How TrueAllele® Works (Part 3)
Kinship, Paternity and Missing Persons

Cybergenetics Webinar
December, 2014

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

Cybergenetics

Kinship

Task. Infer genotypes and match them

Tool. Bayes theorem

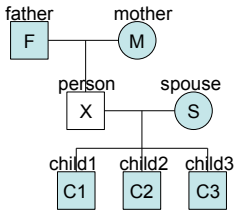
Data. Genotypes of family members

Bayes theorem

$$\text{posterior} \propto \text{likelihood} \times \text{prior}$$

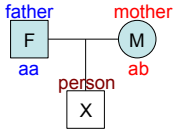
$$\Pr(X = x | data) \propto \Pr(data | X = x) \times \Pr(X = x)$$

Pedigree



Prior probability

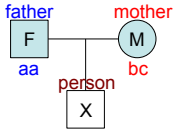
How a variable is affected



$$\Pr\left(X = x \middle| F, M\right)$$

Prior probability

Punnett square



	b	c
a	ab	ac
a	ab	ac

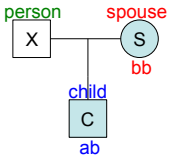
$$\Pr\left(X = x \middle| F = aa, M = bc\right)$$

X = ab 1/2 ac 1/2

Likelihood

How a variable affects others

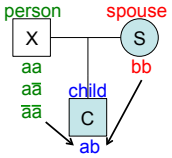
$$\Pr(C \mid X = x, S)$$



Likelihood

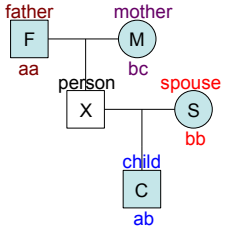
$$\Pr(C = ab \mid X = x, S = bb)$$

$$L = \begin{cases} aa, 1 \\ a\bar{a}, 1/2 \\ \bar{a}\bar{a}, 0 \end{cases}$$



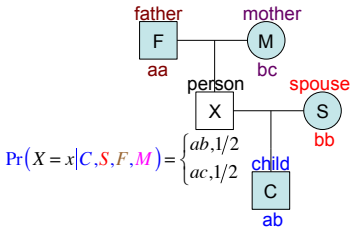
Posterior probability

$$\Pr(X = x \mid C, S, F, M) \propto \Pr(C \mid X = x, S) \times \Pr(X = x \mid F, M)$$



Posterior probability

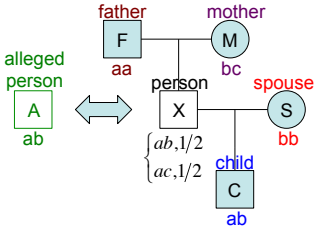
$$\Pr(X = x \mid C, S, F, M) \propto \Pr(C \mid X = x, S) \times \Pr(X = x \mid F, M)$$



$$\Pr(X = x \mid C, S, F, M) = \begin{cases} ab, 1/2 \\ ac, 1/2 \end{cases}$$

$$\Pr(C = ab \mid X = x, S = bb) \times \Pr(X = x \mid F = aa, M = bc)$$

Genotype comparison



Bayes theorem, second round

Hypothesis: The missing and reference persons are the same individual

$$LR = \frac{O(H \mid C, S, F, M)}{O(H)}$$

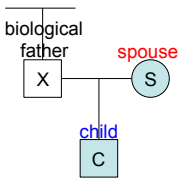
$$= \frac{\Pr(X = ab \mid C, S, F, M)}{\Pr(X = ab)}$$

Posterior to prior genotype probability ratio (Essen-Moller, 1938).

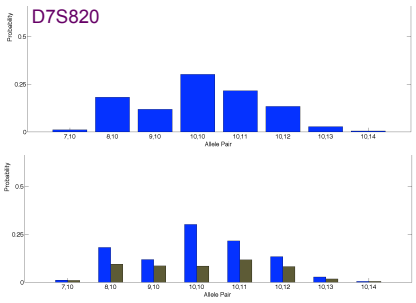
Match strength

$$LR = \frac{\Pr(X = ab \mid C, S, F, M)}{\Pr(X = ab)}$$
$$= \frac{1/2}{2 p_a p_b}$$
$$= \frac{1}{4 p_a p_b}$$

