Familias and Forensic Statistics

Thore.Egeland@umb.no
Daniel Kling
Norwegian University of Life Sciences (UMB) and
Norwegian Institute of Public Health
Manuel García Magariños
Universidade da Coruña
Lourdes Prieto Solla
Comisaría General de Policía Científica

Alcalá, Oct 8-9, 2012

Purpose

- "..practical workshop about the statistical evaluation of forensic cases (paternity and other family relationships) by using Familias"
- Linked markers and haplotypic evidence

Thore Egeland lectures unless otherwise stated.
Manuel García Magariños summarises in Spanish in between; discussion in English or Spanish.

October 8, 2012

09:00 – 10:00 **Part I.** Review of basic forensic genetics. Principles. LR for simple cases.
10:00 – 11:15 **Part II.** Introduction to Familias. Demo.
11:15 – 11:45 Coffee break.
11:45 – 12:45 **Part III.** Familias exercises.
12:45 – 14:00 **Part IV and V.** Non-standard cases. Complex relationships.
14:00 – 15:30 Lunch.
15:30 – 17:00 **Part VI.** Familias exercises.
17:00 – 18:00 Discussion (Spanish/English). More questions, contributions from participants.

October 9, 2012

09:00 – 10:45 **Part VII.** Mutation. Unseen alleles.
10:40 – 11:15 Familias 3, new mutation model **(Daniel Kling).**
11:15 – 11:45 Coffee break.
11:45 – 12:45 **Part VIII.** Familias exercises.
12:45 – 13:30 **Part IX and X.** Theta correction. Silent alleles.
13:30 – 14:00 mtDNA and Y. (**Lourdes**).
14:00 – 15:30 Lunch.
15:30 – 17:00 **Part XI.** The Bayesian alternative. Linked markers. Haplotypic evidence (**Daniel).**
17:00 – 17:30 **Part XII** Demo of FamLink software, freely available from http://famlink.se/ (**Daniel Kling**).
17:30 – 18:00 Discussion (Spanish/English). Summary.

Course homepage *arken.umb.no/~theg/alcala/*

Contents and purpose: Deal with non-standard cases

- Principles
- Statistical evaluation
  - LR (Paternity index)
  - Non-standard cases
    - Theta-correction
    - Mutations
    - Complex pedigrees
    - Silent alleles.
    - Linked markers.Haplotypic evidence
- Familias. Exercises
- FamLink. Demo

Part I. 09:00-10:00

- Review of basic forensic genetics

Principles

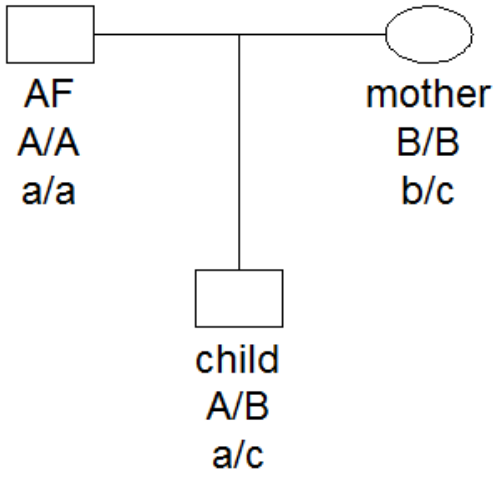
1. To evaluate the uncertainty of any given proposition it is necessary to consider at least one alternative proposition.
2. Scientific interpretation is based on questions of the kind: "What is the probability of the evidence given the proposition?"
3. Scientific evidence is conditioned not only by the competing propositions, but also by the framework of circumstances within which they are to be evaluated

Books: Evett&Weir, Balding, Buckleton et al.

