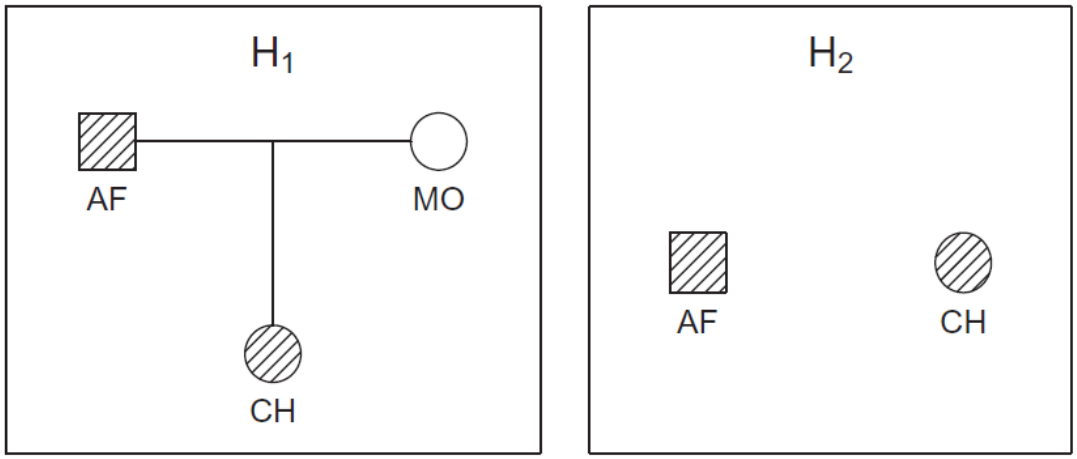
Kinship testing



Thore Egeland
Norwegian University of Life Sciences &
Department of Forensic Medicine, Norway

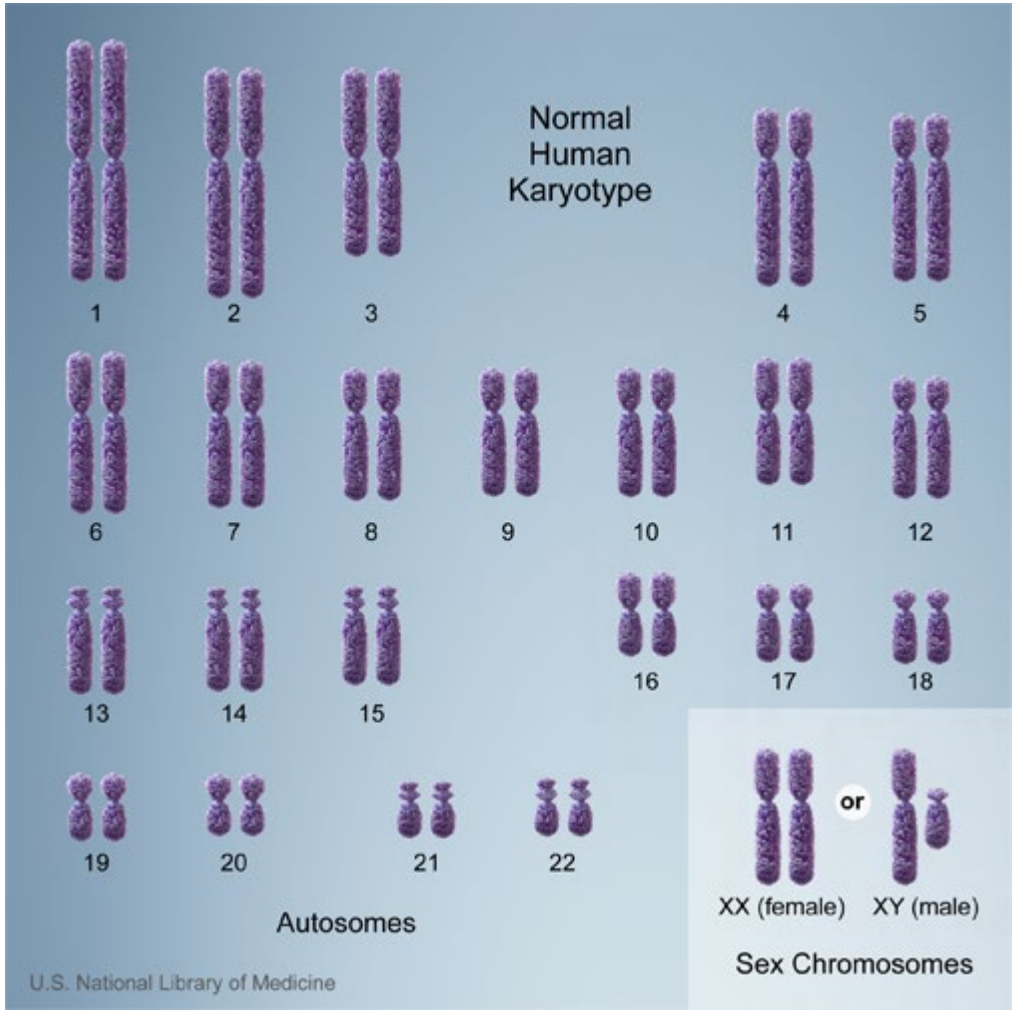
Motivating examples

- Kinship testing
  - Close (paternity) or distant (second cousins)
  - Disaster victim identification (DVI)
  - Pedigree reconstruction
  - …

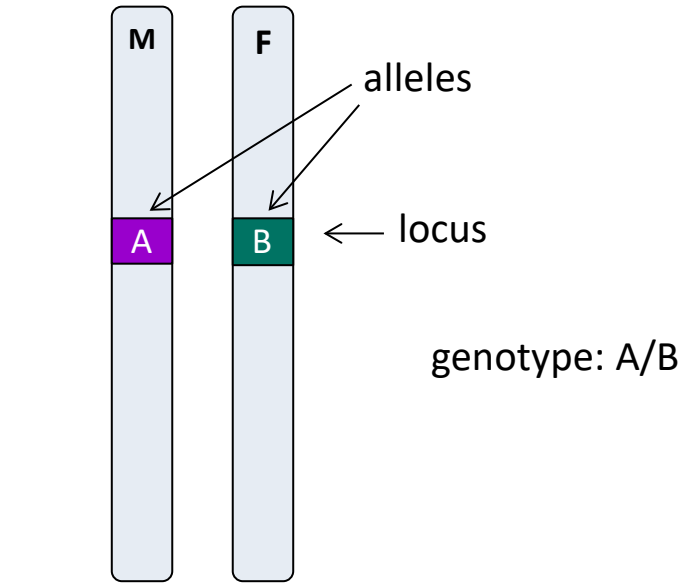
- We distinguish between
  - *kinship testing*, current topic, where a specific set of alternatives are compared, and
  - *relatedness inference* aiming to find the most probable relationship without restrictions

Genetics terminology

- Locus
- Allele
- Genotype
- Genetic markers
  - SNPs
  - microsatellites



Locus, allele, genotype



- **LOCUS** = a specific place in the genome
- **ALLELE** = any of the alternative forms of a locus
- **GENOTYPE** = the set (usually: pair) of alleles carried at a given locus

