




 EVROPSKÁ UNIE **esf** MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY **MFT** **OP Vzdělávání** **KASSG**


 INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ


Vzdělávání v oblasti forenzní genetiky
reg. č. CZ.1.07/2.3.00/09.0080


INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Tento projekt je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.



Familias: A recently upgraded program for determining familial relationships


 Petter Mostad
 Chalmers University, Gothenburg, Sweden
 mostad@chalmers.se



 Universitetet i Oslo

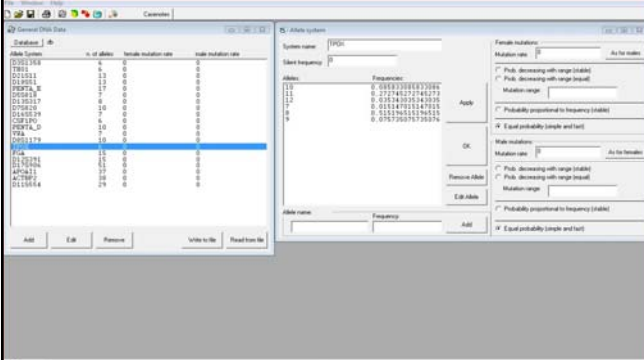

Overview

- Inference of familial relationships with the Familias program
- Newly added features (new homepage etc)
- Using Familias with multiple hypotheses
- Some lesser known features:
 - Generation of pedigrees
 - Priors
 - Etc.
- Future developments: R-package etc



The familias program

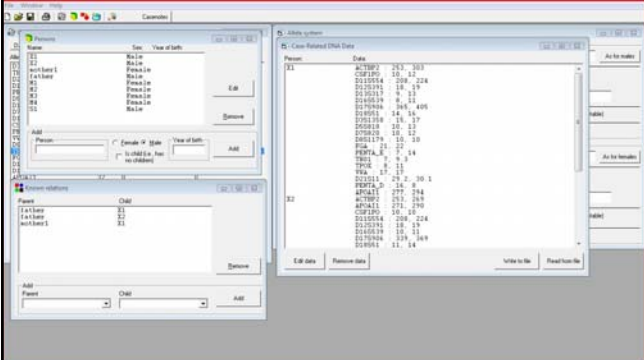
- Project initiated around 1992: computation of probabilities relating to inferring family relations
 - RMI – Forensic medicine in Oslo, Bjørnar Olaisen, Margurethe Stenersen, Berit Mevåg.
 - Thore Egeland & myself at the Norwegian Computing Center
- Released free on the internet 1995
- Validated by Jiří Drábek (2008)