Exercise S1 Standard paternity case

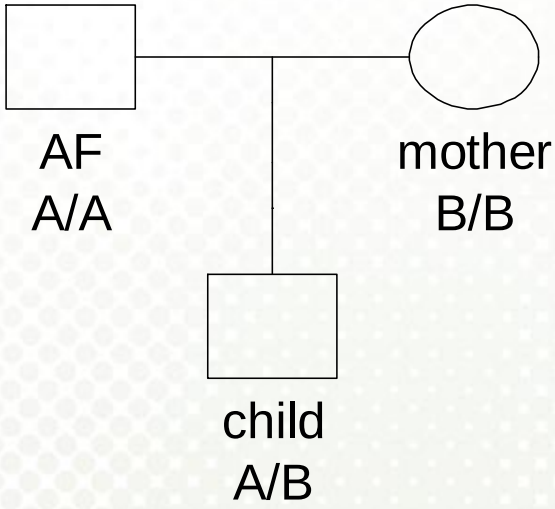
- H_1: The alleged father (AF) is the real father
- H_2: AF and the child are unrelated.

Figure1. Standard paternity case.



Exercise S1 a) Calculate LR. First marker

Standard paternity case. First marker



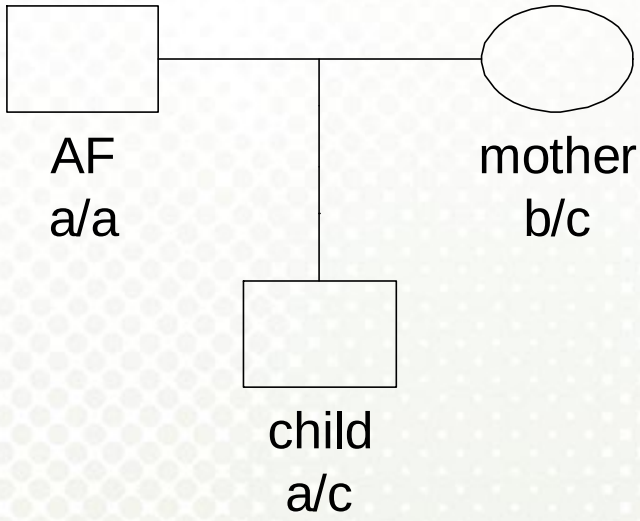
$$LR = \frac{P(child \mid mother, AF)}{P(child \mid mother)} = \frac{1}{p_A} = \frac{1}{0.05} = 20.$$

- *Interpretation:*

"The data is 20 times more likely assuming AF to be the father compared to the alternative that some unknown man is the father".

Exercise S1 c) Calculate LR. Second marker

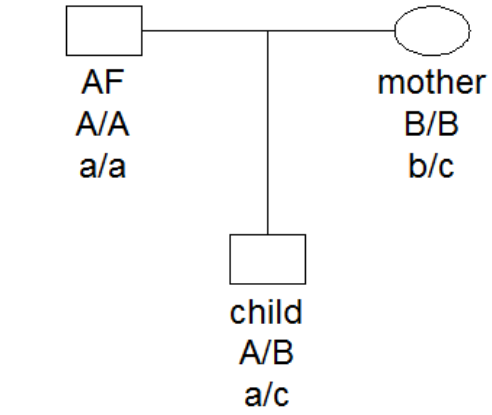
Standard paternity case. Second marker



$$LR = \frac{P(child \mid mother, AF)}{P(child \mid mother)} = \frac{0.5}{0.5 p_a} = \frac{0.5}{0.5 * 0.1} = 10.$$

Exercise S1 c) Calculate LR. Two markers

Figure1. Standard paternity case.



$$LR = 20 * 10 = 200.$$

- *Interpretation:*

"The data is 200 times more likely assuming AF to be the father compared to the alternative that some unknown man is the father".

- *Assumptions*?

Non-standard cases

- Complications
  - Theta-correction
  - Mutations
  - Complex pedigrees
  - Silent alleles

Part II. 10:00-11:15

- Introduction to Familias. Demo.
- Exercises S1-S3, S5. *Plenum*
- *Videos available*

Part III. 1145-1245

- Familias exercises S1-S6

Part IV and V: 12:45-14:00

- Non-standard cases.
- Complex relationships.
- *Introduced today*; more tomorrow

Non-standard cases

- Complications
  - Theta-correction
  - Mutations
  - Complex pedigrees
  - Silent alleles