Genetic markers

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

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

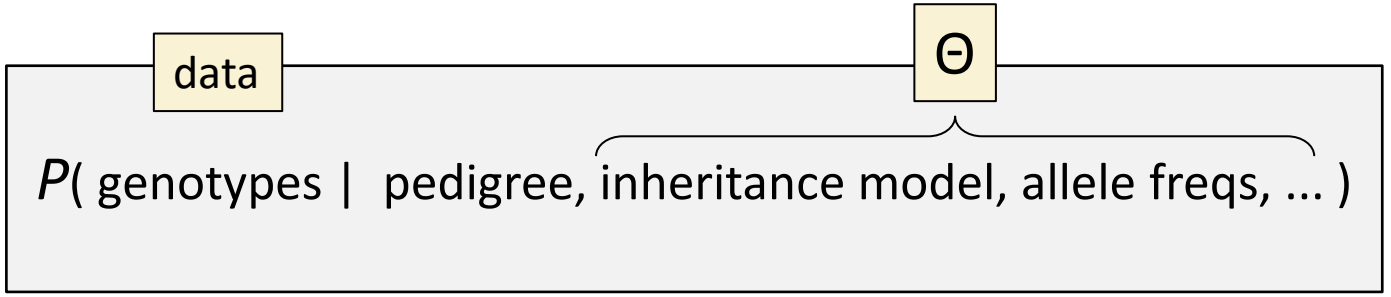
...CCGTTATATGGGC...
...CCGTTAGATGGGC...
...CCGTTATATGGGC...
...CCGTTATATGGGC...
...CCGTTAGATGGGC...

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

...ACG TTAG TTAG TTAG TTAG AAC..
...ACG TTAG TTAG AAC..
...ACG TTAG TTAG TTAG TTAG TTAG AAC..

Pedigree likelihoods

- Many applications involve probabilities of the following form



- Often referred to as a *pedigree likelihood*:

$$L(\text{ pedigree} \mid \text{data}) = P(\text{ data} \mid \text{pedigree}, \Theta)$$

Three principles of interpretation

1. It is necessary to consider at least one alternative proposition

2. Interpretation is based on questions like:
   «What is the probability of the evidence given the proposition»

3. We condition on proposition and known circumstances

- These principles lead to the likelihood ratio (LR):

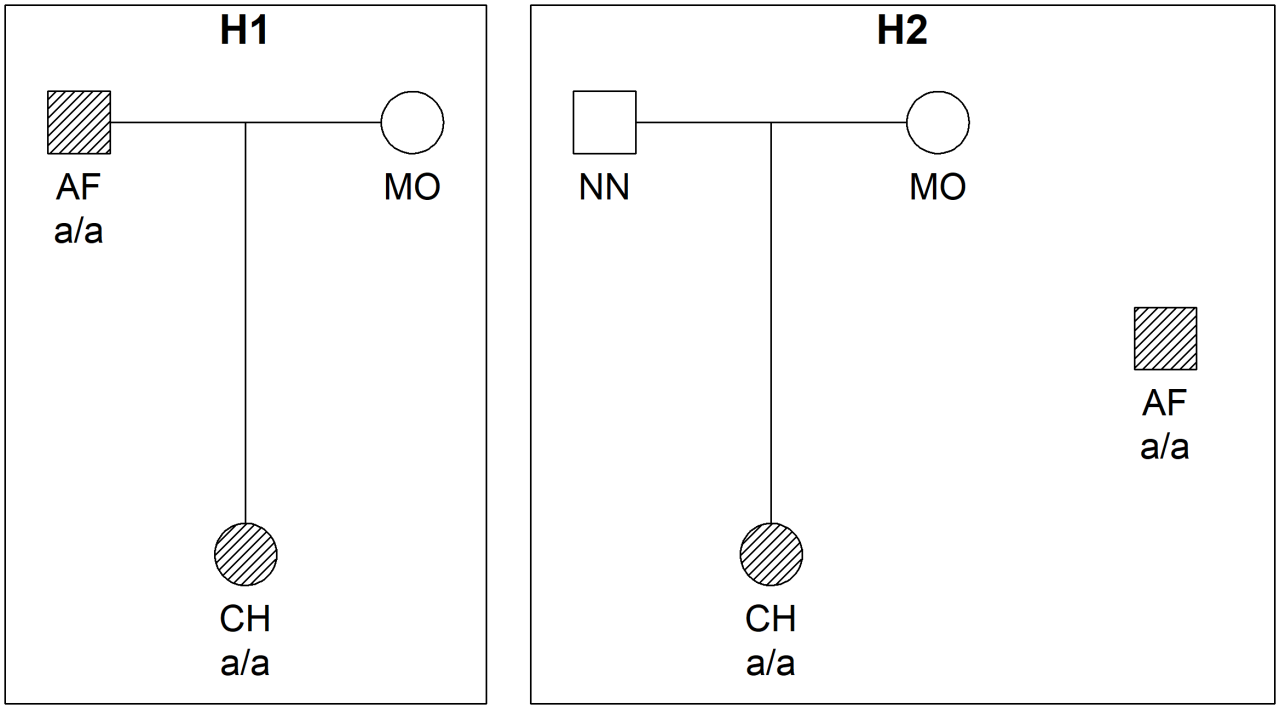
The Likelihood Ratio

- H_1: The individuals are related according to some pedigree $\mathcal{P}_1$.
- H_2: The individuals are related according to a different pedigree $\mathcal{P}_2$.

$$\mathrm{LR} = \frac{P(\text{data} \mid H_1, \Theta)}{P(\text{data} \mid H_2, \Theta)}.$$

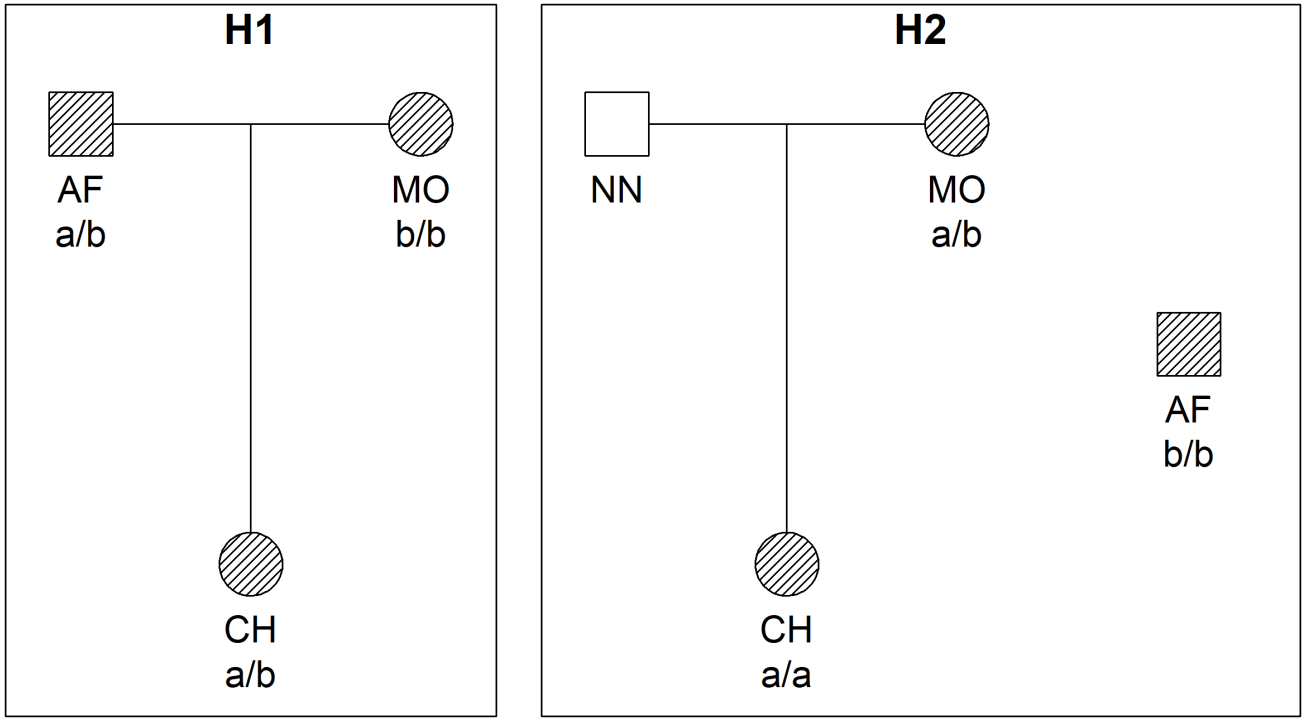
- data: available genotypes
- Θ : fixed model parameters, common to both hypotheses
- Interpretation:
  - The LR measures how well H_1 explains the data compared to H_2
- Default assumptions:
  - ✓ Hardy Weinberg Equilibrium
  - ✓ No mutations
  - ✓ No artefacts (drop out, drop in, genotyping error)
  - ✓ Independence between markers