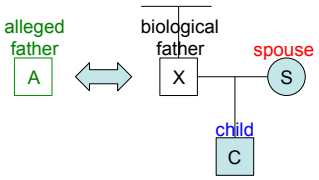
Paternity



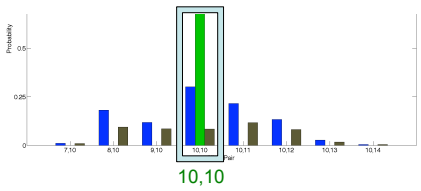
Paternity genotype



Genotype comparison

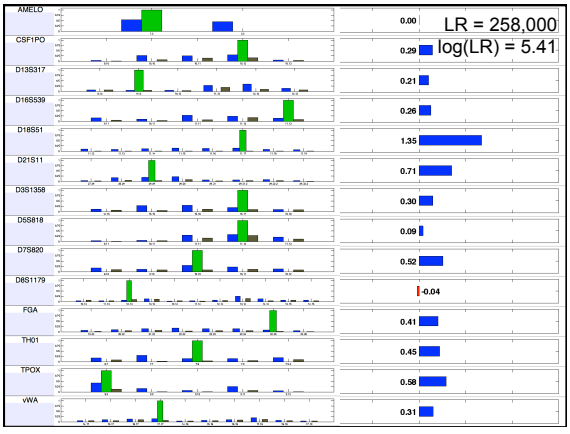


Paternity index

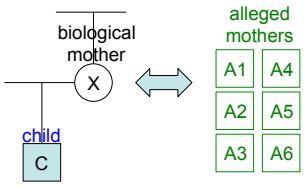


10,10

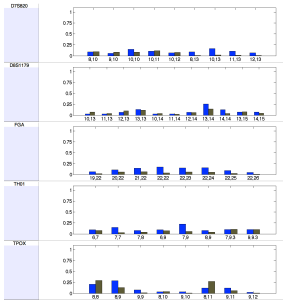
LR = 3.28
log(LR) = 0.52

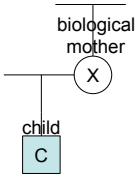


Maternity question

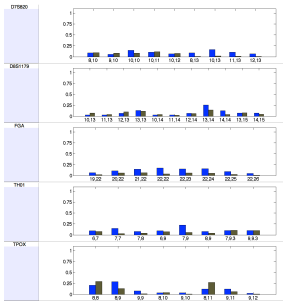


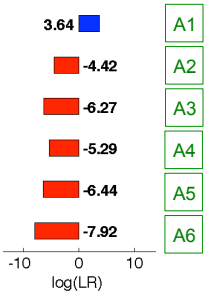
Biological mother genotype



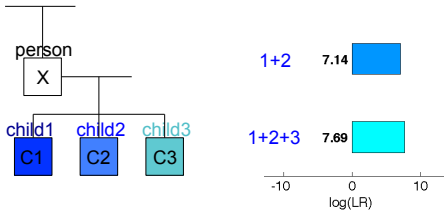


Identifying the mother

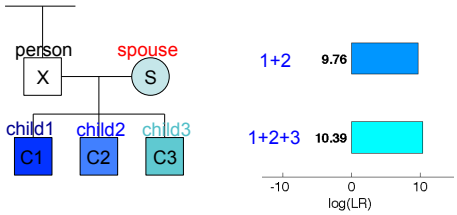




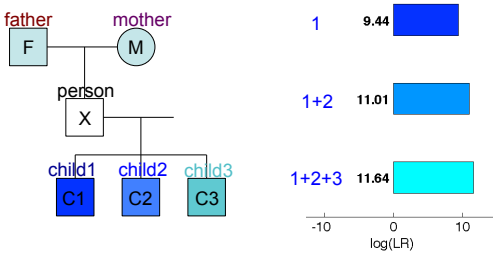
Missing person



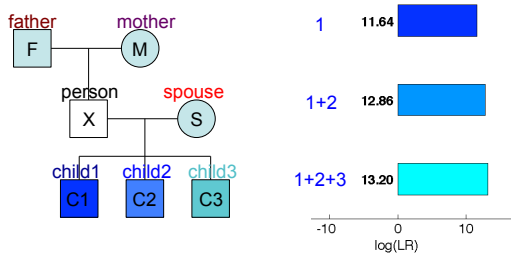
Children + Spouse



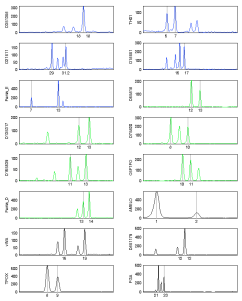
Parents + Children



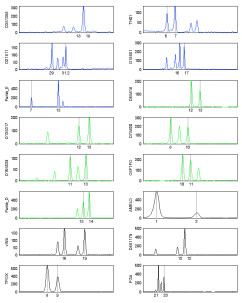
Parents + Children + Spouse

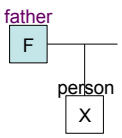


DNA mixture

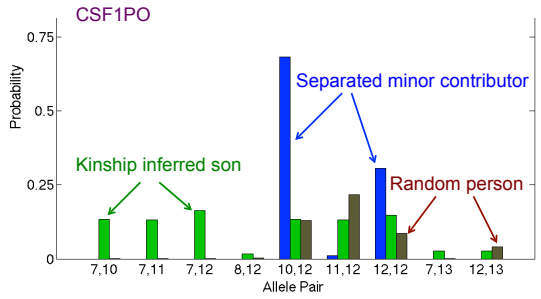


Mixture + Kinship





Mixture + Kinship genotypes



Contributor genotype match

LR = 39,700
log(LR) = 4.60

$$\sum_{x \in G} \frac{\Pr(X = x | d_M) \Pr(Y = x | d_K)}{\Pr(X = x)}$$

mixture kinship

population

-18.00
-20.58
-20.66
-20.76
-22.00
-24.00
-24.00
-24.97
-26.00
-28.00

How TrueAllele Works

Part 1, 16-Oct-2014
Genotype modeling and the likelihood ratio

Part 2, 20-Nov-2014
Degraded DNA and allele dropout

Part 3, 18-Dec-2014
Kinship, paternity and missing persons

Part 4, 15-Jan-2015
Genotype database and DNA investigation

http://www.cybgen.com/information/webinar/page.shtml

Cybergenetics

perlin@cybgen.com