Theta - correction

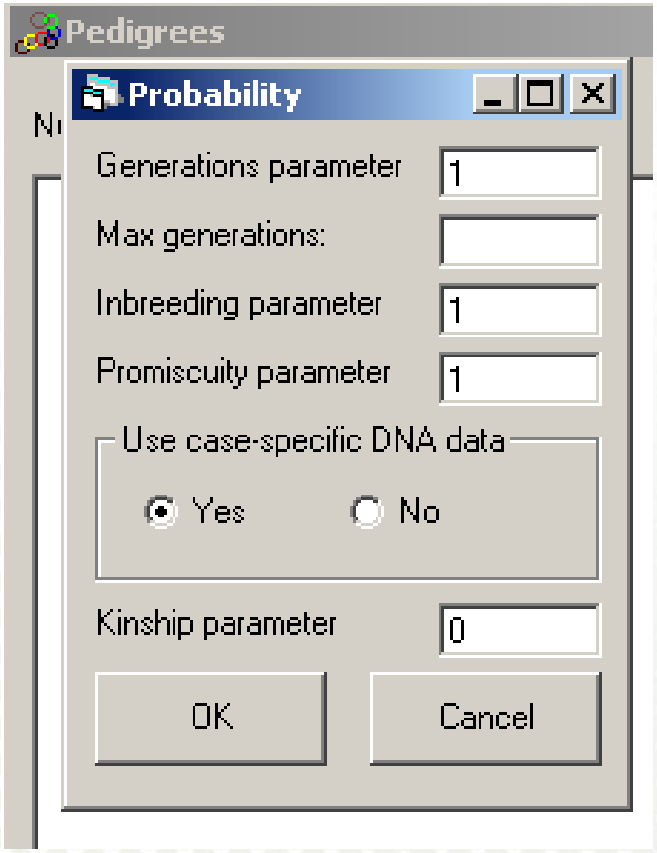
Homozygous A,A: $\theta p_A + p_A^2(1-\theta),$

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

Special cases:

θ=0

θ=1

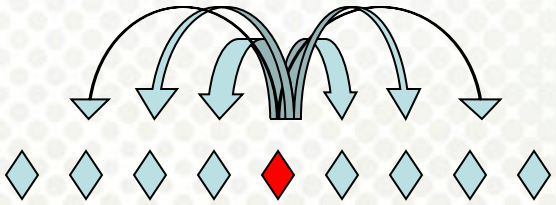


θ

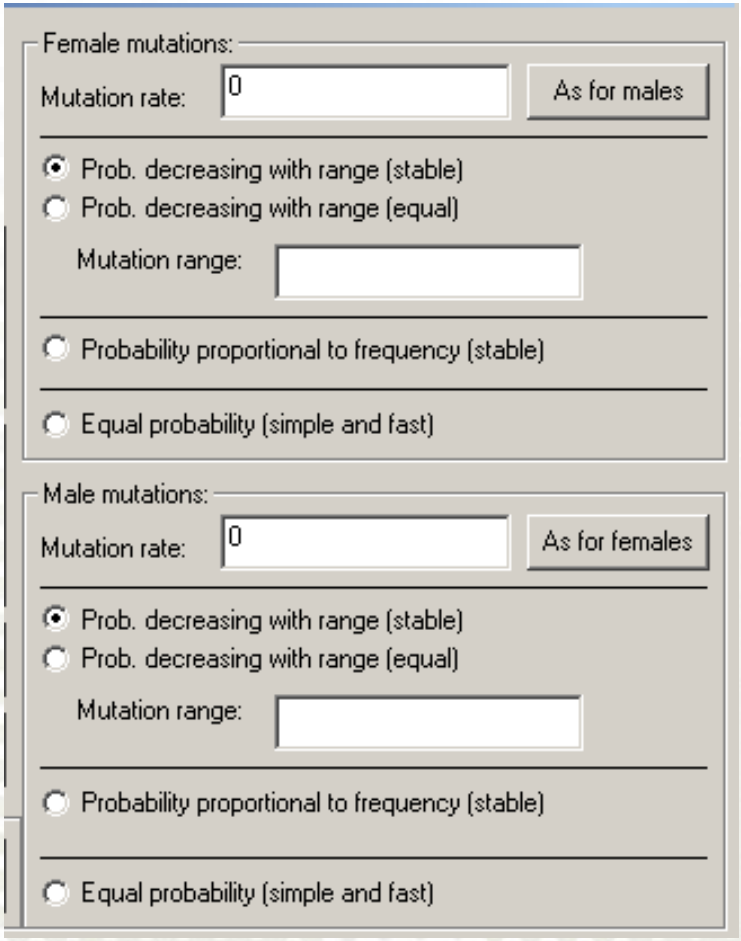
Revisit Exercise S1h (plenum)