Example 1: Paternity case



$$LR_1 = \frac{P(\mathrm{AF} = a/a, \mathrm{CH} = a/a \mid H_1)}{P(\mathrm{AF} = a/a, \mathrm{CH} = a/a \mid H_2)} = \frac{p_a^2 \cdot p_a}{p_a^2 \cdot p_a^2} = \frac{1}{p_a}.$$

Mother genotyped



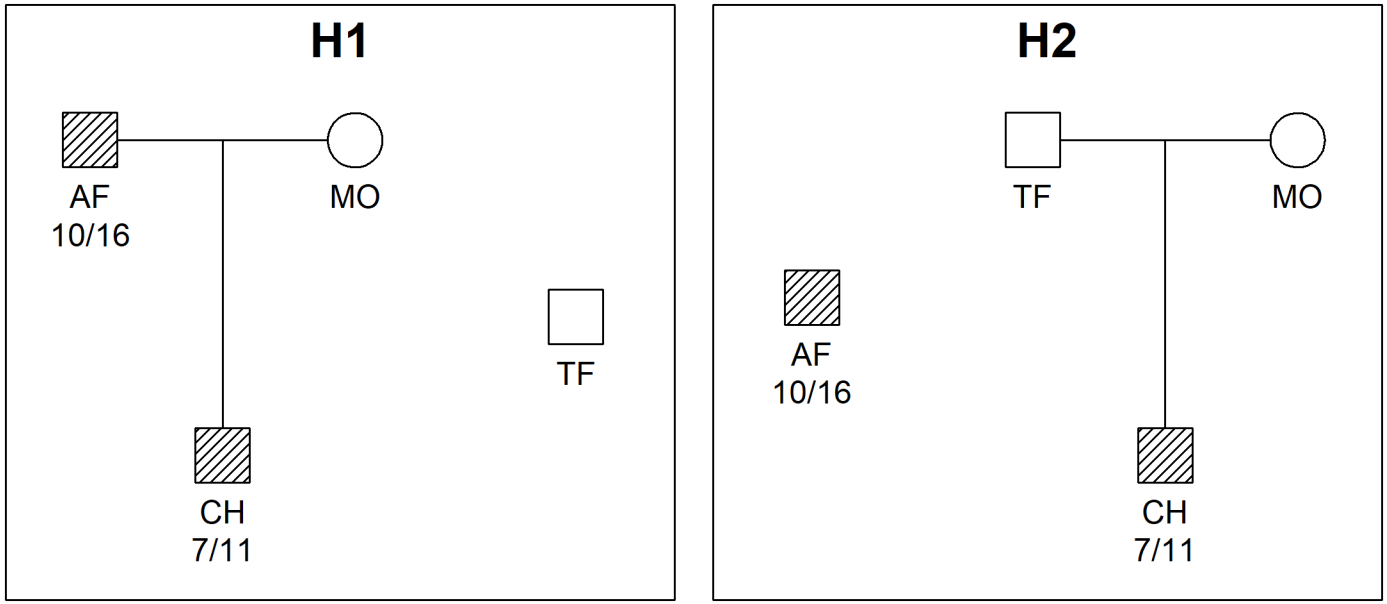
$$LR_2 = \frac{P(\mathrm{AF} = a/b, \mathrm{MO} = b/b, \mathrm{CH} = a/b \mid H_1)}{P(\mathrm{AF} = a/b, \mathrm{MO} = b/b, \mathrm{CH} = a/b \mid H_2)} = \frac{2p_a p_b \cdot p_b^2 \cdot \frac{1}{2}}{2p_a p_b \cdot p_b^2 \cdot p_a} = \frac{1}{2p_a}.$$

Combined LR

- Assume $p_a = 0.05$ for both markers:
  - $LR_1 = \frac{1}{p_a} = 20$
  - $LR_2 = \frac{1}{2p_a} = 10$
- Assuming independence:
  - $LR = LR_1 \cdot LR_2 = 20 \cdot 10 = 200$

- Interpretation:
  The data is 200 times more likely if we assume H_1 to be true rather than H_2

Mutation?

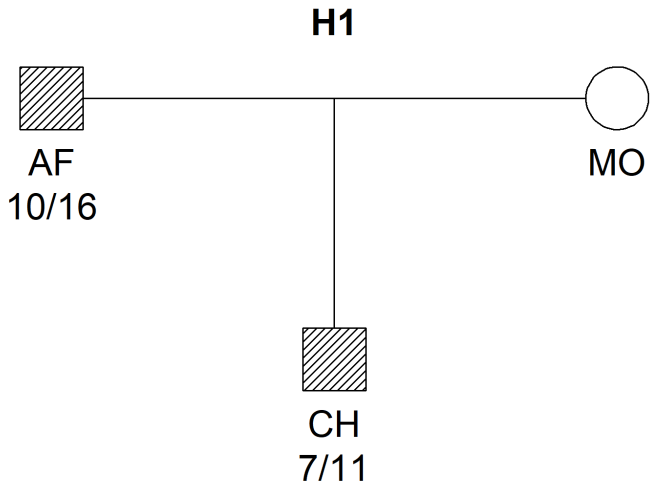


Mutation models

- Mutation rates higher in males.
- Short mutations more likely: One step mutation more likely than two steps and so on.
- Mutation rates:

2024 AABB Relationship Testing Technical Report

Dealing with mutations



Strategies for handling mutations

- Exclude inconsistent markers from the analysis. Not recommended
- Apply mutation modelling only to inconsistent markers
- Apply mutation modelling to *all* markers. Recommended

