



# Marcinelle - Belgium





## Marcinelle coal mine

- 1000 m deep
- 170 000 tonnes/year
- 700+ workers  
(many Italian immigrants)

# 8 August 1956: The Marcinelle mining disaster



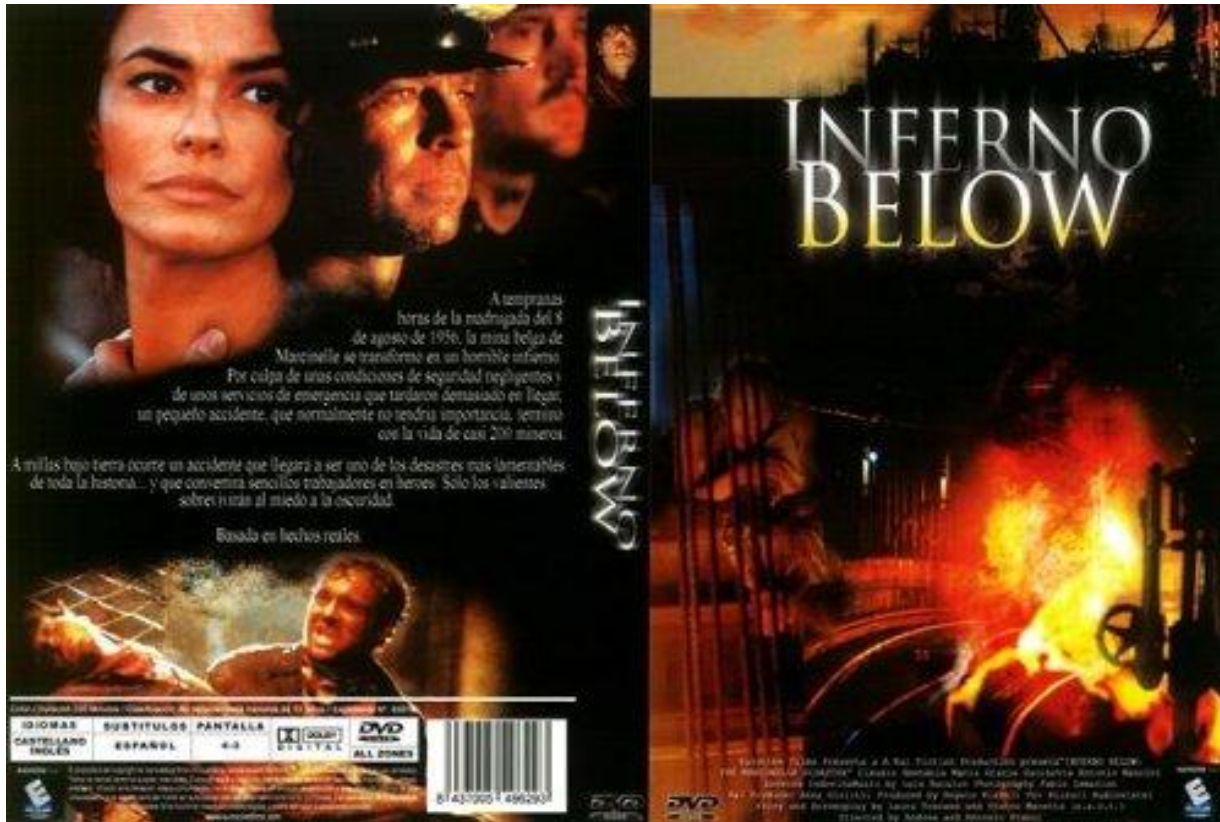
262 dead:

- 136 Italians
- 95 Belgians

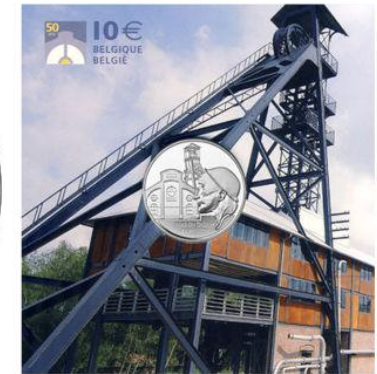


All except 17 identified

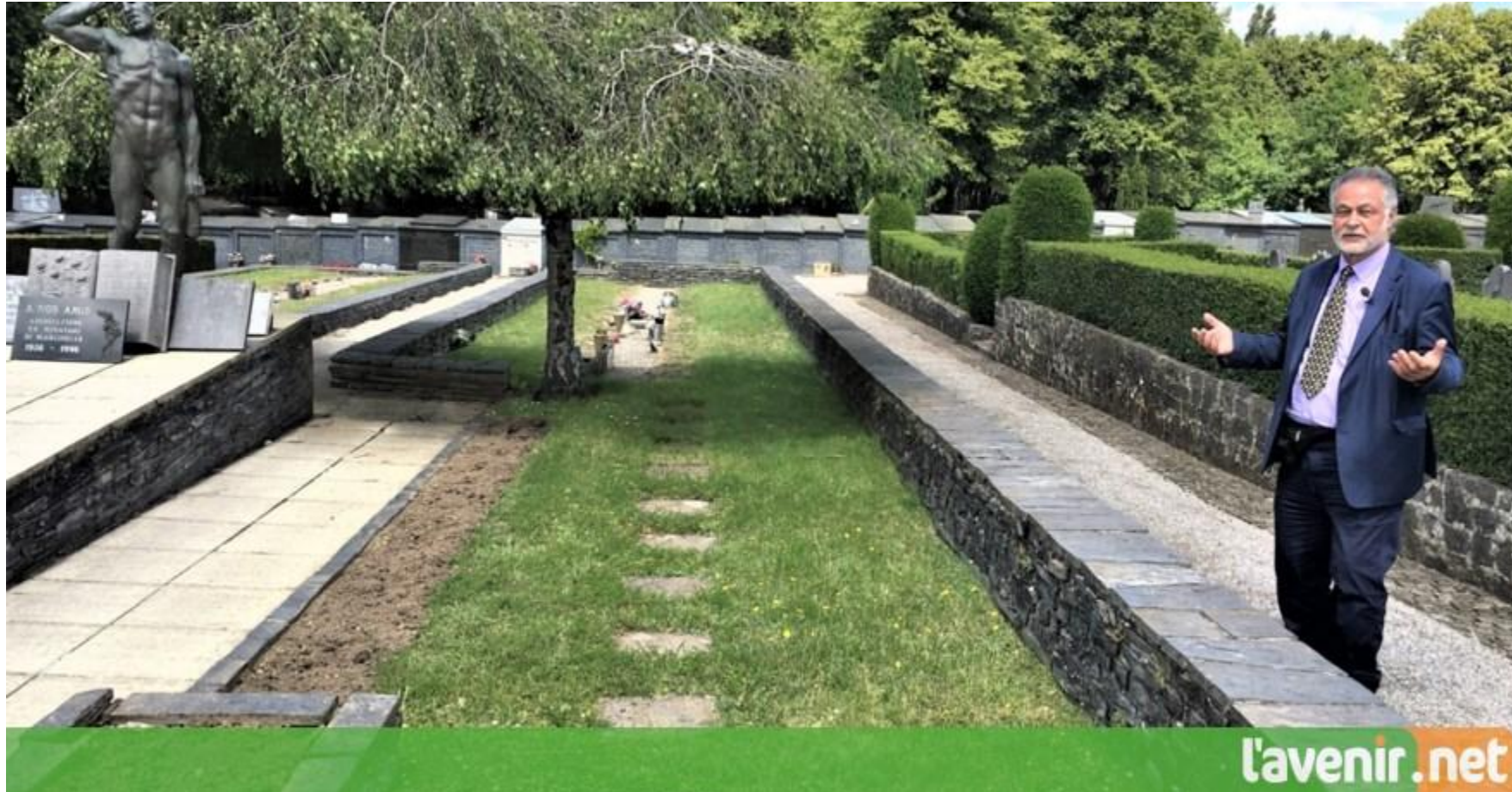
2003 TV series (IMDb: 6.9)



2006 (50th anniversary)

