Workshop: Paternity and kinship testing including X-chromosomal markers

Thore Egeland and Daniel Kling

(1) Norwegian University of Life Sciences, (2) Department of Forensic Services, OUS, Norway

First lecture: Genetical and statistical background

Thore Egeland[(1)], Daniel Kling[(2)]

(1) Norwegian University of Life Sciences, (2) Department of Forensic Services, OUS, Norway

ISFG workshop, Sep 3-4 2018, Catanzaro (Calabria), Italy,
`http://familias.name/isfg-kinship-2018/`

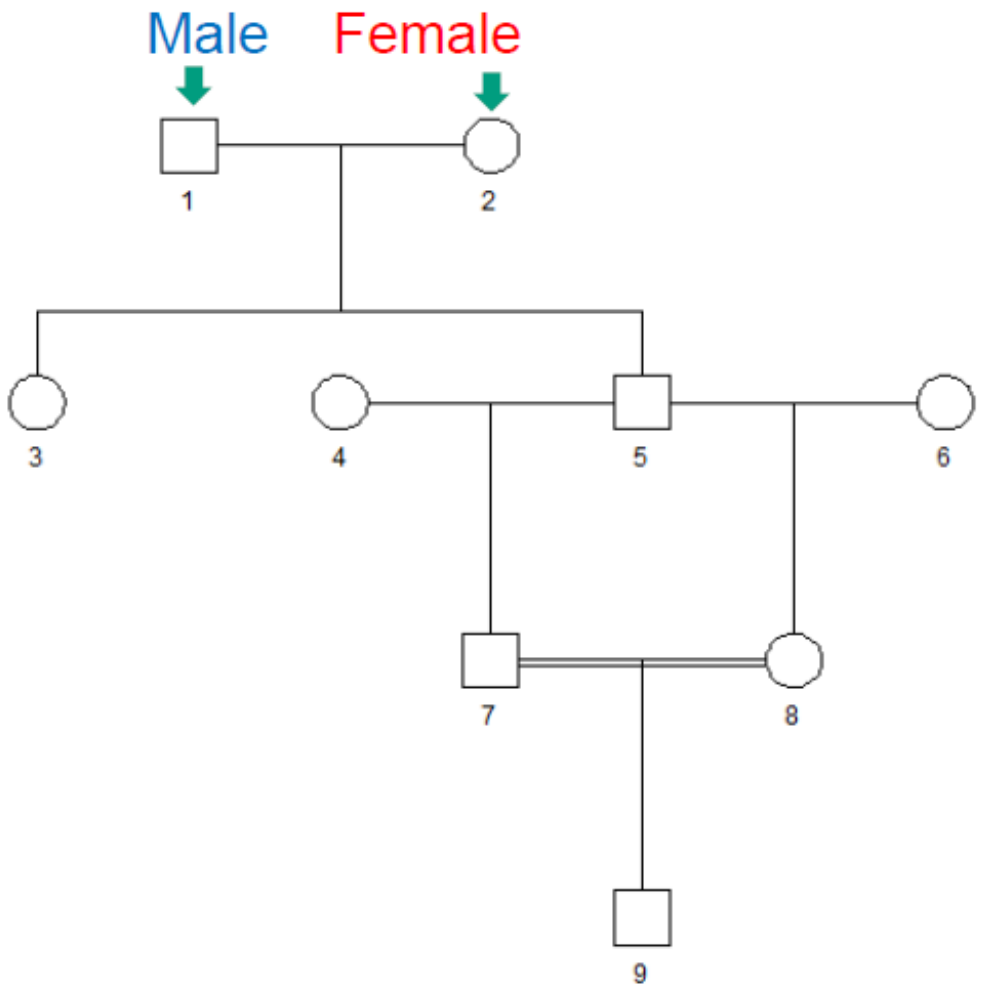
Program: Lectures and exercises

`http://familias.name/isfg-kinship-2018/`

Contents 09:15–10:00 and 10:15–11:30

- Basic forensic genetics very briefly:
  - Mendelian inheritance
  - Markers: autosomal, X, Y, mtDNA, STR-s.
- Weight of evidence. Likelihood Ratio (LR). Assumptions.
- Combining information. Bayes theorem
- Complications:
  - Mutation.
  - Theta correction.
  - Silent alleles.
- Introduction to Familias

Pedigree



Genetic markers I

- Small parts of the genome which ...
  - have known position
  - vary in the population
  - are easy to genotype

allele = a manifestation of a variable locus

- SNPs (single nucleotide polymorphisms)
  - two alleles
  - usual requirement: MAF > 1%
  - very common in the genome (millions!)
  - used in medical genetics +++

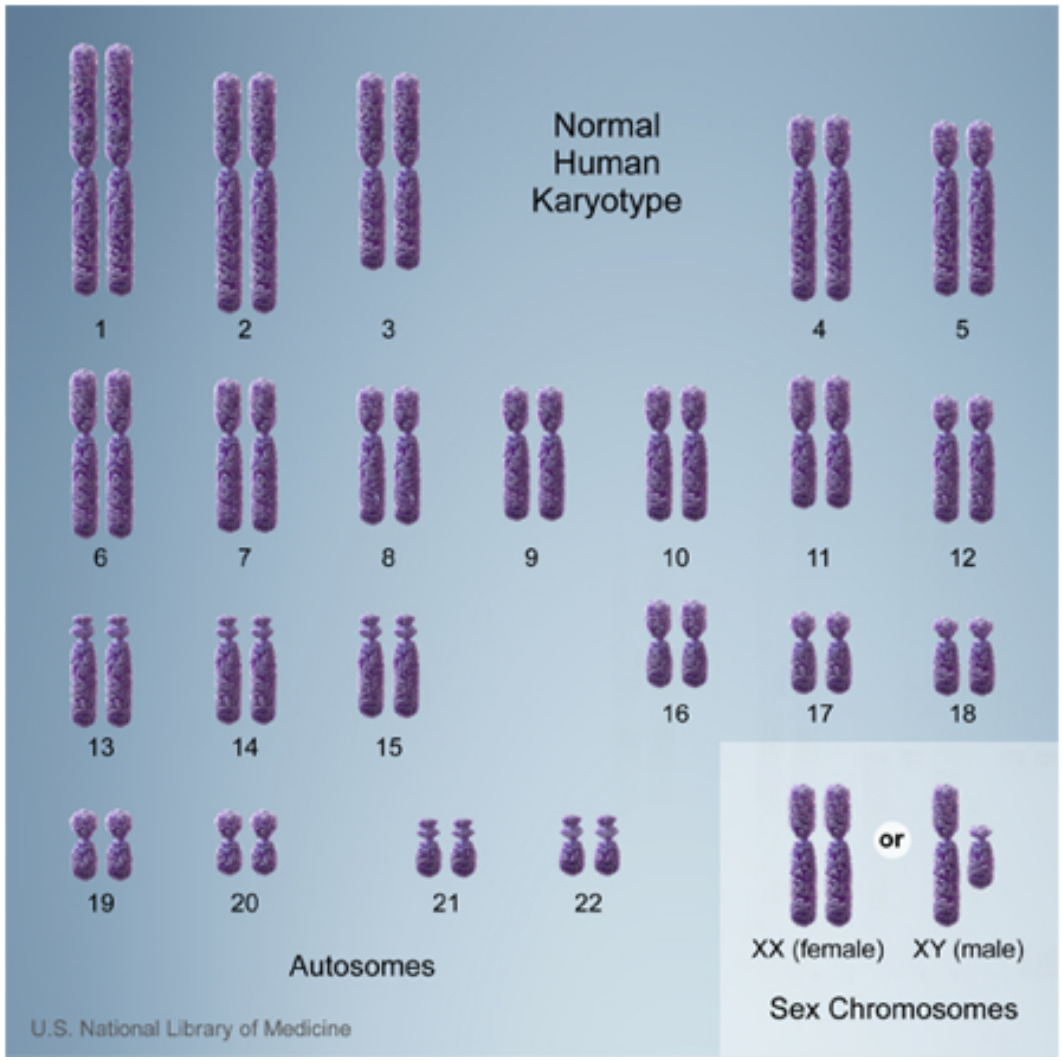
MAF= minor allele frequency

...CCGTTATATGGGC...
...CCGTTAGATGGGC...
...CCGTTATATGGGC...
...CCGTTATATGGGC...
...TTAGATGGGC...