Kinship testing with linked markers

Forensic Science International: Genetics 86 (2027) 103578

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsigen

Research Paper

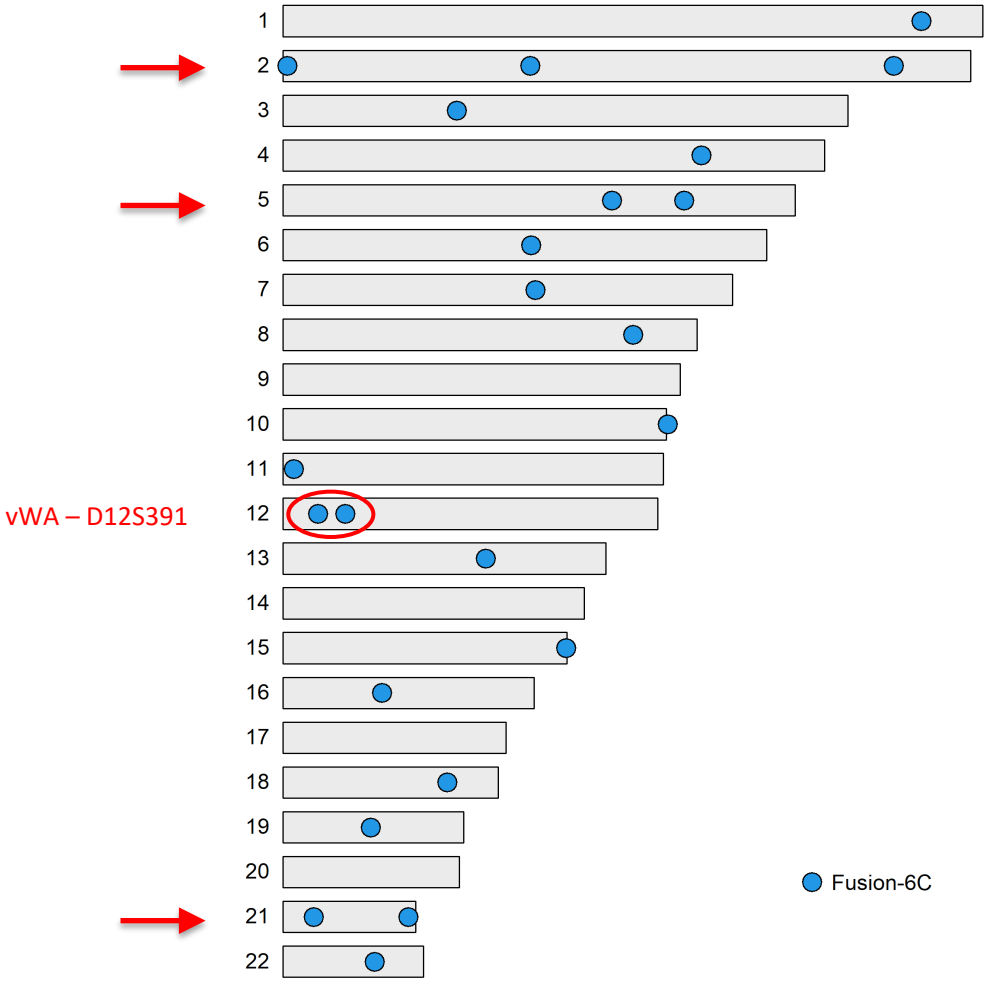
KLINK: A program for kinship testing with pairwise linked STR markers

Magnus D. Vigeland *, Siv Gilfillan

Department of Forensic Sciences, Oslo University Hospital, Oslo, 0424, Norway

PowerPlex® Fusion 6C

- 27 autosomal STR markers
- 1 close pair (< 25 cM)
- 4 linkage groups

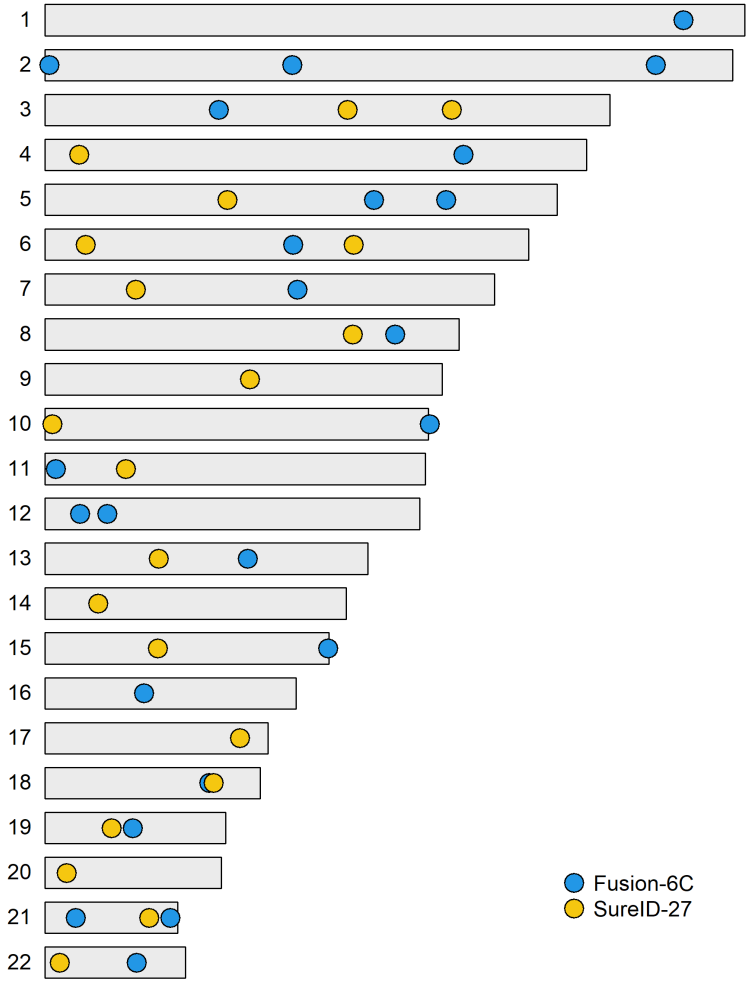


PowerPlex® Fusion 6C

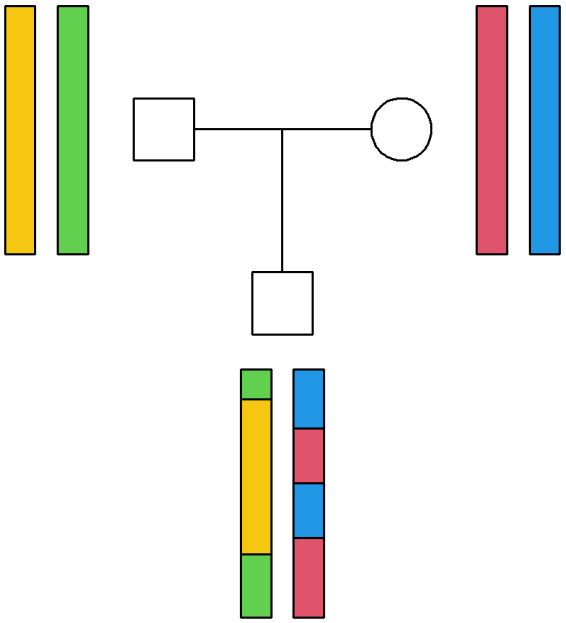
- 27 autosomal STR markers
- 1 close pair (< 25 cM)
- 4 linkage groups