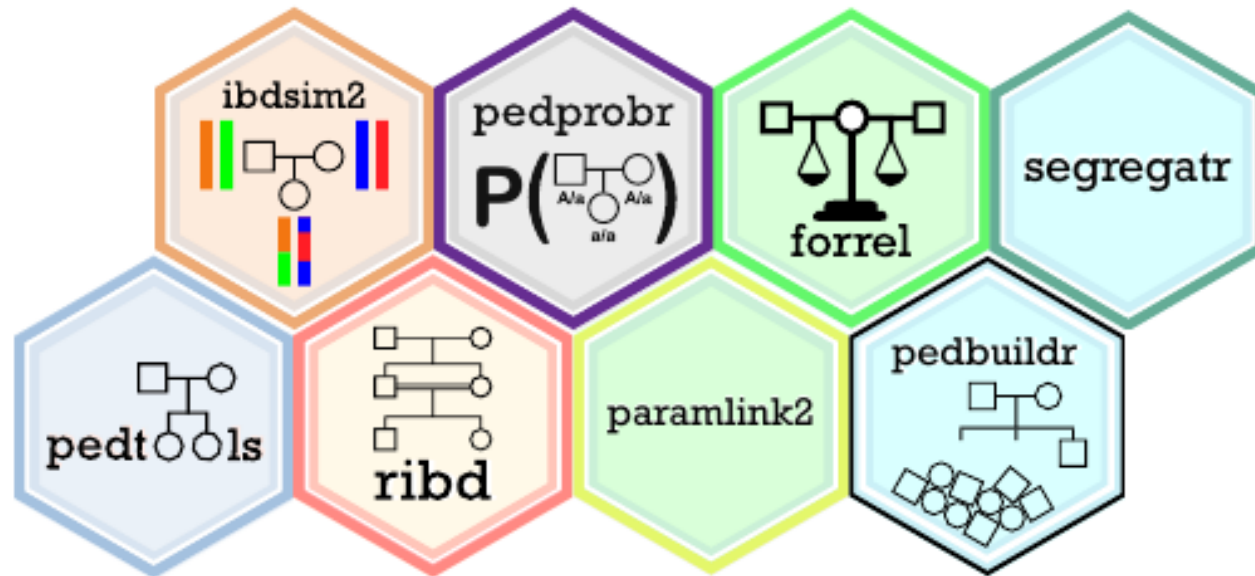


# 2021: Exhumation of the 17 unidentified



# The ped suite:

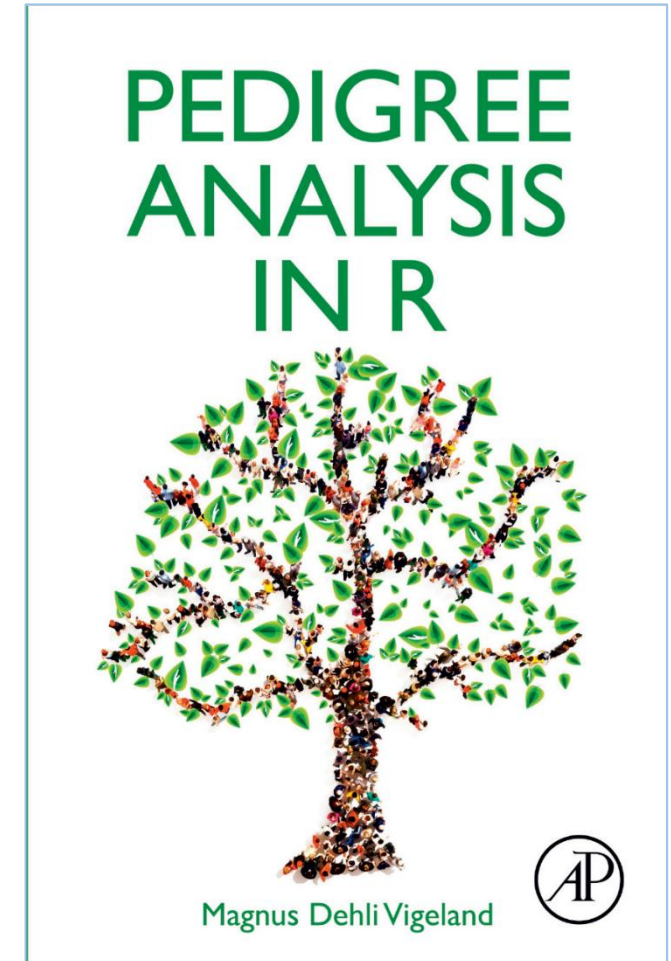
A collection of packages for pedigree analysis in R