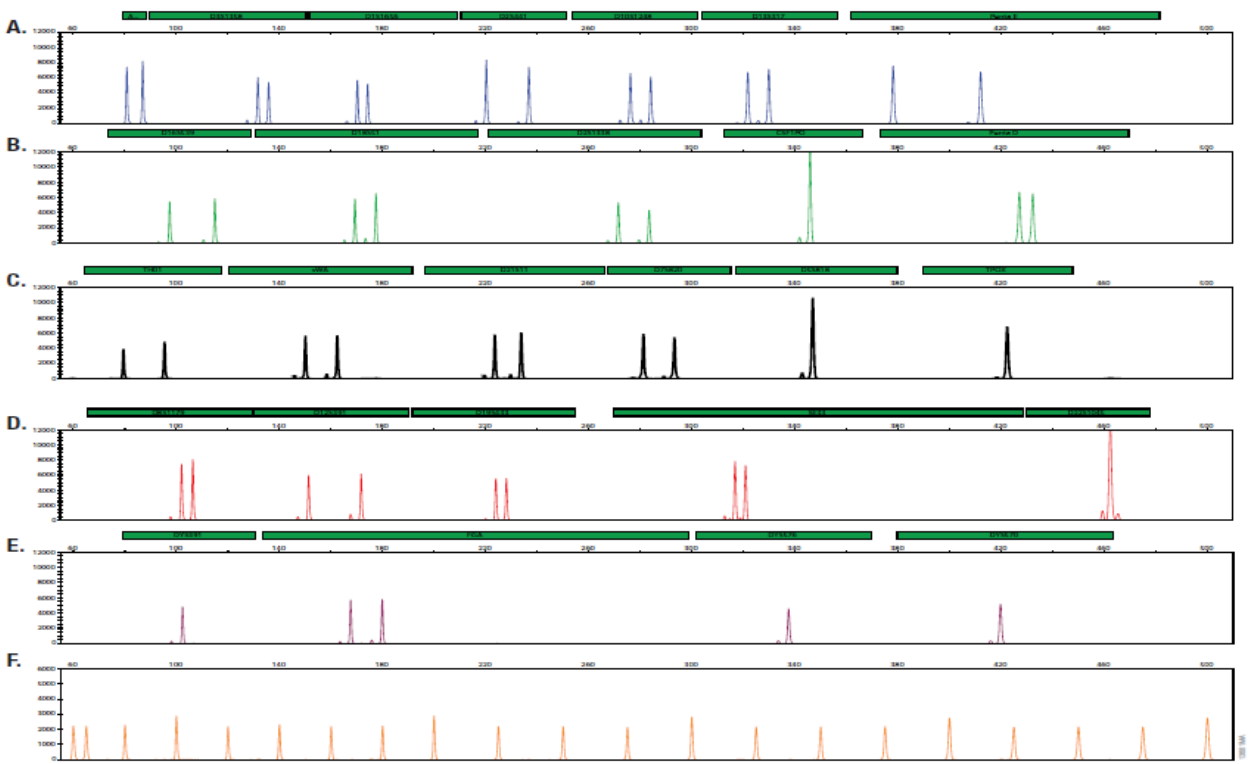
- STRs (short tandem repeats) = microsatellites
  - consecutive repeats of 2-5 bases
  - multiallelic: 5 - 50 alleles
  - allele names: # repeats
  - used in forensics

Genetic markers II

- chrom 1 - 22:
  - called the *autosomes*
  - all have 2 alleles at each locus
  - autosomal markers
- X chromosome
  - males have 1 allele
  - females have 2
  - X-linked markers
- Y chromosome
  - males have 1 allele
  - females 0
  - Y-linked markers



. Genetic markers III. Example: Fusion 6C



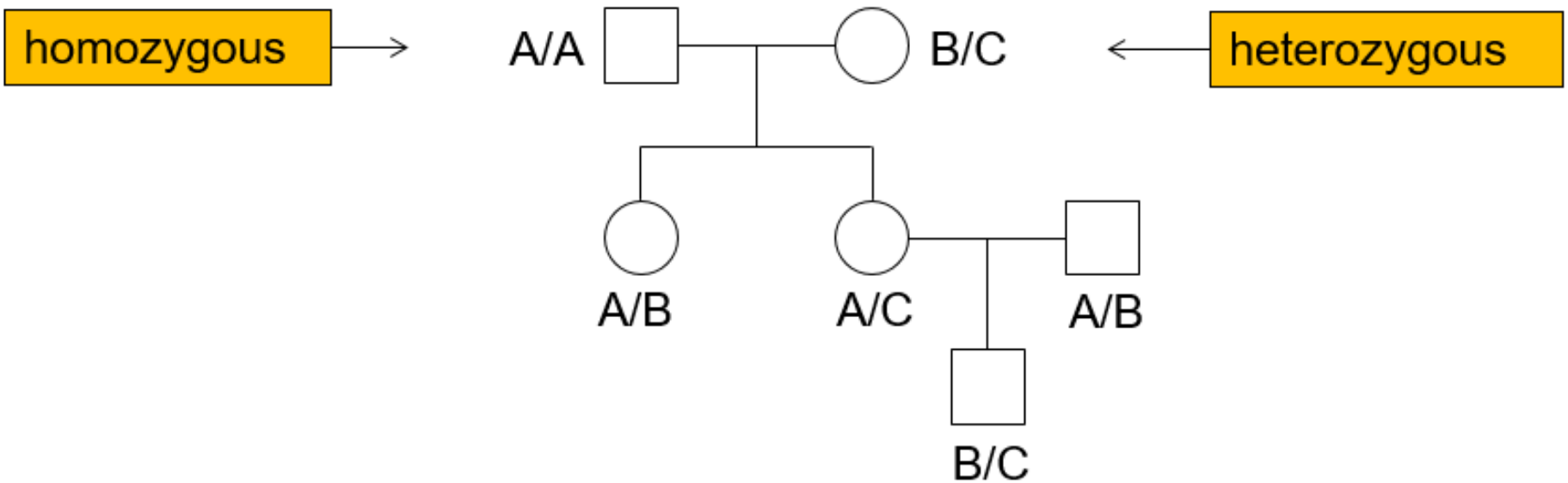
Promega

Promega Corporation · 2800 Woods Hollow Road · Madison, WI 53711-5399 USA · Toll Free in USA 800-356-9526 · 608-274-4330 · Fax 608-277-2516
TMD045 · Revised 4/17
www.promega.com

Figure 27. The PowerPlex® Fusion 6C System. The 2800M Control DNA (1.0ng) was amplified using the PowerPlex® Fusion 6C System and 29 cycles. Amplification products were mixed with WEN Internal Lane Standard 500 and analyzed using an Applied Biosystems® 3500xL Genetic Analyzer and a 1.2kV, 24-second injection. Results were analyzed using GeneMapper® *ID*-X software, version 1.4. **Panel A.** An electropherogram showing the peaks of the FL-6C-labeled loci: Amelogenin, D3S1358, D1S1656, D2S441, D10S1248, D13S317 and Penta E. **Panel B.** An electropherogram showing the peaks of the JOE-6C-labeled loci: D16S539, D18S51, D2S1338, CSF1PO and Penta D. **Panel C.** An electropherogram showing the peaks of the TMR-6C-labeled loci: TH01, vWA, D21S11, D7S820, D5S818, and TPOX. **Panel D.** An electropherogram showing the peaks of the CXR-6C-labeled loci: D8S1179, D12S391, D19S433, SE33 and D22S1045. **Panel E.** An electropherogram showing the TOM-6C-labeled loci: DYS391, FGA, DYS576 and DYS570. **Panel F.** An electropherogram showing the 60bp to 500bp fragments of the WEN Internal Lane Standard 500.