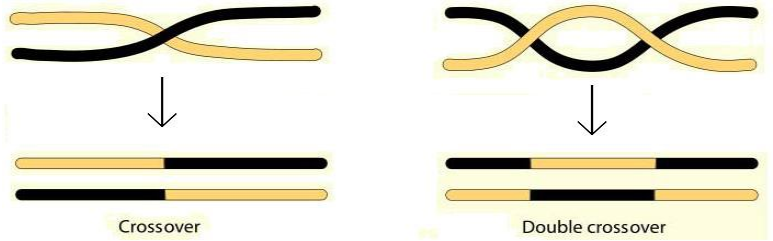
Fusion 6C + SureID® 27

- 49 markers
- 6 close pairs (< 25 cM)
- 16 linkage groups



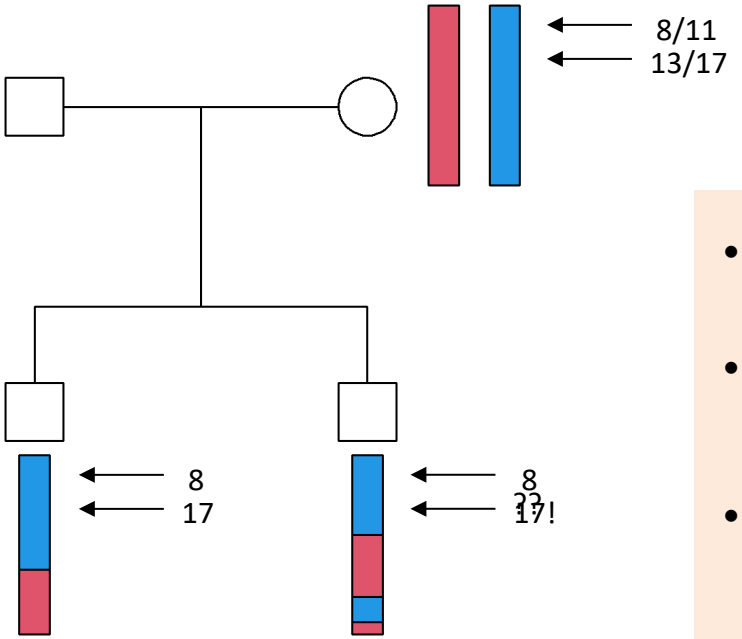
Meiotic recombination





- Genetic distance between two loci:
  = average crossovers per meiosis
- Units:
  - 1 Morgan (M) = 1 crossover per meiosis
  - 1 centiMorgan (cM) = 0.01 M
- The human genome: Ca 30 Morgan

Why do we care about linkage?



- Linkage → dependent inheritance
- Irrelevant in paternity cases (duo/trio)
  (unless *phased* data)
- Relevant if ≥ 2 informative, *non-independent* meioses

Software handling linked markers

(No linkage)

.fam ← Familias.no

MERLIN

Pros
- **Gold standard** for linkage
- Arbitrary linkage

Cons
- Command-line based
- Max 30 alleles
- No mutation modelling

FamLink2

Pros
- Designed for forensics
- Good with SNPs
- GUI + many features
- *Quick analysis* mode

Cons
- No mut. models in QA

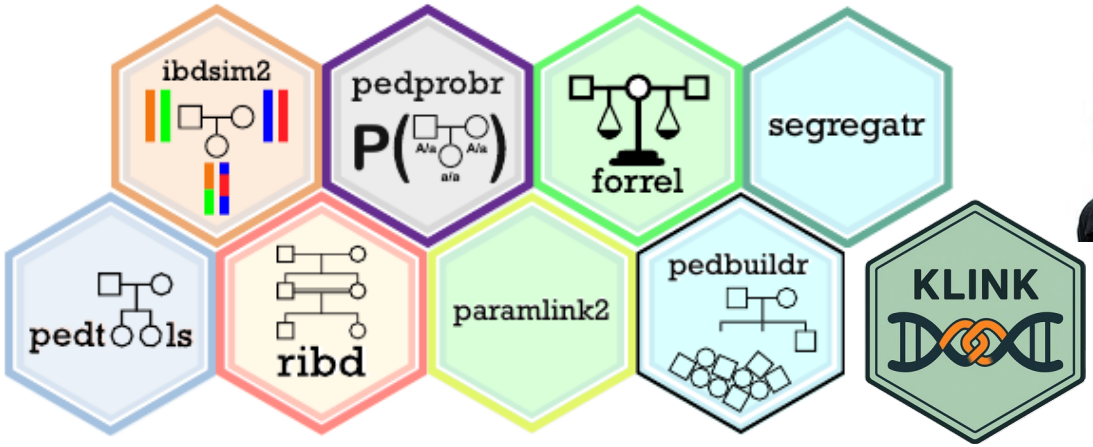
KLINK

Pros
- Designed for kinship STRs
- GUI with visualisations
- Any mutation models
- Tailor-made Excel reports

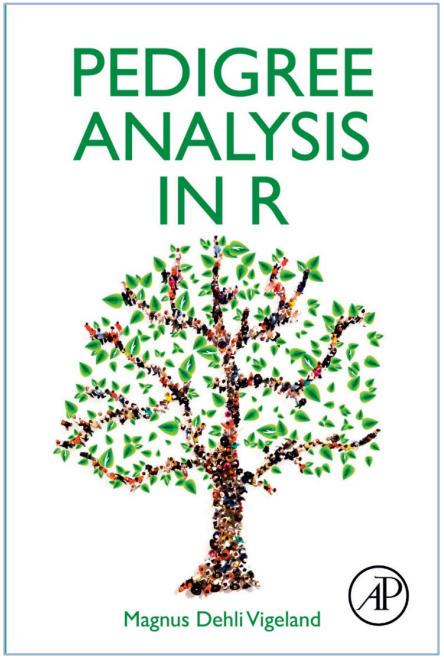
Cons
- Only pairwise linkage

Home page: https://magnusdv.github.io/pedsuite
Source code: https://github.com/magnusdv

The **pedsuite**:
A collection of packages for pedigree analysis in R



Siv Gilfillan

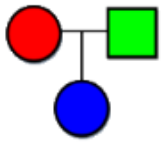


KLINK is available in two ways:

- From R: `KLINK::launchApp()`
- Online: https://magnusdv.shinyapps.io/klink/
- Input. Familias file:

Familias

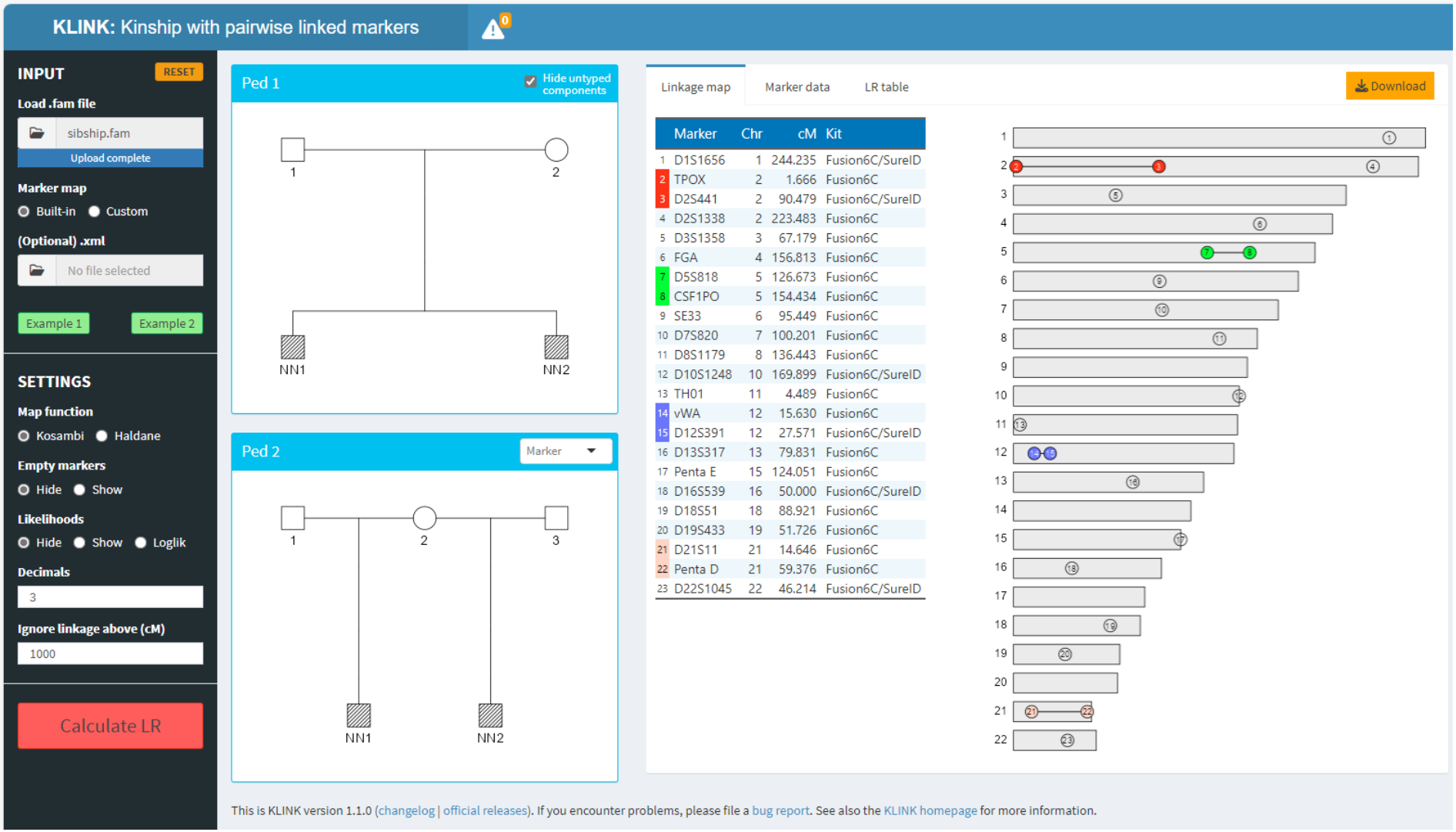
- New homepage merging familias.name into https://familias.no/
- New maintainer, developer: Franco Marsico