Modelling mutations

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



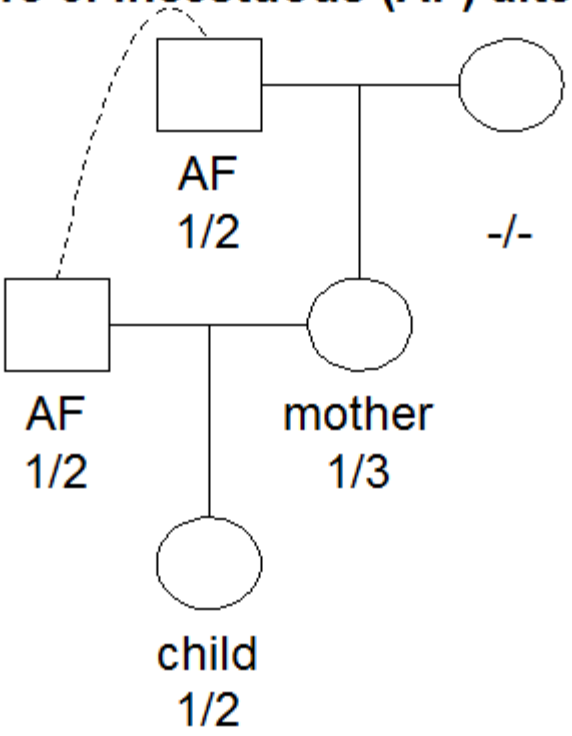
Mutations



Revisit Exercise S2, S7

Complex pedigrees I. Exercise S6

Figure 5. Incestuous (AF) alternative