Mendelian inheritance

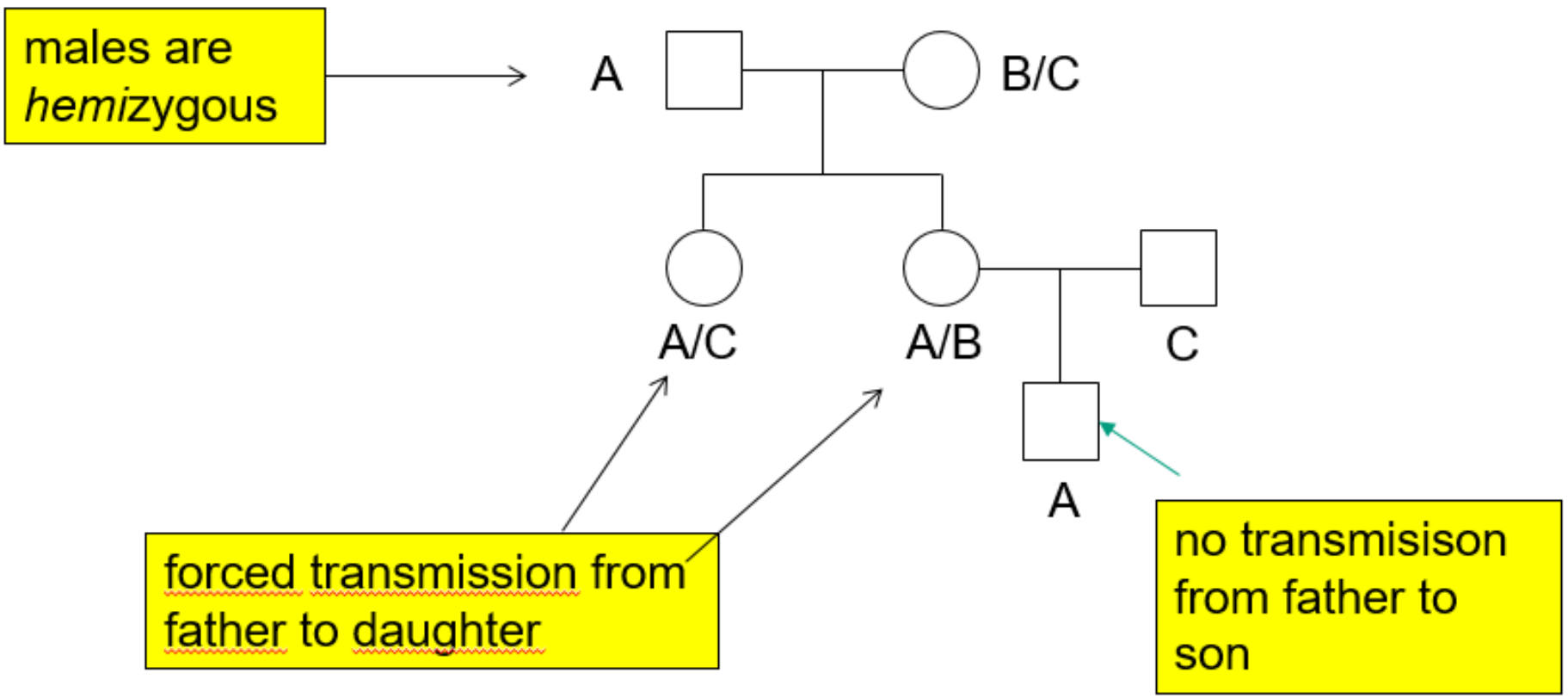
Example: autosomal marker with 3 alleles: A, B, C



Probability of transmitting either allele: **always 50%**

X linked inheritance

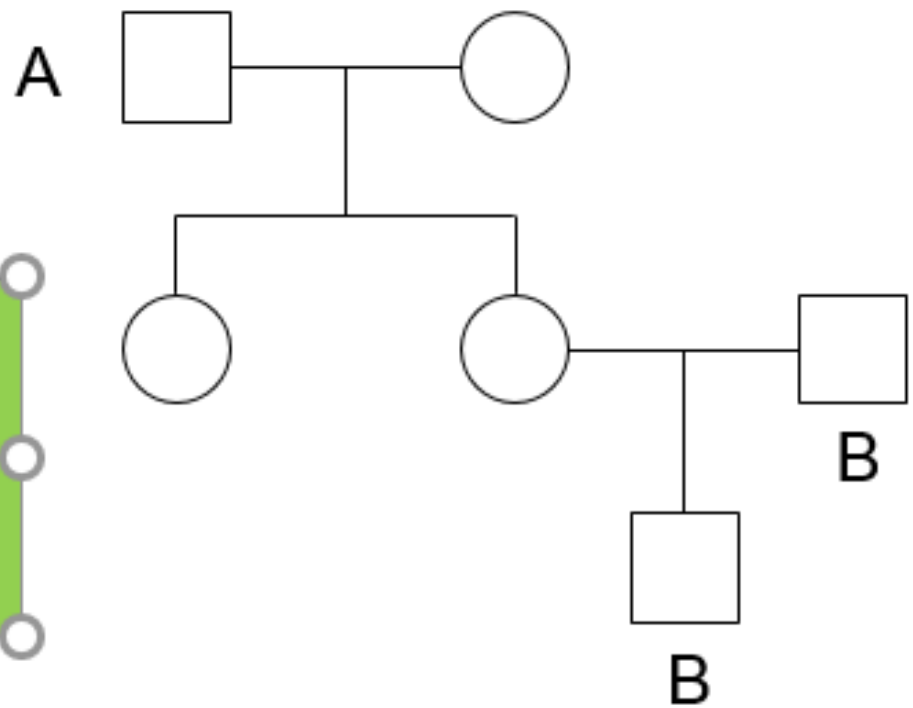
Example: X-linked marker with 3 alleles: A, B, C