Familias - Tutorial

By: Daniel Kling, Lourdes Prieto,
Franco Marsico and Thore Egeland

Petter Mostad


KLINK: Kinship with pairwise linked markers
0
INPUT
RESET
Load .fam file
sibship.fam
Upload complete
Marker map
Built-in
Custom
(Optional) .xml
No file selected
Example 1
Example 2
SETTINGS
Map function
Kosambi
Haldane
Empty markers
Hide
Show
Likelihoods
Hide
Show
Loglik
Decimals
3
Ignore linkage above (cM)
1000
Calculate LR
Ped 1
Hide untyped components
1
2
NN1
NN2
Ped 2
Marker
1
2
3
NN1
NN2
Linkage map
Marker data
LR table
Download
Marker Chr cM Kit
1 D1S1656 1 244.235 Fusion6C/SureID
2 TPOX 2 1.666 Fusion6C
3 D2S441 2 90.479 Fusion6C/SureID
4 D2S1338 2 223.483 Fusion6C
5 D3S1358 3 67.179 Fusion6C
6 FGA 4 156.813 Fusion6C
7 D5S818 5 126.673 Fusion6C
8 CSF1PO 5 154.434 Fusion6C
9 SE33 6 95.449 Fusion6C
10 D7S820 7 100.201 Fusion6C
11 D8S1179 8 136.443 Fusion6C
12 D10S1248 10 169.899 Fusion6C/SureID
13 TH01 11 4.489 Fusion6C
14 vWA 12 15.630 Fusion6C
15 D12S391 12 27.571 Fusion6C/SureID
16 D13S317 13 79.831 Fusion6C
17 Penta E 15 124.051 Fusion6C
18 D16S539 16 50.000 Fusion6C/SureID
19 D18S51 18 88.921 Fusion6C
20 D19S433 19 51.726 Fusion6C
21 D21S11 21 14.646 Fusion6C
22 Penta D 21 59.376 Fusion6C
23 D22S1045 22 46.214 Fusion6C/SureID
This is KLINK version 1.1.0 (changelog | official releases). If you encounter problems, please file a bug report. See also the KLINK homepage for more information.

KLINK: Kinship with pairwise linked markers

INPUT RESET

Load .fam file

sibship.fam

Upload complete

Marker map

Built-in Custom

(Optional) .xml

No file selected

Example 1 Example 2

SETTINGS

Map function

Kosambi Haldane

Empty markers

Hide Show

Likelihoods

Hide Show Loglik

Decimals

3

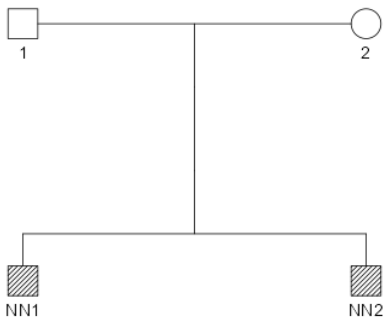
Ignore linkage above (cM)