Home page: <https://magnusdv.github.io/pedsuite>

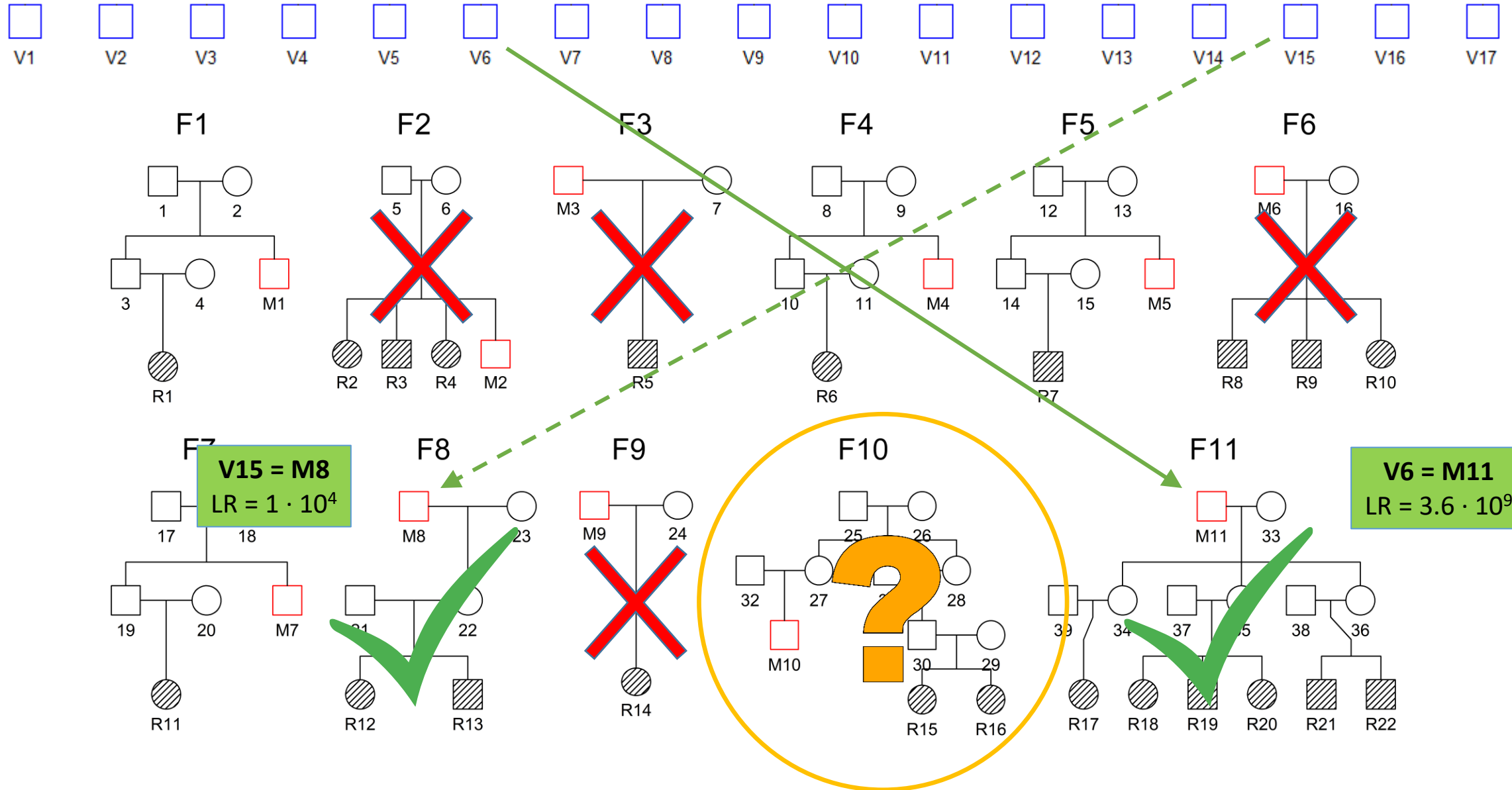
Source code: <https://github.com/magnusdv>