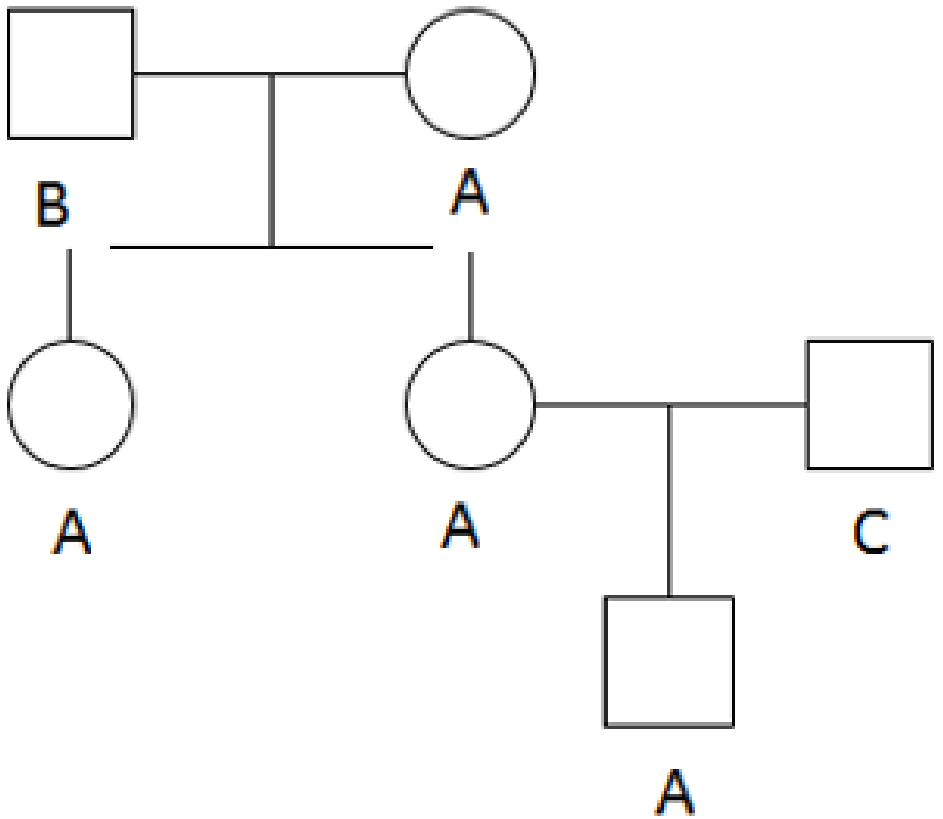
Y linked inheritance

Example: Y-linked marker with 2 alleles: A, B

- no transmission involving females
- father-son forced

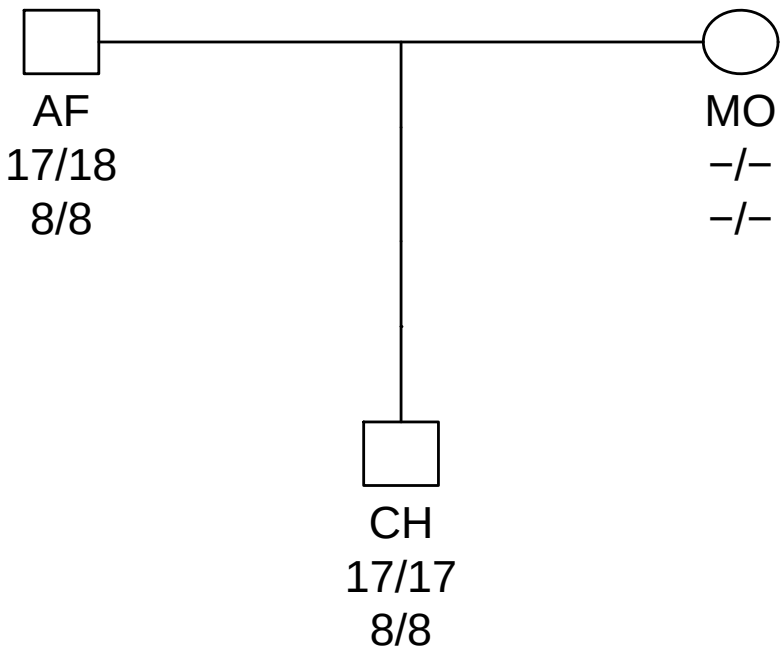


Mitochondrial (mtDNA) inheritance



Passed on from *mother* to all children

Hypotheses



- H_1: AF biological father of CH.
- H_2: AF and CH unrelated.
- Notation. Sometimes:
- $H_1 = H_P$:
  "prosecution hypothesis",
- $H_2 = H_D$:
  "defence hypothesis".

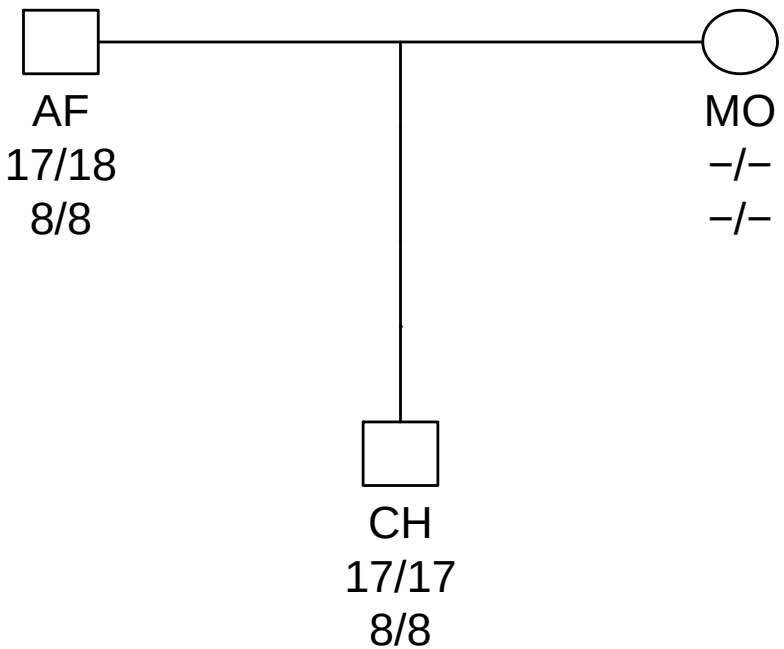
Likelihood ratio. Definition

Forensic framework

$$LR = LR_{H_1,H_2}(E) = \frac{P(E \mid H_1)}{P(E \mid H_2)}$$

is the likelihood ratio for evidence E with respect to the two hypotheses H_1 and H_2. The LR measures how much better H_1 explains the evidence E than H_2.

Likelihood Ratio. Example



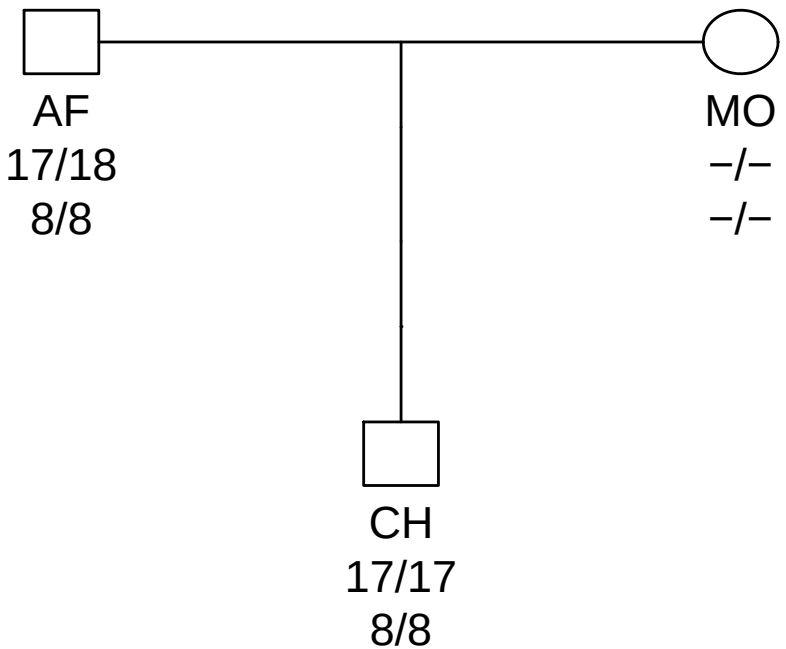
$$LR = \frac{P(E \mid H_1)}{P(E \mid H_2)} = \cdots = \frac{P(g_{CH} \mid g_{AF})}{P(g_{CH})}$$

$$LR_1 = \frac{\frac{1}{2}p_{17}}{p_{17}^2} = \frac{1}{2 \times 0.204} = 2.45$$

$$LR_2 = \frac{p_8}{p_8^2} = \frac{1}{0.554} = 1.81$$

$$LR = LR_1 \times LR_2 = 2.45 \times 1.81 = 4.4.$$

Likelihood Ratio. Interpretation and assumptions



- Interpretation LR=4.4: The data is 4.4 times more likely if AF is assumed to be the father compared to the unrelated alternative.
- Assumptions
  - Hardy–Weinberg Equilibrium (HWE).
  - Independent markers.
  - No artefacts: (no mutation, no silent alleles, no drop–out/in, no error).

Realistic number of markers

Marker	CH	AF	LR	LR(mut)
D3S1358	17/17	17/18	2.45	2.45
TPOX	8/8	8/8	1.81	1.80
D6S474	16/17	14/15	0.000	0.001
...	...	...	...	...
D19S433	12/15	12/14	3.34	3.34
Total			0	25070642