1000

Calculate LR

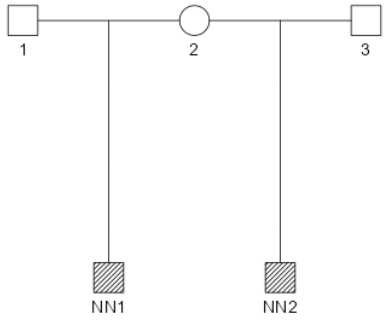
Ped 1

Hide untyped components



Ped 2

Marker

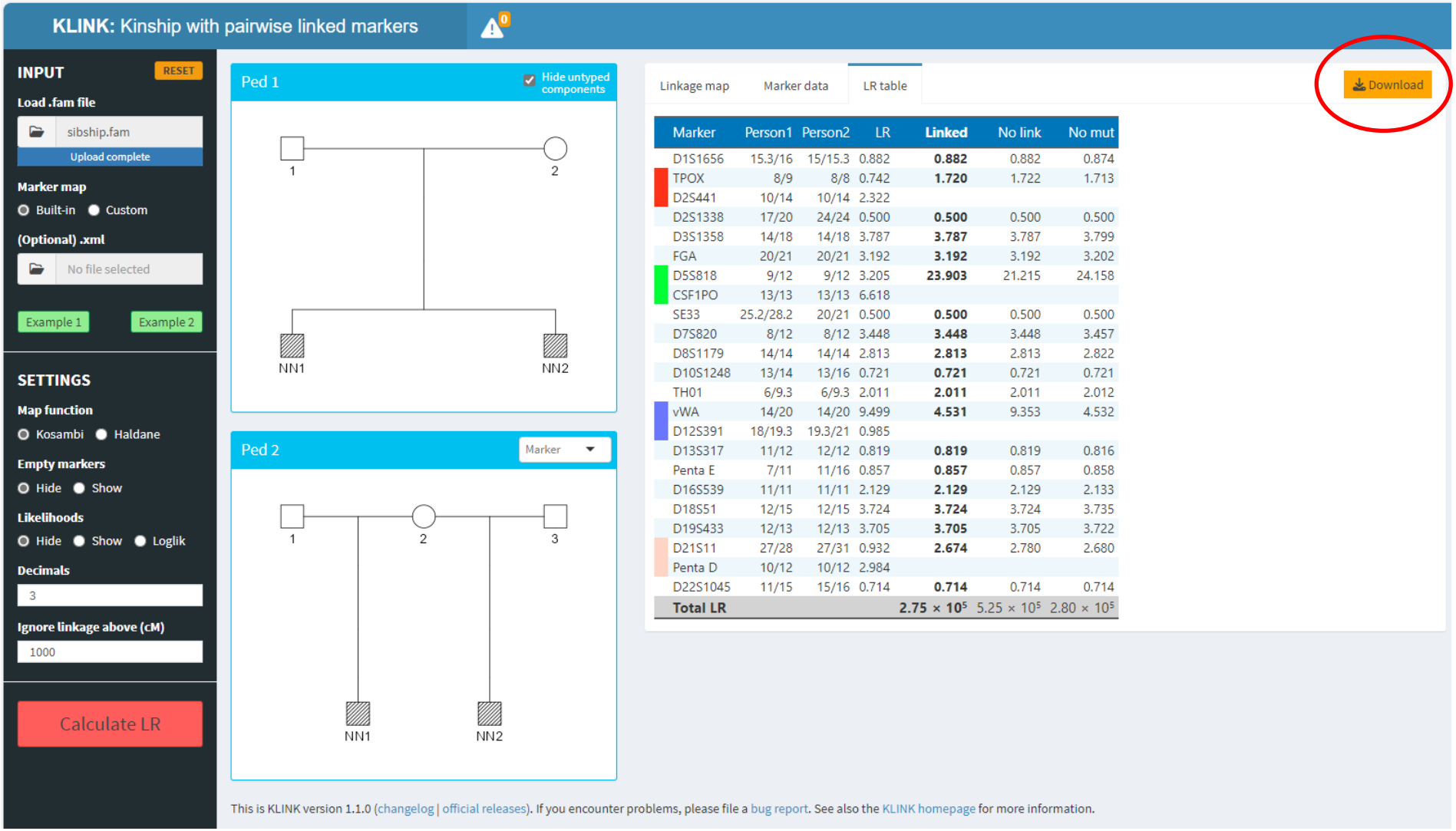


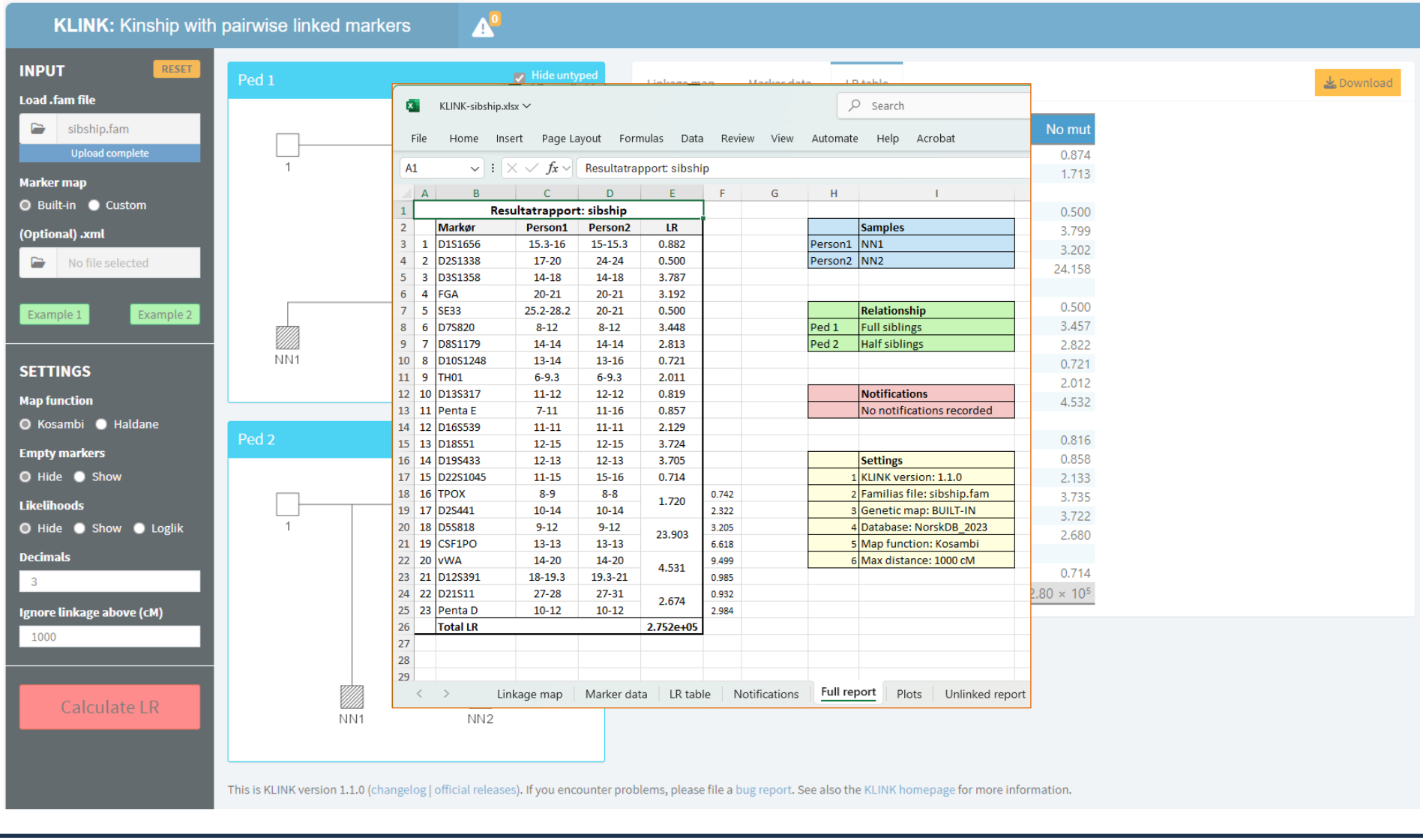
Linkage map | Marker data | LR table

Download

Marker	Person1	Person2	Alleles	PIC	MinFreq	Model	Rate	Rate2	Range	Stat	Lump
D1S1656	15.3/16	15/15.3	17	0.891	0.001	Stepwise	0.001/0.002	1e-06	0.1	No	No
TPOX	8/9	8/8	9	0.569	0.001	Stepwise	0.001/0.002	1e-06	0.1	No	No
D2S441	10/14	10/14	13	0.696	0.001	Stepwise	0.001/0.002	1e-06	0.1	No	No
D2S1338	17/20	24/24	13	0.863	0.001	Stepwise	0.001/0.002	1e-06	0.1	No	No
D3S1358	14/18	14/18	12	0.766	0.001	Stepwise	0.001/0.002	1e-06	0.1	No	No
FGA	20/21	20/21	25	0.855	0.001	Stepwise	0.001/0.002	1e-06	0.1	No	No
D5S818	9/12	9/12	9	0.645	0.001	Stepwise	0.001/0.002	1e-06	0.1	No	No
CSF1PO	13/13	13/13	11	0.688	0.001	Stepwise	0.001/0.002	1e-06	0.1	No	No
SE33	25.2/28.2	20/21	55	0.947	0.001	Stepwise	0.001/0.002	1e-06	0.1	No	No
D7S820	8/12	8/12	19	0.794	0.001	Stepwise	0.001/0.002	1e-06	0.1	No	No
D8S1179	14/14	14/14	12	0.779	0.001	Stepwise	0.001/0.002	1e-06	0.1	No	No
D10S1248	13/14	13/16	9	0.718	0.001	Stepwise	0.001/0.002	1e-06	0.1	No	No
TH01	6/9.3	6/9.3	10	0.734	0.001	Stepwise	0.001/0.002	1e-06	0.1	No	No
vWA	14/20	14/20	12	0.775	0.001	Stepwise	0.001/0.002	1e-06	0.1	No	No
D12S391	18/19.3	19.3/21	23	0.884	0.001	Stepwise	0.001/0.002	1e-06	0.1	No	No
D13S317	11/12	12/12	9	0.779	0.001	Stepwise	0.001/0.002	1e-06	0.1	No	No
Penta E	7/11	11/16	21	0.890	0.001	Stepwise	0.001/0.002	1e-06	0.1	No	No
D16S539	11/11	11/11	9	0.737	0.001	Stepwise	0.001/0.002	1e-06	0.1	No	No
D18S51	12/15	12/15	23	0.866	0.001	Stepwise	0.001/0.002	1e-06	0.1	No	No
D19S433	12/13	12/13	17	0.739	0.001	Stepwise	0.001/0.002	1e-06	0.1	No	No
D21S11	27/28	27/31	26	0.822	0.001	Stepwise	0.001/0.002	1e-06	0.1	No	No
Penta D	10/12	10/12	24	0.807	0.001	Stepwise	0.001/0.002	1e-06	0.1	No	No
D22S1045	11/15	15/16	9	0.674	0.0012	Stepwise	0.001/0.002	1e-06	0.1	No	No

This is KLINK version 1.1.0 (changelog | official releases). If you encounter problems, please file a bug report. See also the KLINK homepage for more information.