Developers: Magnus D. Vigeland, Thore Egeland

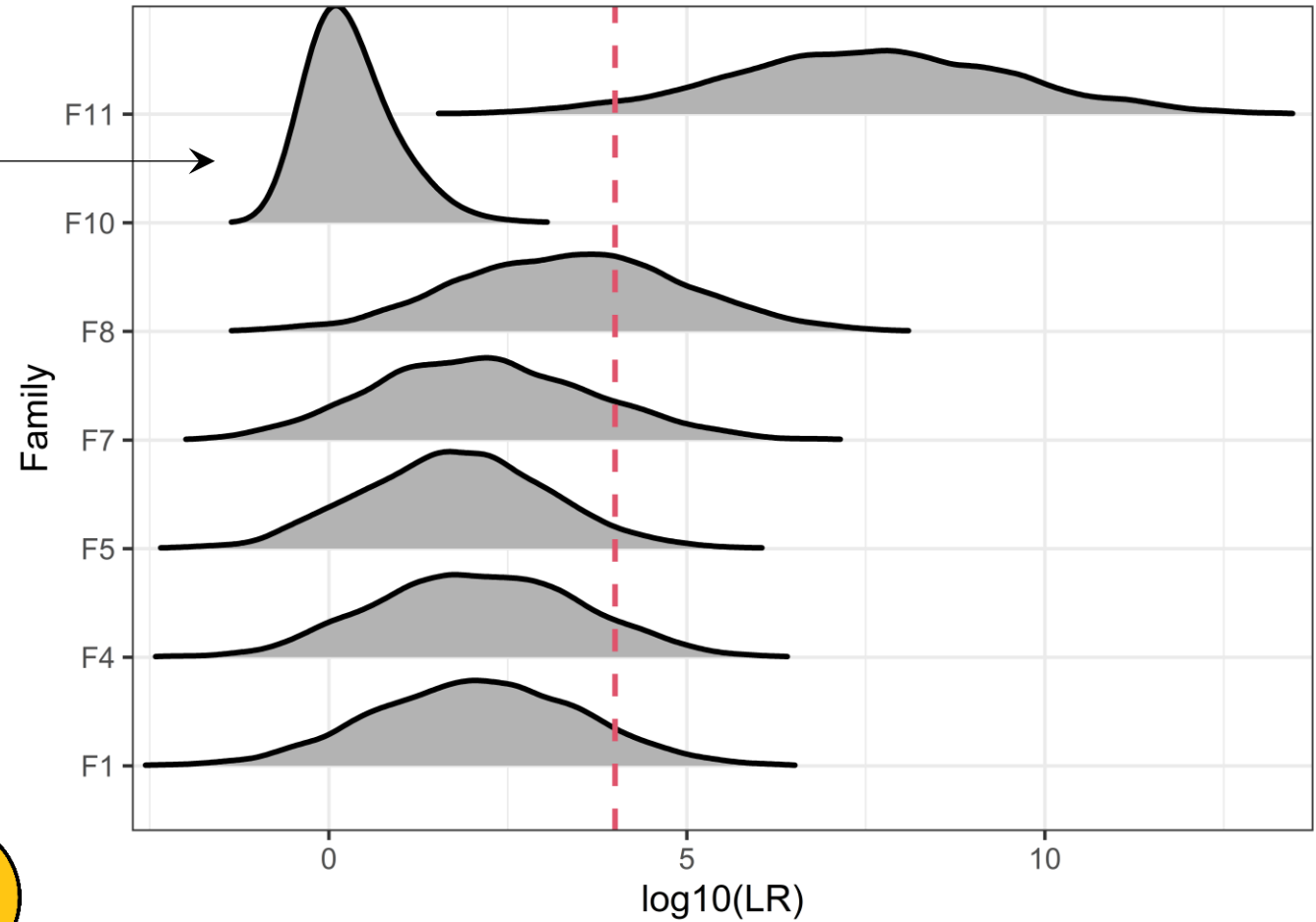
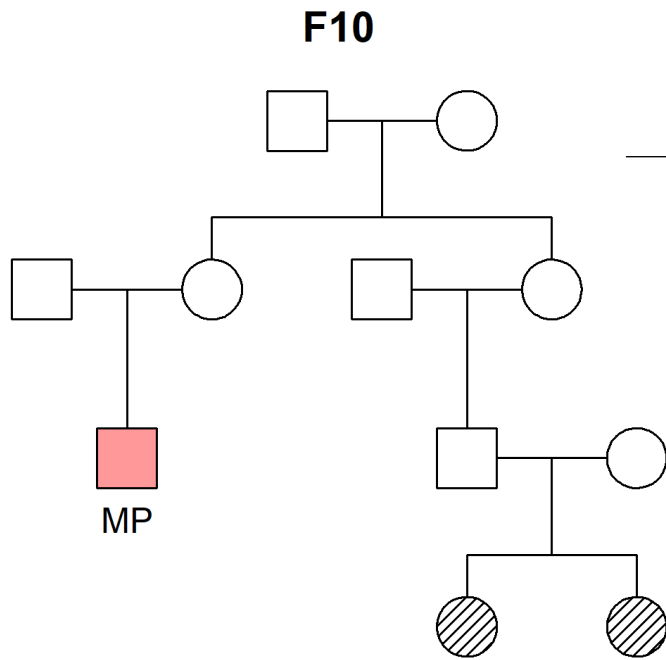