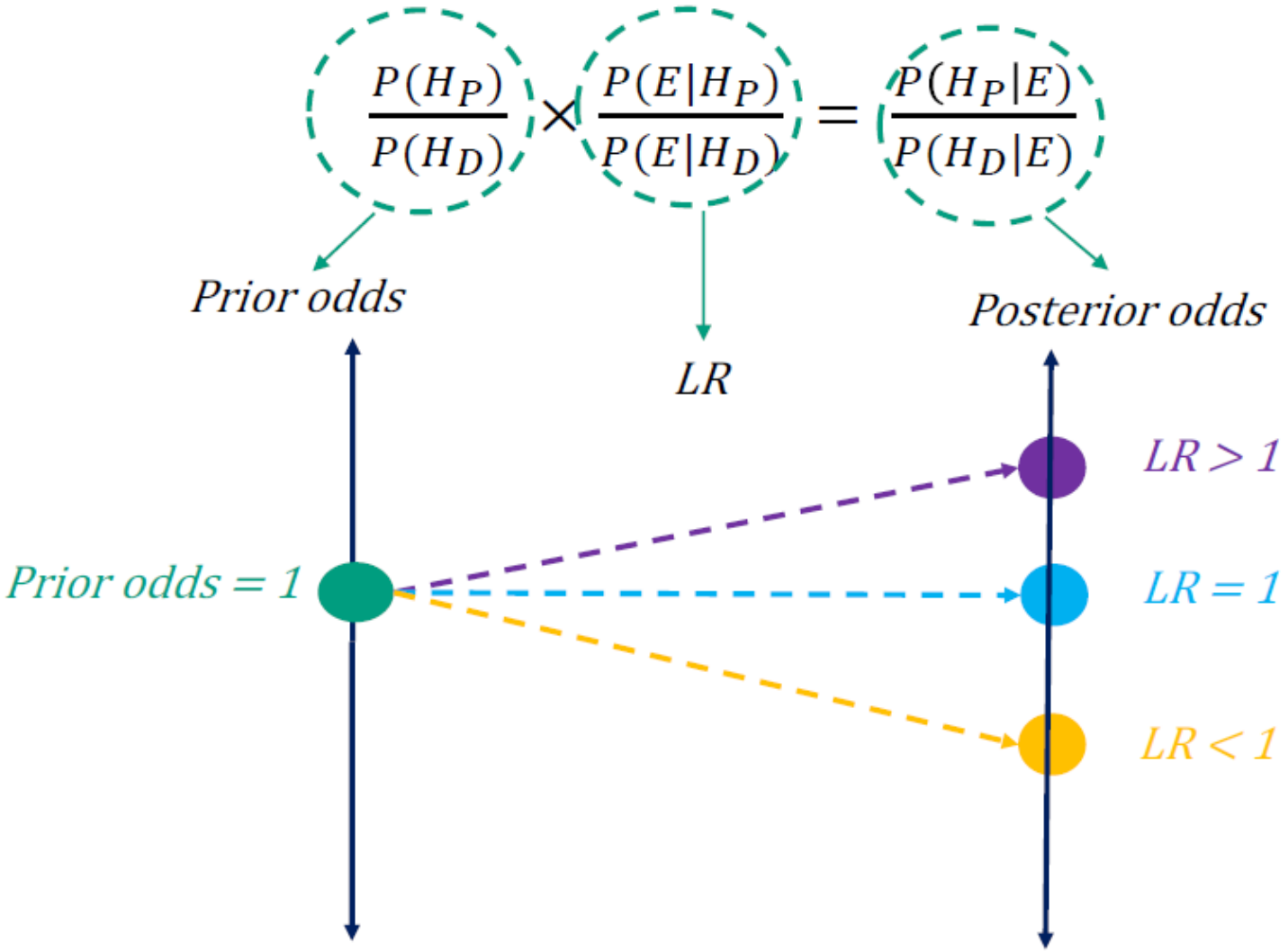
W = Posterior probability of paternity

- Assume *prior probabilities* $P(H_P) = P(H_D) = 0.5$ (*reasonable?*)
- *Prior odds* $\frac{P(H_P)}{P(H_D)} = 1$.

Then

$$W = P(H_P \mid E) = \frac{LR}{LR+1} = \frac{25070642}{25070642+1}$$
$$= 0.99999996 = \text{"Probability of } H_P \text{ given evidence"}$$

Bayes theorem on odds form



Blackstone ratio

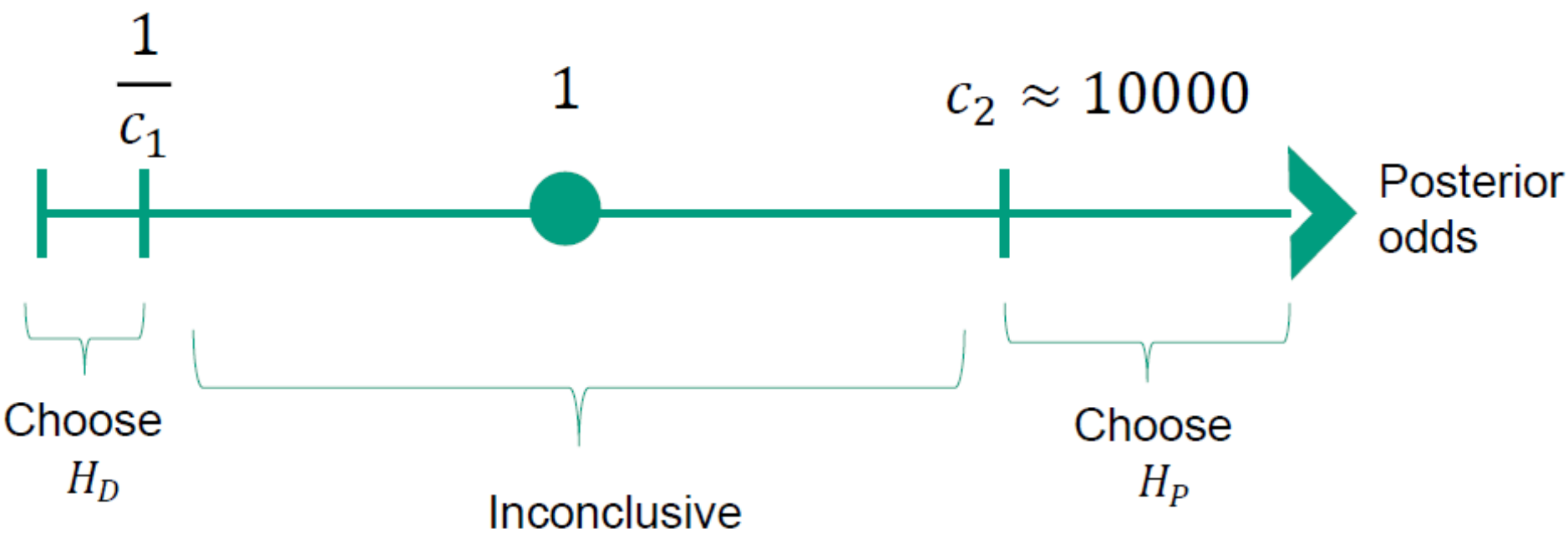
- Blackstone's ratio:

$(1 + c_2)/(1 + c_1) = 10$ (in practice much higher.)

		TRUTH	
		Guilt H_P	Innocence H_D
VERDICT	Guilt H_P	0	$1 + c_2$
	Innocence H_D	$1 + c_1$	0

Make no decision: cost = 1

Optimal decision rule



If c_1 and c_2 are specified, an optimal decision rule can be determined.
See Tillmar and Mostad (2014) for an application