Complex pedigrees II. Exercise S6

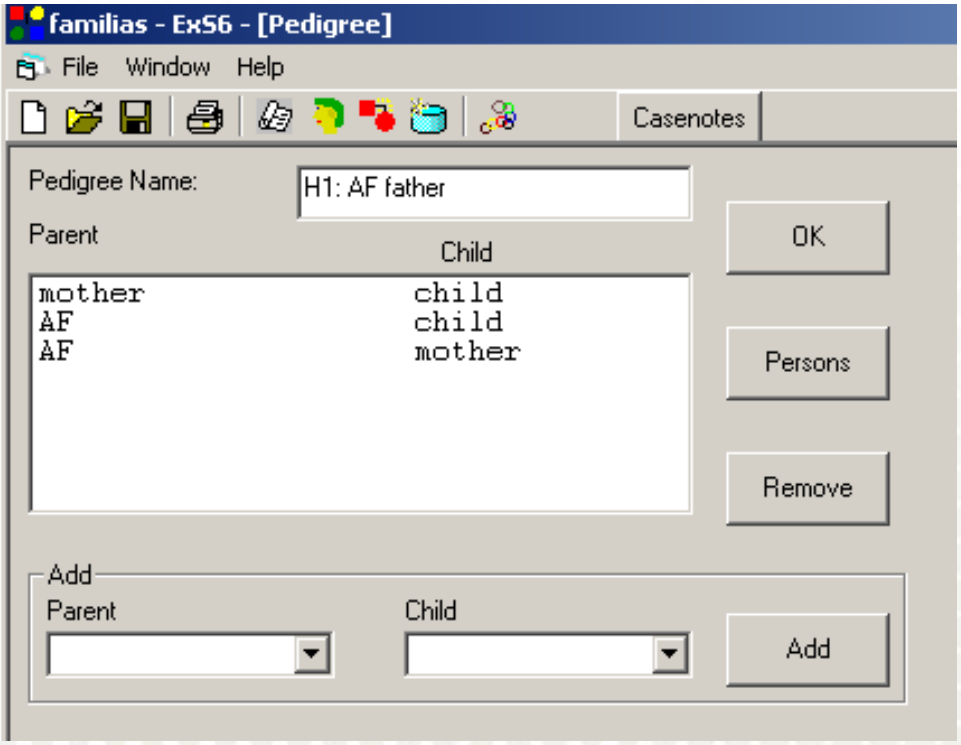
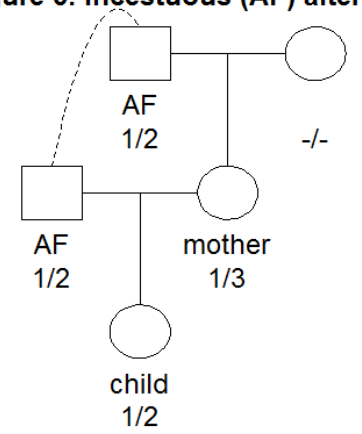
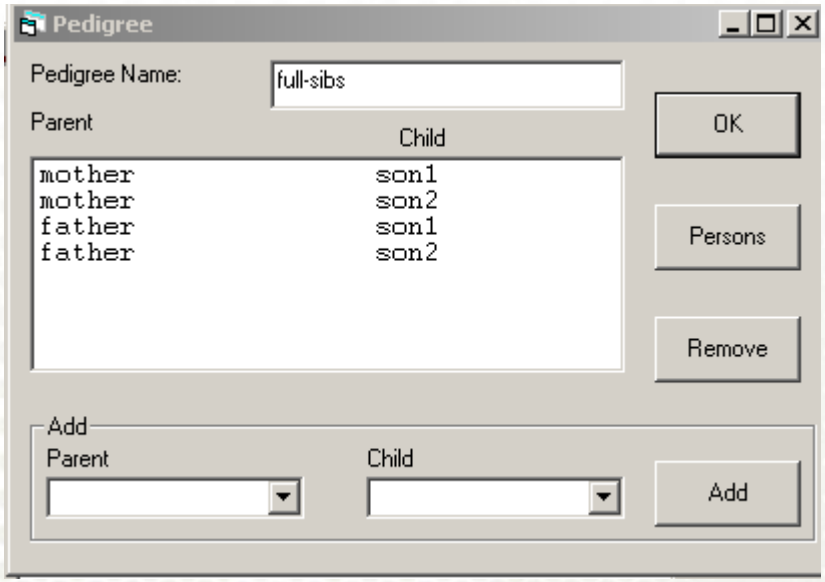
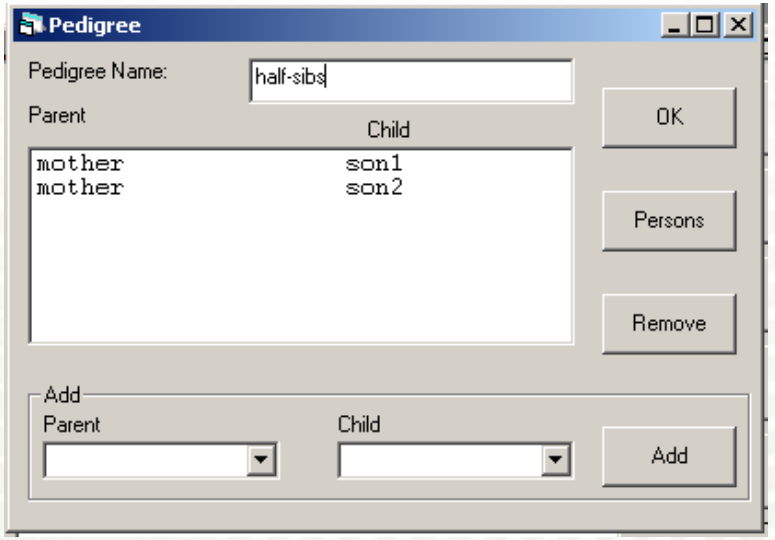