Familias user interface



Marker	no. of alleles	autosomal mutation rate	Y-DNA mutation rate
D1S11	8	0.000000	0.000000
D2S1338	10	0.000000	0.000000
D3S1358	10	0.000000	0.000000
D5S818	8	0.000000	0.000000
D7S822	8	0.000000	0.000000
D8S1179	10	0.000000	0.000000
D10S1248	10	0.000000	0.000000
D12S1719	10	0.000000	0.000000
D13S322	10	0.000000	0.000000
D15S657	10	0.000000	0.000000
D16S1131	10	0.000000	0.000000
D17S11	10	0.000000	0.000000
D18S51	10	0.000000	0.000000
D19S43	10	0.000000	0.000000
D21S11	10	0.000000	0.000000
D22S1044	10	0.000000	0.000000
D23S448	10	0.000000	0.000000
D24S241	10	0.000000	0.000000
D25S25	10	0.000000	0.000000
D27S1628	10	0.000000	0.000000
D28S142	10	0.000000	0.000000
D29S11	10	0.000000	0.000000
D30S11	10	0.000000	0.000000
D31S11	10	0.000000	0.000000
D32S11	10	0.000000	0.000000
D33S11	10	0.000000	0.000000
D34S11	10	0.000000	0.000000
D35S11	10	0.000000	0.000000
D36S11	10	0.000000	0.000000
D37S11	10	0.000000	0.000000
D38S11	10	0.000000	0.000000
D39S11	10	0.000000	0.000000
D40S11	10	0.000000	0.000000
D41S11	10	0.000000	0.000000
D42S11	10	0.000000	0.000000
D43S11	10	0.000000	0.000000
D44S11	10	0.000000	0.000000
D45S11	10	0.000000	0.000000
D46S11	10	0.000000	0.000000
D47S11	10	0.000000	0.000000
D48S11	10	0.000000	0.000000
D49S11	10	0.000000	0.000000
D50S11	10	0.000000	0.000000
D51S11	10	0.000000	0.000000
D52S11	10	0.000000	0.000000
D53S11	10	0.000000	0.000000
D54S11	10	0.000000	0.000000
D55S11	10	0.000000	0.000000
D56S11	10	0.000000	0.000000
D57S11	10	0.000000	0.000000
D58S11	10	0.000000	0.000000
D59S11	10	0.000000	0.000000
D60S11	10	0.000000	0.000000
D61S11	10	0.000000	0.000000
D62S11	10	0.000000	0.000000
D63S11	10	0.000000	0.000000
D64S11	10	0.000000	0.000000
D65S11	10	0.000000	0.000000
D66S11	10	0.000000	0.000000
D67S11	10	0.000000	0.000000
D68S11	10	0.000000	0.000000
D69S11	10	0.000000	0.000000
D70S11	10	0.000000	0.000000
D71S11	10	0.000000	0.000000
D72S11	10	0.000000	0.000000
D73S11	10	0.000000	0.000000
D74S11	10	0.000000	0.000000
D75S11	10	0.000000	0.000000
D76S11	10	0.000000	0.000000
D77S11	10	0.000000	0.000000
D78S11	10	0.000000	0.000000
D79S11	10	0.000000	0.000000
D80S11	10	0.000000	0.000000
D81S11	10	0.000000	0.000000
D82S11	10	0.000000	0.000000
D83S11	10	0.000000	0.000000
D84S11	10	0.000000	0.000000
D85S11	10	0.000000	0.000000
D86S11	10	0.000000	0.000000
D87S11	10	0.000000	0.000000
D88S11	10	0.000000	0.000000
D89S11	10	0.000000	0.000000
D90S11	10	0.000000	0.000000
D91S11	10	0.000000	0.000000
D92S11	10	0.000000	0.000000
D93S11	10	0.000000	0.000000
D94S11	10	0.000000	0.000000
D95S11	10	0.000000	0.000000
D96S11	10	0.000000	0.000000
D97S11	10	0.000000	0.000000
D98S11	10	0.000000	0.000000
D99S11	10	0.000000	0.000000
D100S11	10	0.000000	0.000000


Familias user interface, cont.



esf Familias user interface, cont.

The screenshot shows the Familias software interface. On the left, there is a 'Pedigree' list with columns for 'Pedigree Name' and 'File'. Below it are buttons for 'Add', 'Edit', 'Remove', 'Remove All', 'LR', 'Probability', 'Options', 'Help', and 'Exit'. In the center, a 'Case Related Data' window is open, displaying a table with columns for 'Pedigree Name', 'File', 'Date', 'Time', 'User', and 'Version'. The table contains several rows of data. On the right, there are 'Add' and 'Remove' buttons for the data table.

esf Some Familias features

- General pedigrees
- Computations of likelihoods, LR, posterior probabilities.
- Multiple hypotheses
- Different mutation models
- Kinship
- Silent alleles
- Automatic pedigree generation
- Simple Windows user interface

The screenshot shows the top part of the Familias software interface, including the menu bar with 'File', 'Window', and 'Help' options, and a toolbar with various icons for file operations and analysis.

esf Examples of uses

- "Deficient" paternity cases, where data is available only from more distant relatives
- Disaster identification (Olaisen 1996)
- Immigration cases (Tillmar 2007)
- Animal breeding
- ...and probably many uses we don't know about

esf Largest user problem: Defining pedigrees correctly

- Example: Identifying remains of man based on tests of his mother and his children
- In alternative pedigree, an alternative missing man must take his place

The pedigree diagram shows a family structure. At the top, a grey square (man) and a red circle (woman) are connected by a horizontal line. A vertical line descends from the man, and another from the woman, meeting at a horizontal line. From this junction, a vertical line descends to a green square (man) and a grey circle (woman). Below them, another horizontal line connects a red square (man) and a red circle (woman). This represents a man, his mother, and his children.

Tento projekt je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.

esf The recent familias upgrade

- Improved input and output of data formats (Excel-type)
- Improved presentation of results, and improved design of some forms
- Ongoing improvements in new homepage www.familias.name
- Improved stability
 - Two bugs in user interface fixed
- Larger extensions has been moved to a coming R implementation

esf Two or more hypotheses?

