

Theta correction, mutations, silent alleles

Thore Egeland^{(1),(2)}

(1) Norwegian University of Life Sciences, (2) NIPH

thore.egeland@gmail.com

Beyond standard cases

- ▶ **Complex pedigrees** Large, inbred
- ▶ **Theta correction** If e.g. Hardy-Weinberg does not apply
- ▶ **Mutation** If e.g. parent and child share no alleles
- ▶ **Silent alleles** Homozygote or silent allele?
- ▶ **Drop out** Homozygote or drop out?

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, Massachusetts
Forensic Science Research and Training Center, FBI Laboratory, FBI Academy, Quantico, Vi

Hardy Weinberg

Two alleles A, B .

$$p_A = 0.4, p_B = 0.6.$$

$$\text{Fraction } A/A: 0.4^2 = 0.160$$

$$\text{Fraction } A/B: 2 * 0.4 * 0.6 = 0.480$$

$$\text{Fraction } B/B: 0.6^2 = 0.360$$

$$\text{Sum} = 1.000$$

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

Solution: theta-correction

Theta - correction

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

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

$\theta = 0.1$ (extreme case)

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

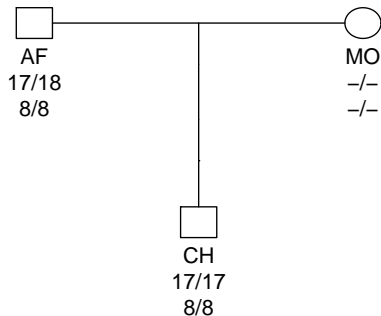
Fraction A/B: $2 * 0.4 * 0.6 * (1-0.1) = 0.432 < 0.480$

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

Pedigrees

Project name: Demo1 Number of pedi

Pedig...	Pr...	Posterior	Likelihood ...	Ln like...
H1 AF ...	0.5	0.8020685	4.05225374	-6.87...
H2: un...	0.5	0.1979315	1	-8.27...

Set Parameters ✕

Theta (Fst) parameter

Prior parameters

Generation parameter

Maximum generations

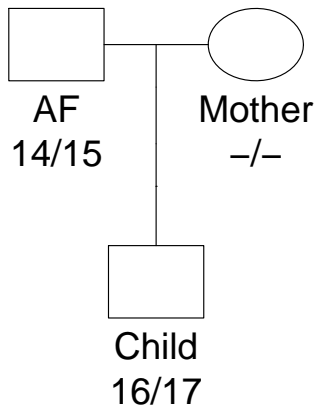
Inbreeding parameter

Promiscuity parameter

Use Case-specific DNA data
 Yes No

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

Mutation: Exercise 2.2, 2.7



Mutation: Results, similar 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 (explained 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:** $P(14 \rightarrow 17) = P(15 \rightarrow 16) = 0.0002$.

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 options

Male mutation model

Model: 3. Stepwise (Unstationary) ▾

Rate: 0.001

Range: 0.5

Rate 2:

Female mutation model

Model: 3. Stepwise (Unstationary) ▾

Rate: Same as Male
1. Equal probability (Simple)
2. Proportional to freq.
3. Stepwise (Unstationary) **Selected**
4. Stepwise (Stationary)
5. Extended stepwise

Range:

Rate 2:

Change model only

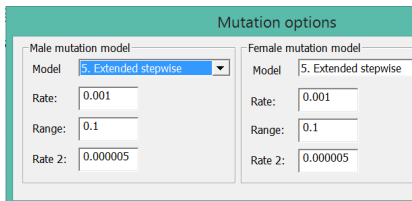
Apply to selected Apply to all Close

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	0.000153
17	4.3478...	8.695...	0.000173...	0.00034783	0.999	0.000347
18	3.2258...	6.451...	0.000129...	0.00025806	0.00051613	0.999

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

Extended stepwise model: Generally recommended, consistent for microvariants



The image shows a dialog box titled "Mutation options" with two columns: "Male mutation model" and "Female mutation model". Each column contains a "Model" dropdown menu set to "5. Extended stepwise", and three input fields for "Rate:", "Range:", and "Rate 2:". The values in the input fields are 0.001, 0.1, and 0.000005 respectively for both male and female models.

Mutation options	
Male mutation model	Female mutation model
Model: 5. Extended stepwise	Model: 5. Extended stepwise
Rate: 0.001	Rate: 0.001
Range: 0.1	Range: 0.1
Rate 2: 0.000005	Rate 2: 0.000005

- ▶ **Rate Integer mutations:**
9 → 10, 9.3 → 10.3
- ▶ **Range:** As before.
1 step $1/0.1 = 10$ times more likely than two steps, etc.
- ▶ **Rate2 Fractional Mutations:**
9.3 → 10, 9 → 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: $\Pr(\text{"no mut"}) = 1 - (0.001 + 0.000005) = 0.998995$.

$$\Pr(9 \rightarrow 10) = 0.0009 >$$

$$\Pr(9 \rightarrow 9.3) = 2.5e - 006 = 0.0000025$$

$$\frac{\Pr(9 \rightarrow 11)}{\Pr(9 \rightarrow 10)} = \frac{0.00009}{0.0009} = 0.1$$

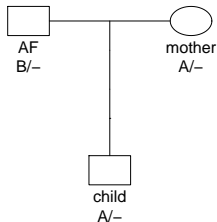
Silent alleles. Input: General DNA data > ... > Options

The screenshot displays three overlapping windows from a software application. The 'Options' window (top left) has a title bar with a close button and contains three input fields: 'Silent Allele Frequency' with the value 0.05, 'Database size' with 2500, and 'Dropout' with 0. Below these are 'Save' and 'Cancel' buttons. The 'Edit Marker' window (center) has a title bar with a close button and a 'System name' field containing 'S1'. It features a table with the following data:

Name	Frequency
Silent Allele	0.05
A	0.1
B	0.1
Extra allele	0.75

To the right of the table are buttons for 'Save', 'Close', 'Options', 'Mutation models', 'Edit', and 'Remove'. At the bottom of this window is an 'Add allele' section with 'Name' and 'Frequency' input fields and an 'Add' button. The third window (right) is partially obscured and shows a vertical stack of buttons: 'Add', 'Edit', 'Remove', 'Mutations', 'Export', 'Import', and 'Close'.

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>