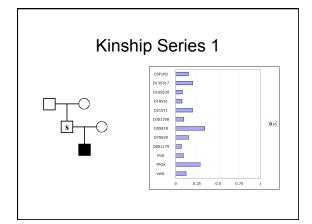
Subject genotype S, with probability function s(x)

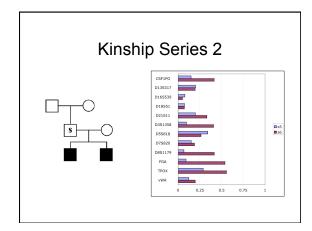$$s(x) \propto prior(x) \cdot likelihood(x)$$
$$= \Pr\{S = x \mid F = y, M = z\} \cdot \prod_{k=1}^{K} \Pr\{C_k = v_k \mid S = x, W = w\}$$
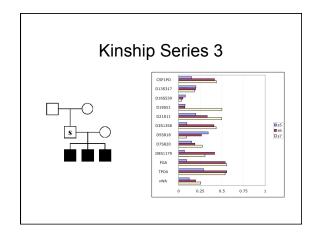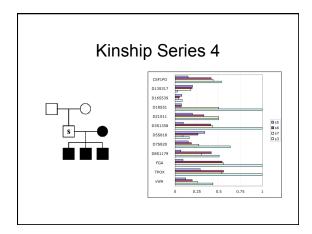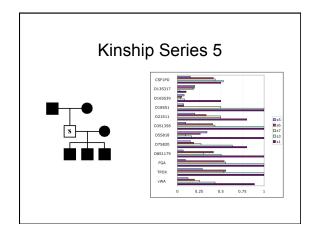
Parental Kinship

Subject genotype S, with probability function s(x)

	c	d
a	1/4	1/4
b	1/4	1/4

s(x)

Kinship Series 1

Kinship Series 2

Kinship Series 3

Kinship Series 4

Kinship Series 5

Match Information

Likelihood Ratio $= \dfrac{Pr(Q=S)}{Pr(Q=R)}$

log(LR)

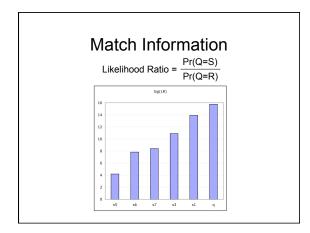
Match Statement

$LR = \dfrac{Pr(Q=S)}{Pr(Q=R)}$

A match between the questioned remains and the missing subject
is N times more probable than
a match between the questioned remains and a random person.

World Trade Center

- Q: Victim remains
  18,251 samples; 79,084 lanes
- S: Missing persons
  personal effects (2,386)
  family references (6,660 for 2,347)
- *Infer* uncertain genotypes (Q, R, S)
- *Match* likelihood ratio: $\dfrac{Pr(Q=S)}{Pr(Q=R)}$
- Computer implementation
  24 processor TrueAllele® supercomputer
  VUIer™ visual user interface software
  analyst asks questions, computer answers
  productivity: set up 30 samples per hour