Figure 5. Incestuous (AF) alternative



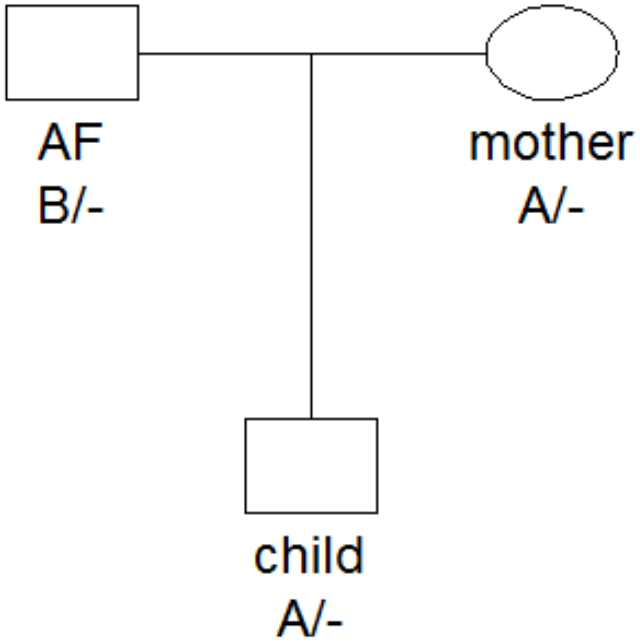
Difference between full-sibs and half sibs





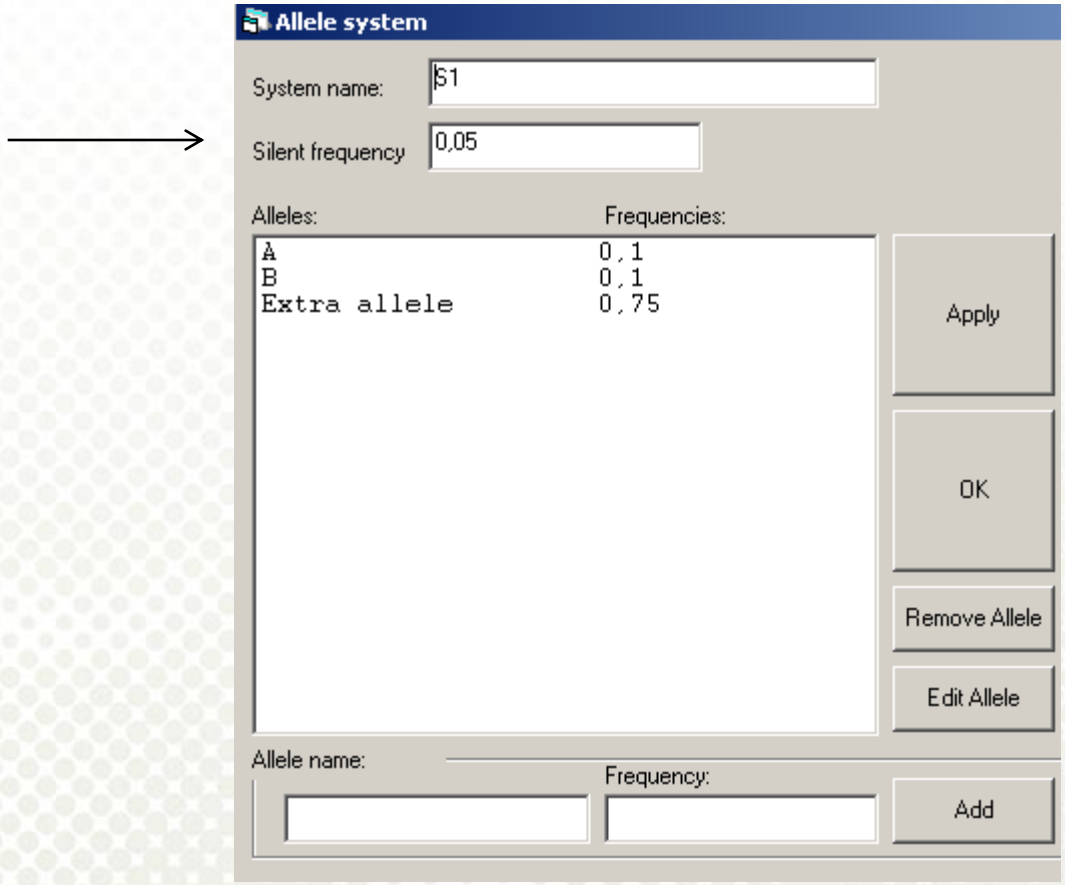
Silent alleles I. Exercise S11

Figure 5. Silent allele?