# Victims and 11 reference families: genotyped with 23 STRs



# LR distributions (5000 simulations of MP)



Conclusion for F10:  
**Terrible power**



# What to do?

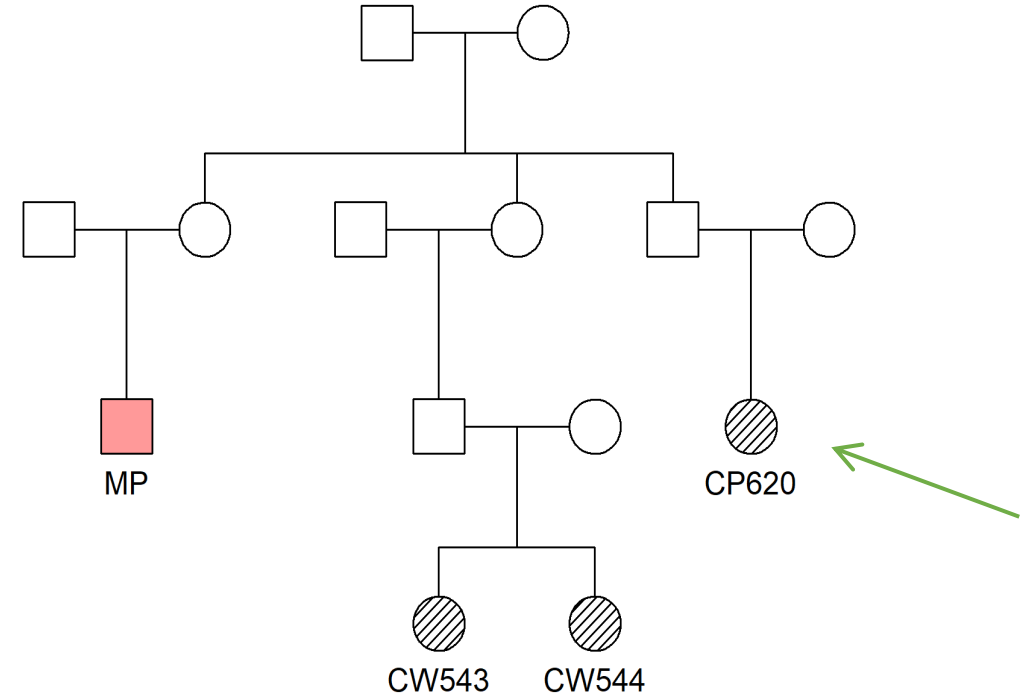
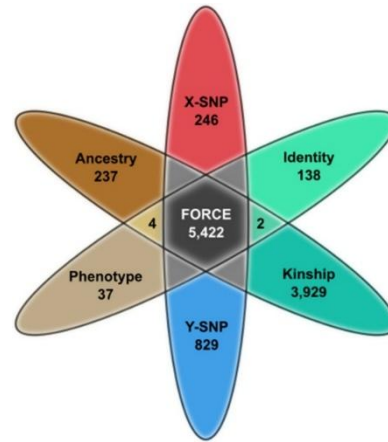
- Recruit additional relatives?