Adding evidence I

- If prior odds $= 0$ or $LR = 0$

  ⟹ posterior odds $= 0$

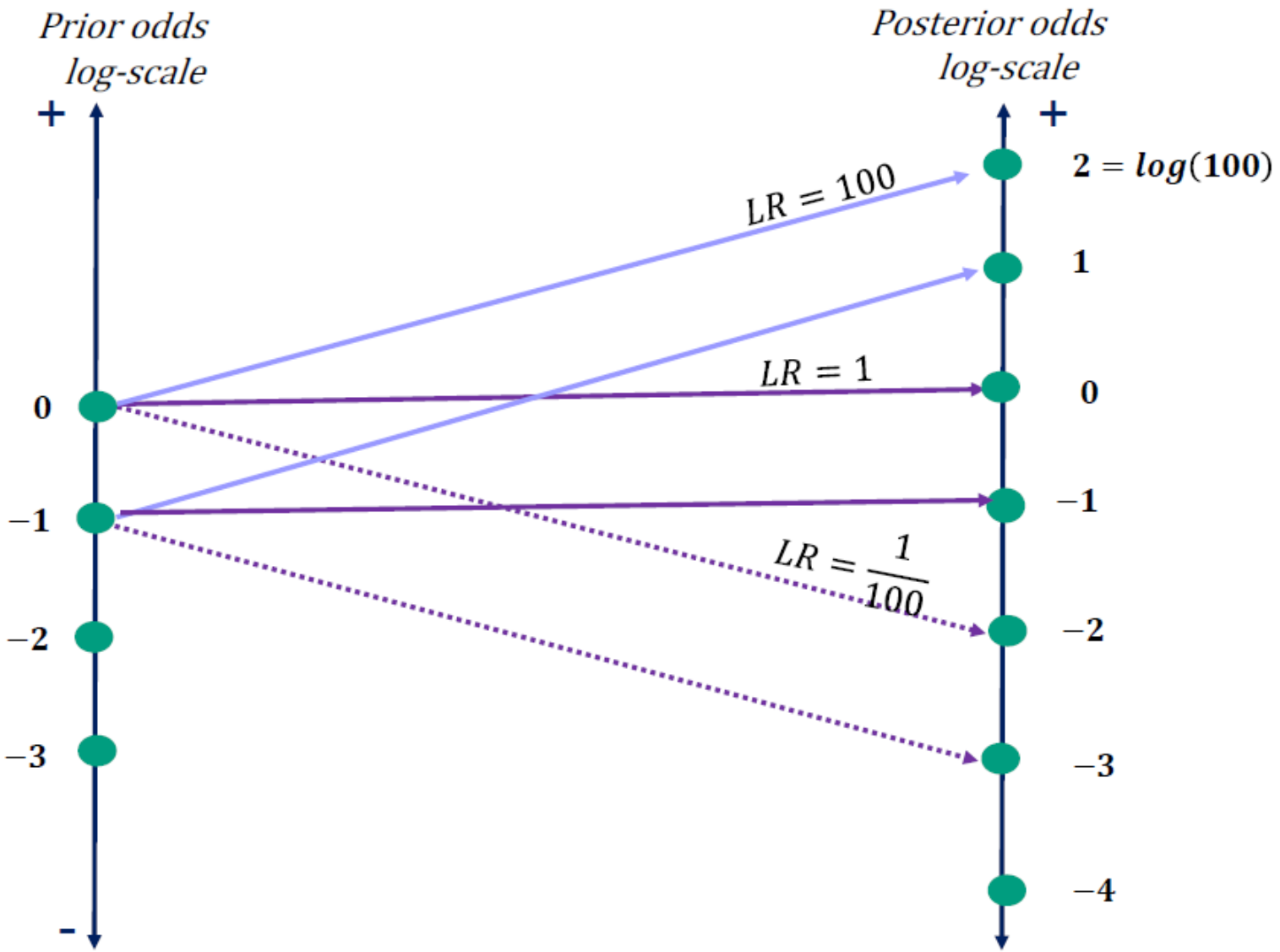
- Assume prior odds > 0 and $LR > 0$. Then

  $$\log(\text{prior odds}) + \log(\text{LR}) = \log(\text{posterior odds})$$

- $\log(\text{LR}) = \log_{10}(\text{LR})$ (unit called "ban" - Alan Turing)

*Good IJ (1985)

Adding evidence II



Theta correction. Dispute laid to rest

nature

SEARCH JOURNAL Go

Journal Home
Current Issue
AOP
Archive

THIS ARTICLE

Download PDF
References

Export citation
Export references

Send to a friend

commentary

Nature **371**, 735 - 738 (27 October 2002); doi:10.1038/371735a0

DNA fingerprinting dispute laid to rest

ERIC S. LANDER & BRUCE BUDOWLE

Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachu
Forensic Science Research and Training Center, FBI Laboratory, FBI Academy, Quantico, Vi

Hardy Weinberg

Two alleles A, B.

$p_A = 0.4, p_B = 0.6.$

Fraction A/A: 0.4^2	$= 0.160$
Fraction A/B: $2*0.4*0.6$	$= 0.480$
Fraction B/B: 0.6^2	$= 0.360$
Sum	$= 1.000$

Problem: Above requires HW, not valid if 'unrelated people' are slightly related

Solution: theta-correction

Theta - correction

$$Homozygous \quad A,A: \ \theta p_A + p_A^2(1-\theta),$$

$$Heterozygous \quad A,B: \ 2p_A p_B(1-\theta).$$

$\theta = 0.1$ (extreme case)

$$\text{Fraction A/A: } 0.1*0.4+0.4^2*(1-0.1) = 0.184 > 0.160$$

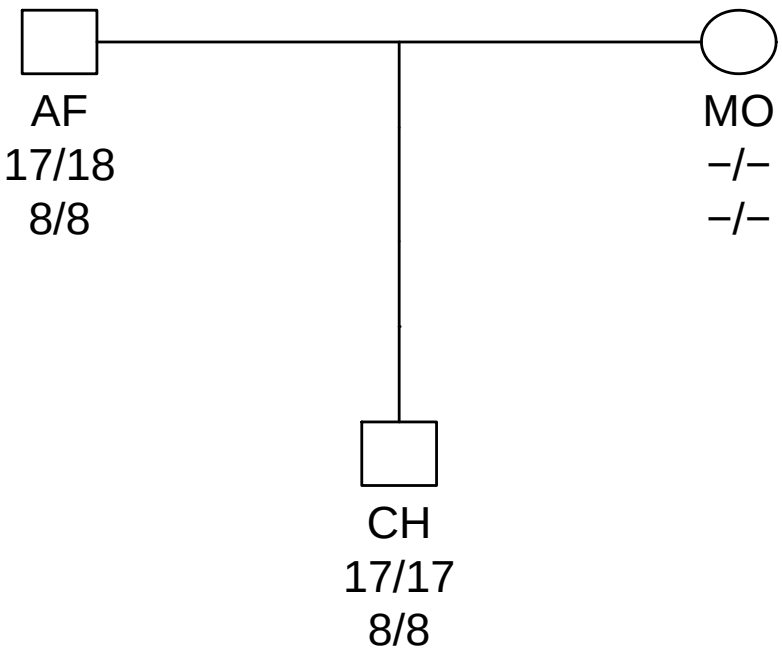
$$\text{Fraction A/B: } 2*0.4*0.6*(1-0.1) = 0.432 < 0.480$$