Silent alleles II. Familias

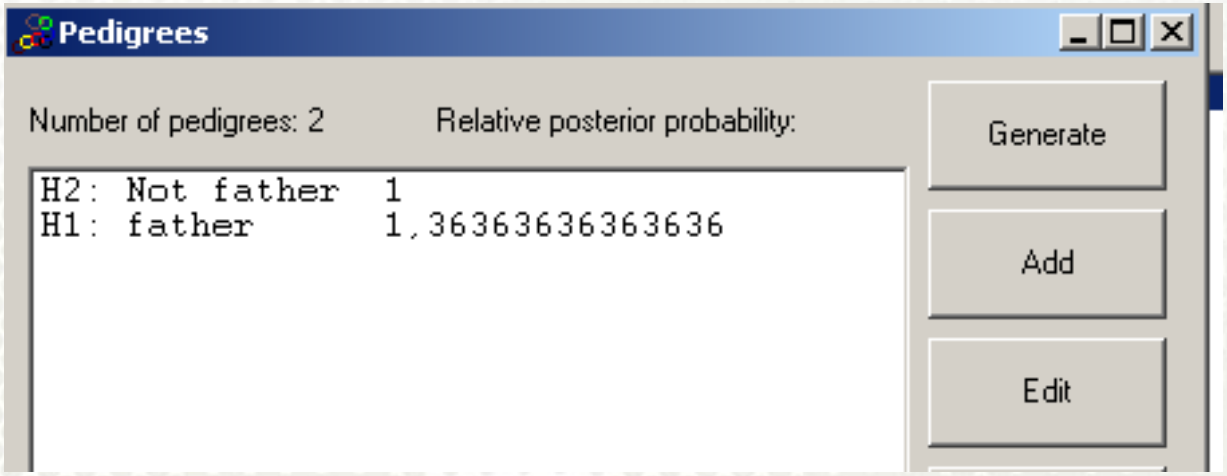
Silent allele



Silent alleles III. Exercises S11 (S12 theoretical)

$$LR = \frac{p_s\left(p_A + p_s\right)}{\left(p_A + p_s\right)^2\left(p_B + 2p_s\right) + p_s p_A\left(p_B + 2p_s\right)}$$

$$= \frac{0.05 \cdot 0.15}{0.15^2 \cdot 0.2 + 0.05 \cdot 0.1 \cdot 0.2} = 1.36.$$



Part VI. 15:30-17:00

- Familias exercises: S4, S6, S9, S10, S11
  If time: 'Extra' exercises on homepage:
  arken.umb.no/~theg/alcala/

Discussion 17-18

- Discussion of exercises.
- Spanish summary.
- Questions?
- Exercises S1h, S4c, S5d, S9d

Day 2, Oct 9

Part VII. 09:00-10:45

- Mutations:
  - Exercise S2,S7,S8 discussed, see solutions
  - Two challenges:
    - Unstable models
    - LR may differ depending on whether a *minimum database* (i.e., only alleles appearing for marker are named) is used or not
- *Unseen alleles*:
  - What allele frequency should we assign new alleles?

Unstable models I

Exercise S7f

Comment:

• Model 1 (Probability decreasing with range (stable)) and model 3 (Probability proportional to frequency) and 4 are *stable* whereas model 2 and 4 are *unstable.*

• If a model is stable, introducing a new untyped person, say the father of the alleged father, does not change the *LR*. This is a reasonable property of a model as introducing irrelevant information should not change the result.

For an unstable model, however, the *LR* will change slightly**.**

• Why?

Familias 3, new mutation model 10:45-11:15 **(Daniel Kling).**

Part VIII. 11:45-12:45

- Familias exercises

Part IX and X. 12:45-13:30

- Theta correction. Silent alleles.
- Exercises in plenum: S12, S13.

mtDNA and Y. Lourdes. 13:30-14:00

Part XI. 15:30-17:00

- The Bayesian alternative. Linked markers.
  - (Autosomal and X)
  - **Daniel**

Part XII. 17:00-17:30

- Demo of FamLink software, freely available from http://famlink.se/ . **Daniel**

17:30-18:00

- Summary

References

- Familias reference:
  Egeland, Mostad et al., see also
http://familias.name/