- Bayesian philosophical view: Separation between DNA-based evidence and non-DNA evidence.
- There may be *several* different scenarios, each with its *separate* probability for observing the given data.
- Each should be investigated and reported.

esf **Example: The brother of the alleged father is the father**

EVROPSKA UNIE

OP Vzdělávání pro konkurenceschopnost

ČSSFG

- The LR between a man and a "randomly selected man" in a certain population may be quite high.
- A close relative may be expected to fit almost equally well
- Case-dependent non-DNA evidence indicate whether close relatives should be considered

esf **Example: Immigration cases**

EVROPSKA UNIE

OP Vzdělávání pro konkurenceschopnost

ČSSFG

- Immigration rules (in Sweden) specify exactly which relationships a residence permit may be based on
- For some immigrants, biological parenting may be less important than social parenting
- For an immigrant family, a number of possible pedigrees may have to be investigated

esf **Immigration cases**

EVROPSKA UNIE

OP Vzdělávání pro konkurenceschopnost

ČSSFG

Tillmar et al (2007) investigated the risk of erroneous conclusions in immigration cases, using simulation (and Familias)

Cutoff (%)	Exclusion error rate (%)	Total error rate (%)
90	~1.5	~1.5
95	~1.2	~1.2
99	~1.0	~1.0
99.9	~1.5	~1.5
99.99	~5.0	~5.0
99.999	10.0	10.0

- Using mutations in computations lower error rates somewhat
- Improper databases increases rates slightly (but case LR values may be quite influenced)
- Including an investigation of a standard set of 5 hypotheses decreases rates slightly

esf **Example: Large number of possible pedigrees**

EVROPSKA UNIE

OP Vzdělávání pro konkurenceschopnost

ČSSFG

- "Given DNA-data on a set of persons, what is their relationship?"
- A set of *possible* pedigrees must be specified (and given prior probabilities)
- Familias can generate the set of possible pedigrees, given some constraints.
- A crude type of prior can also be computed.

esf **Example: The Romanov family**

EVROPSKA UNIE

OP Vzdělávání pro konkurenceschopnost

ČSSFG

- The Romanov family (parents, five children, and their doctor) disappeared in connection with the Russian revolution.
- The remains in a grave was later shown to contain the family, by using mtDNA and STR markers (Gill et al, 1994)
- In Egeland et al (2000) it was shown how to use Familias to combinatorially find the right pedigree.

Tento projekt je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.


esf **An example extending Familias**

EVROPSKA UNIE

OP Vzdělávání pro konkurenceschopnost





ČSSFG

- When using non-zero mutation rates, computations with Familias may become too slow on large pedigrees with "loops".
- If mutations are disregarded, some loops may be replaced by simple conditional probabilities of whether or not alleles are IBD
- Computations may be implemented by combining Familias with some external calculations, for example in Excel.




Example, continued.

- Case: Do two men share the same father or not? Their mothers are cousins once removed.
- Solution: condition on whether paternal and/or maternal alleles are IBD. Sum and multiply over all autosomal markers
- Can be extended to similar situations
- BUT: This also illustrates the limitations of Familias and its user interface










Tento projekt je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.




Future extensions of Familias planned in new framework

- We are working on an implementation of the core functions of Familias in an R package: OpenFamilias
- This will make future extensions much faster and simpler to do
- More flexible for advanced users!
- A bit more difficult for novices

Tento projekt je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.







Some current work which may be implemented this way

- Computational methods for X chromosome markers in LD (Tillmar, submitted)
- Computational models for SNP data
- Evaluation of STR mixture data

Future work:

- Support for multiple databases
- Inclusion of Y data
- Inclusion of mtDNA data
- Simulations for power computations
- Missing persons evaluation in catastrophe scenarios

Tento projekt je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.




Acknowledgements

- Oslo, Norway: Thore Egeland, Öyvind Skare, Margurethe Stenersen, Bente Mevåg, Björn Olaisen, ...
- Linköping, Sweden: Gunilla Holmlund, Andreas Tillmar, Bertil Lindblom, Jenny Velander, Daniel Kling, ...
- Göteborg, Sweden (where I am): Iftikhar Ahmad, Qi Cao, ...










Tento projekt je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.



References

- Gill et al (1994): Identification of the remains of the Romanov family by DNA analysis (Nat.Genet.)
- Olaisen (1996): Identification by DNA analysis of the victims of the August 1996 Spitsbergen civil aircraft disaster (Nature Genetics)
- Egeland et al (2000): Beyond traditional paternity and identification cases: Selecting the most probable pedigree (FSI)
- Tillmar (2007): DNA testing for immigration cases: The risk of erroneous conclusions (FSI)
- Jiří Drábek (2008): Validation of software for calculating the likelihood ratio for parentage and kinship (FSI)

Tento projekt je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.