KLINK: Kinship with pairwise linked markers
INPUT
RESET
Load .fam file
sibship.fam
Upload complete
Marker map
Built-in
Custom
(Optional) .xml
No file selected
Example 1
Example 2
SETTINGS
Map function
Kosambi
Haldane
Empty markers
Hide
Show
Likelihoods
Hide
Show
Loglik
Decimals
3
Ignore linkage above (cM)
1000
Calculate LR
Ped 1
Hide untyped components
1
2
NN1
NN2
Ped 2
Marker
1
2
3
NN1
NN2
Linkage map
Marker data
LR table
Download
Marker Person1 Person2 LR Linked No link No mut
D1S1656 15.3/16 15/15.3 0.882 0.882 0.882 0.874
TPOX 8/9 8/8 0.742 1.720 1.722 1.713
D2S441 10/14 10/14 2.322
D2S1338 17/20 24/24 0.500 0.500 0.500 0.500
D3S1358 14/18 14/18 3.787 3.787 3.787 3.799
FGA 20/21 20/21 3.192 3.192 3.192 3.202
D5S818 9/12 9/12 3.205 23.903 21.215 24.158
CSF1PO 13/13 13/13 6.618
SE33 25.2/28.2 20/21 0.500 0.500 0.500 0.500
D7S820 8/12 8/12 3.448 3.448 3.448 3.457
D8S1179 14/14 14/14 2.813 2.813 2.813 2.822
D10S1248 13/14 13/16 0.721 0.721 0.721 0.721
TH01 6/9.3 6/9.3 2.011 2.011 2.011 2.012
vWA 14/20 14/20 9.499 4.531 9.353 4.532
D12S391 18/19.3 19.3/21 0.985
D13S317 11/12 12/12 0.819 0.819 0.819 0.816
Penta E 7/11 11/16 0.857 0.857 0.857 0.858
D16S539 11/11 11/11 2.129 2.129 2.129 2.133
D18S51 12/15 12/15 3.724 3.724 3.724 3.735
D19S433 12/13 12/13 3.705 3.705 3.705 3.722
D21S11 27/28 27/31 0.932 2.674 2.780 2.680
Penta D 10/12 10/12 2.984
D22S1045 11/15 15/16 0.714 0.714 0.714 0.714
Total LR 2.75 × 10^5 5.25 × 10^5 2.80 × 10^5
This is KLINK version 1.1.0 (changelog | official releases). If you encounter problems, please file a bug report. See also the KLINK homepage for more information.



Resultatrapport: sibship

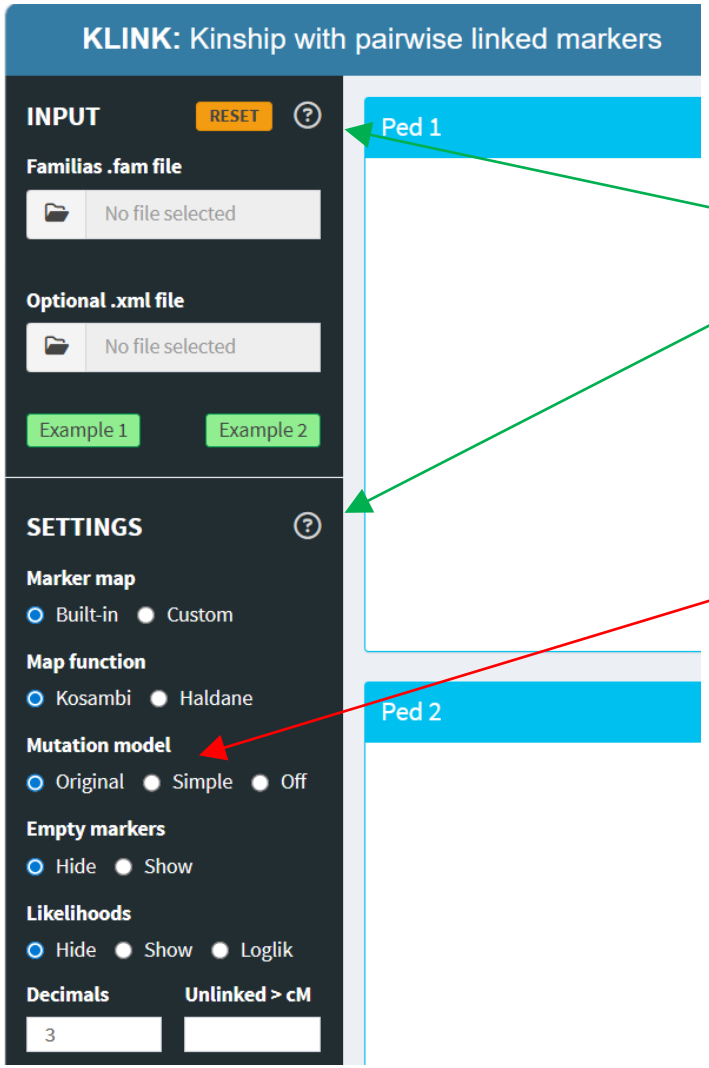
	Markør	Person1	Person2	LR		
1	D1S1656	15.3-16	15-15.3	0.882		
2	D2S1338	17-20	24-24	0.500		
3	D3S1358	14-18	14-18	3.787		
4	FGA	20-21	20-21	3.192		
5	SE33	25.2-28.2	20-21	0.500		
6	D7S820	8-12	8-12	3.448		
7	D8S1179	14-14	14-14	2.813		
8	D10S1248	13-14	13-16	0.721		
9	TH01	6-9.3	6-9.3	2.011		
10	D13S317	11-12	12-12	0.819		
11	Penta E	7-11	11-16	0.857		
12	D16S539	11-11	11-11	2.129		
13	D18S51	12-15	12-15	3.724		
14	D19S433	12-13	12-13	3.705		
15	D22S1045	11-15	15-16	0.714		
16	TPOX	8-9	8-8	1.720	0.742	
17	D2S441	10-14	10-14		2.322	
18	D5S818	9-12	9-12	23.903	3.205	
19	CSF1PO	13-13	13-13		6.618	
20	vWA	14-20	14-20	4.531	9.499	
21	D12S391	18-19.3	19.3-21		0.985	
22	D21S11	27-28	27-31	2.674	0.932	
23	Penta D	10-12	10-12		2.984	
	Total LR			**2.752e+05**		

	Samples
Person1	NN1
Person2	NN2

	Relationship
Ped 1	Full siblings
Ped 2	Half siblings

	Notifications
	No notifications recorded

	Settings
1	KLINK version: 1.1.0
2	Familias file: sibship.fam
3	Genetic map: BUILT-IN
4	Database: NorskDB_2023
5	Map function: Kosambi
6	Max distance: 1000 cM



Most options can be left untouched

Click on question marks to see explanations

Mutation modelling:

- "Original" = use models in the .fam file (if any)
- "Simple" = Equal model with rate 0.001 (all markers)