- Lineage markers? (X, ~~Y~~, ~~mtDNA~~)



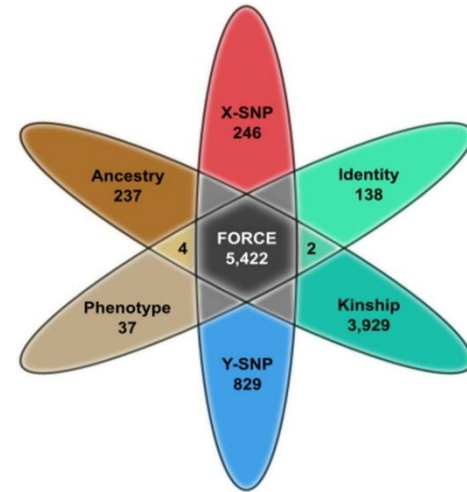
- Use SNP panel?



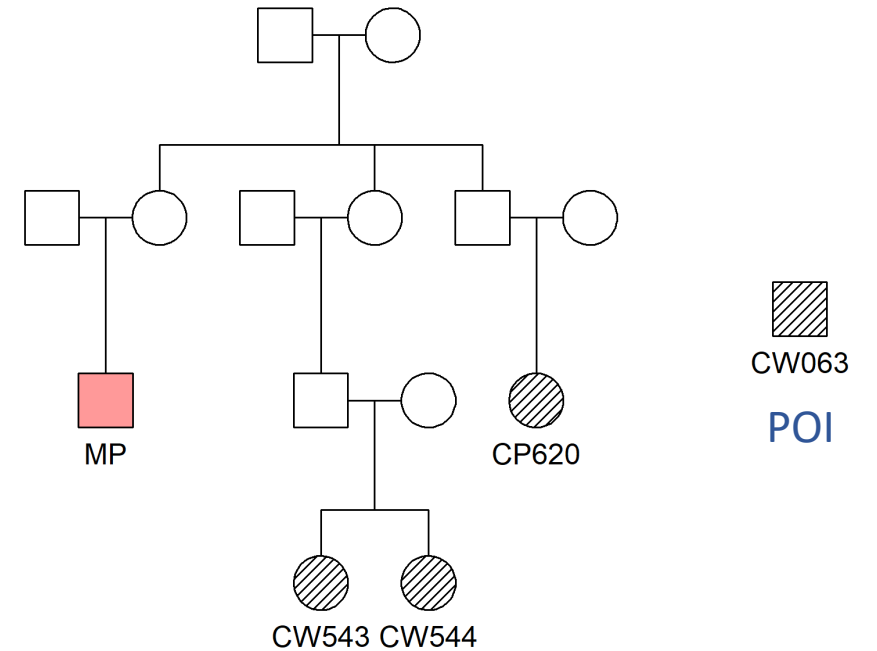
Updated STR analysis:  
LR = 40 for victim V9

# The FORCE panel

- 3931 **autosomal** SNPs for kinship
- 4 samples typed
- After QC: 3836 SNPs (98%)