$$\text{Fraction B/B: } 0.1*0.6+0.6^2*(1-0.1) = 0.384 > 0.360$$

Sum	1.000	1.000

- Fraction homozygotes increases with θ

Previous example

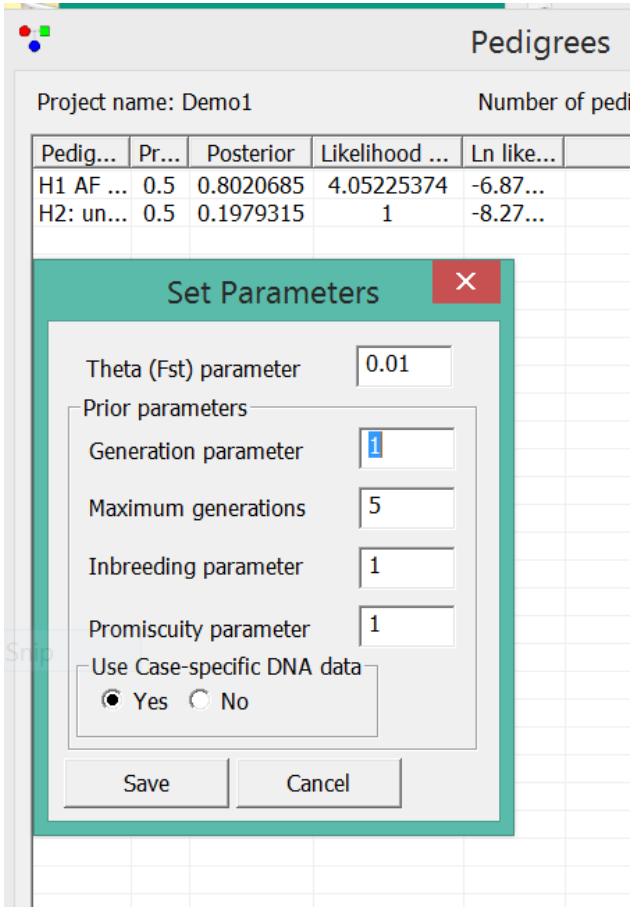


$$LR_1 = \frac{\frac{1}{2}p_{17}}{p_{17}^2} = \frac{1}{2 \times 0.204} = 2.45$$

$$LR_2 = \frac{p_8}{p_8^2} = \frac{1}{0.554} = 1.81$$

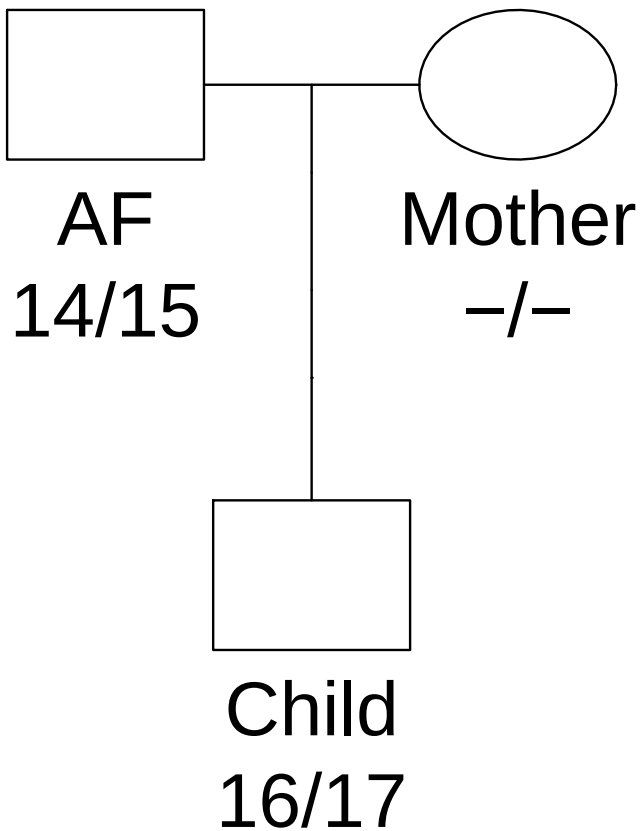
$$LR = LR_1 \times LR_2 = 2.45 \times 1.81 = 4.4.$$

Fst=Theta=0.01. Input: Pedigrees > Parameters



- $LR = 2.298 \cdot 1.764 = 4.1$ when $\theta = 0.01$
- $LR = 2.45 \cdot 1.81 = 4.4$ when $\theta = 0.00$.

Mutation: Exercise 2.2, 2.7



Mutation: Results, simillar to Exercise 2.9

System	Child	AF	LR	LR(mut)
D3S1358	17/17	17/18	2.450	2.449
TPOX	8/8	8/8	1.805	1.804
D6S474	16/17	14/15	0.000	0.001
3 markers			0.000	0.0044
	...	...	...	...
22 (1 mut)			0.000	25070642

Table 1: Genotype data for a Child and alleged father (AF) along with LR-s. The rightmost column is based on a stepwise unstationary mutation model (expained below) with mutation rate 0.001 and range 0.5 for all markers.

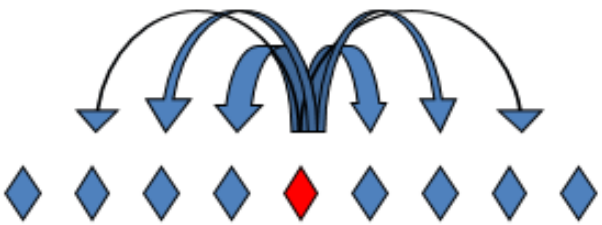
Equal model

	Mutation to					
Al...	13	14	15	16	17	18
13	0.999	0.0002	0.0002	0.0002	0.0002	0.0002
14	0.0002	0.999	0.0002	0.0002	0.0002	0.0002
15	0.0002	0.0002	0.999	0.0002	0.0002	0.0002
16	0.0002	0.0002	0.0002	0.999	0.0002	0.0002
17	0.0002	0.0002	0.0002	0.0002	0.999	0.0002
18	0.0002	0.0002	0.0002	0.0002	0.0002	0.999

- $P(\text{no mutation}) = 0.999$
- Problem: $P(14 \to 17) = P(15 \to 16) = 0.0002.$

Mutation: Biology

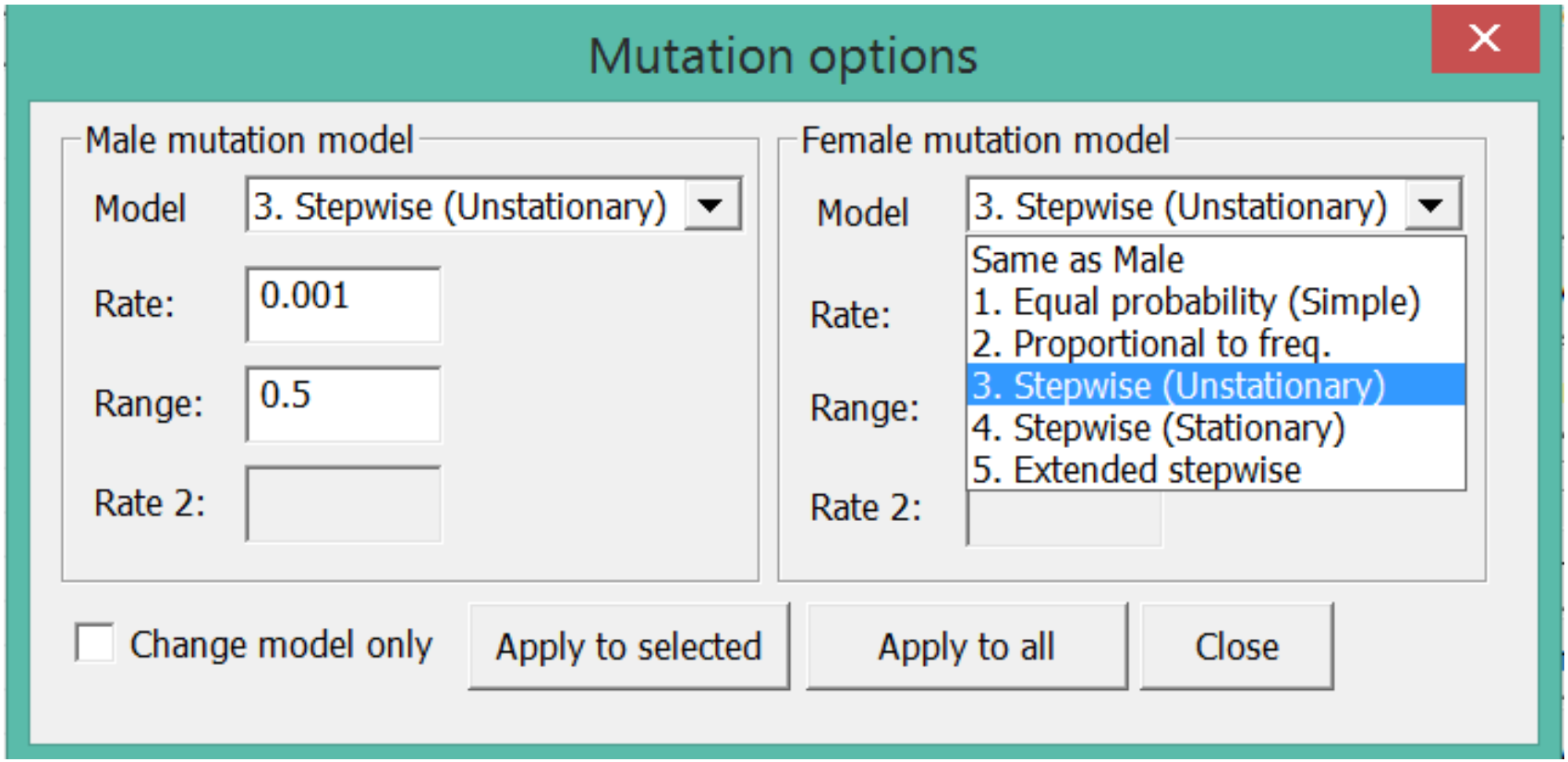
- Mutation rate varies with
  - Sex of parent and locus.
    Alleles tend to mutate to close alleles:



  - Several models

- Mutation rates:
  `http://www.cstl.nist.gov/strbase/mutation.htm`

Mutation: Input

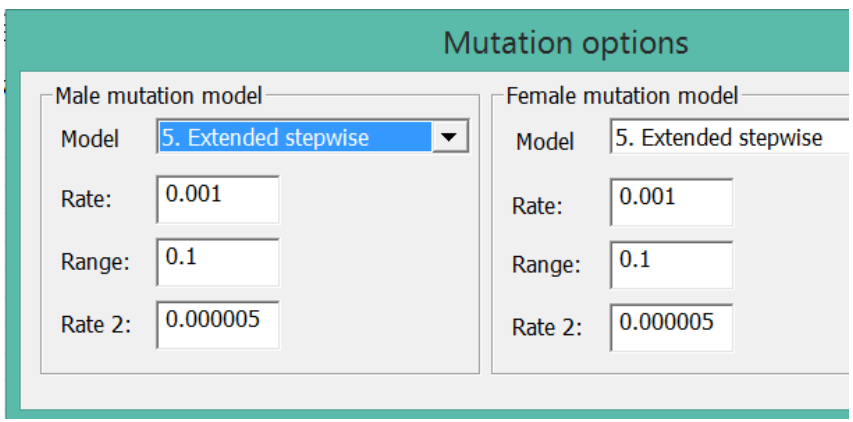


Mutation Matrix. Range

	Mutation to					
Al...	13	14	15	16	17	18
13	0.999	0.000...	0.000258...	0.00012903	6.4516e-005	3.2258e-
14	0.0003...	0.999	0.000347...	0.00017391	8.6957e-005	4.3478e-
15	0.0001...	0.000...	0.999	0.00030769	0.00015385	7.6923e-
16	7.6923...	0.000...	0.000307...	0.999	0.00030769	0.00015
17	4.3478...	8.695...	0.000173...	0.00034783	0.999	0.00034
18	3.2258...	6.451...	0.000129...	0.00025806	0.00051613	0.999

$$\text{Range} = \frac{15 \to 17}{15 \to 16} = \frac{0.00015385}{0.00030769} = 0.5$$

Extended stepwise model: Generally recommended, consistent for microvariants



- Rate Integer mutations: $9 \rightarrow 10$, $9.3 \rightarrow 10.3$
- Range: As before. 1 step $1/0.1 = 10$ times more likely than two steps, etc.
- Rate2 Fractional Mutations: $9.3 \rightarrow 10$, $9 \rightarrow 9.3$

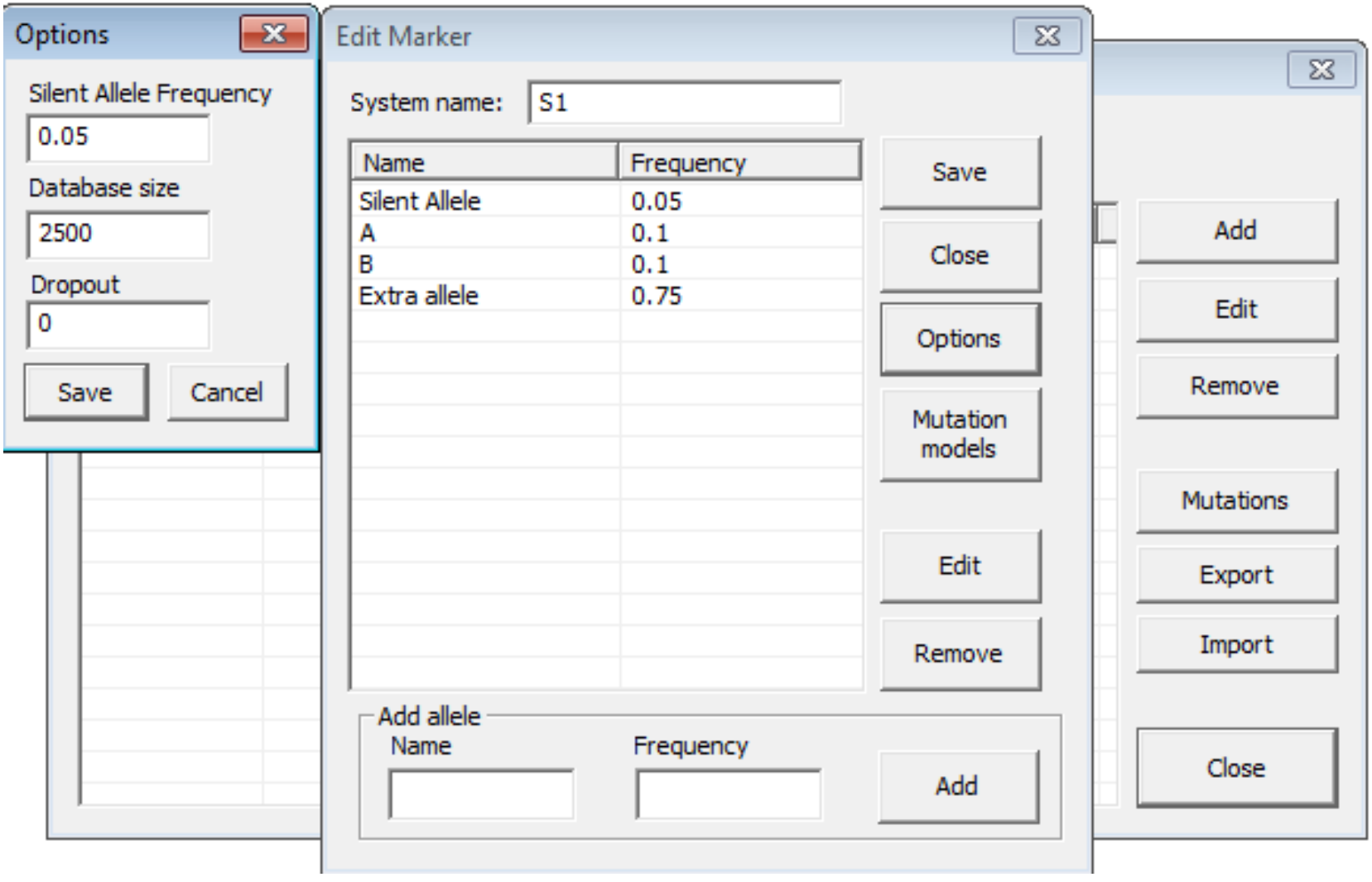
Extended stepwise model: Example

	Mutation to				
A...	9	9.3	10	10.3	11
9	0.998995	2.5e-006	0.0009090909091	2.5e-006	9.090909091e-005
9.3	1.666666667e-006	0.998995	1.666666667e-006	0.001	1.666666667e-006
10	0.0005	2.5e-006	0.998995	2.5e-006	0.0005
10.3	1.666666667e-006	0.001	1.666666667e-006	0.998995	1.666666667e-006
11	9.090909091e-005	2.5e-006	0.0009090909091	2.5e-006	0.998995

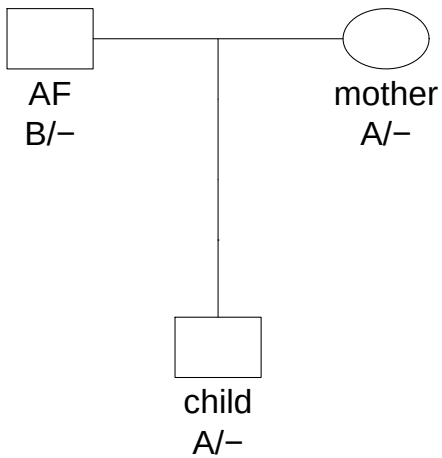
- Recall: Rate $= 0.001$, Range $= 0.1$, Rate2 $= 0.000005$.
- Note: $P(\text{“no mut”}) = 1 - (0.001 + 0.000005) = 0.998995$.

$$P(9 \to 10) = 0.0009 >$$
$$P(9 \to 9.3) = 2.5e-006 = 0.0000025$$
$$\frac{P(9 \to 11)}{P(9 \to 10)} = \frac{0.00009}{0.0009} = 0.1$$

Silent alleles. Input: General DNA data > . . . > Options



Exercise 2.11



- Enter genotypes as homozygous.
- $p_B = 0.1,\ p_S = P(\text{“Silent allele”}) = 0.05$. Now e.g.
- $P(B/-) = p_B^2 + 2 \cdot p_B p_S = 0.02$.
- Statistics for silent alleles:
  `http://www.cstl.nist.gov/strbase/NullAlleles.htm`


T Egeland, D Kling, and P Mostad.
Relationship Inference with Familias and R: Statistical Methods in Forensic Genetics.
Academic Press, 2015.


IJ Good.
Weight of evidence: A brief survey.
Bayesian Statistics, 1985.


A Tillmar and P Mostad.
Choosing supplementary markers in forensic casework.
Forensic Science International: Genetics, 13:128–133, 2014.