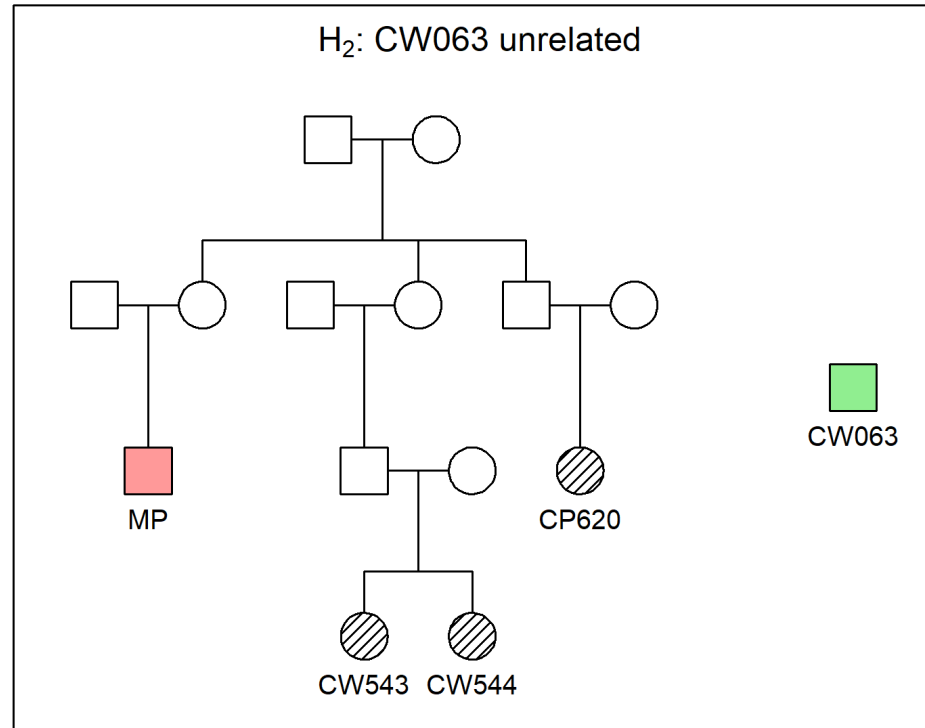
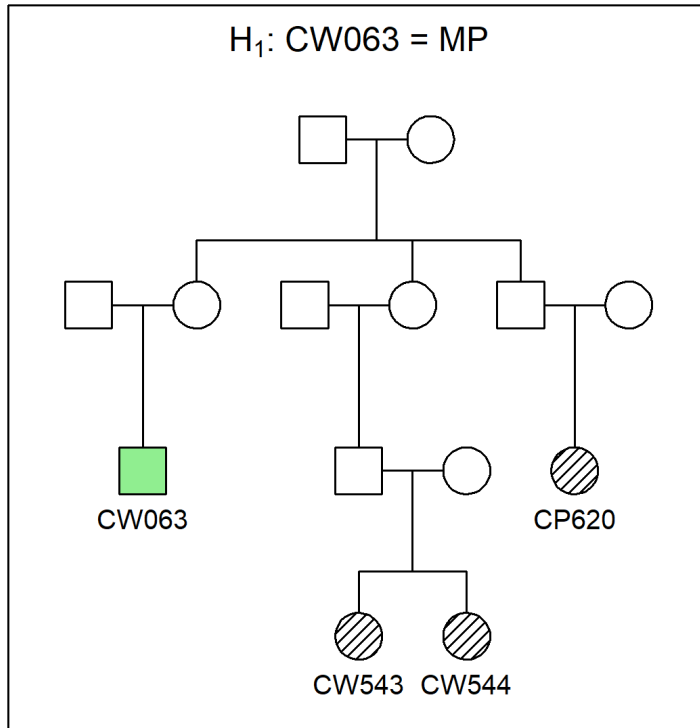


Andreas Tillmar



CHR	MARKER	CM	A1	A2	FREQ1	CW543	CW544	CP620	CW063
1	rs944..	0.00	G	A	0.61	AG	GG	AG	AG
1	rs464..	0.33	C	A	0.64	AA	AC	AC	AC
1	rs1091..	1.24	T	C	0.63	CC	CT	TT	CC
1	rs669..	2.00	C	T	0.57	CC	CC	CC	CC
1	rs376..	4.81	T	G	0.59	GG	GT	GT	GG
1	rs736..	5.86	C	T	0.67	CC	CC	CT	CT
1	rs667..	6.77	G	A	0.56	AG	AG	GG	AA

# LR approach 1

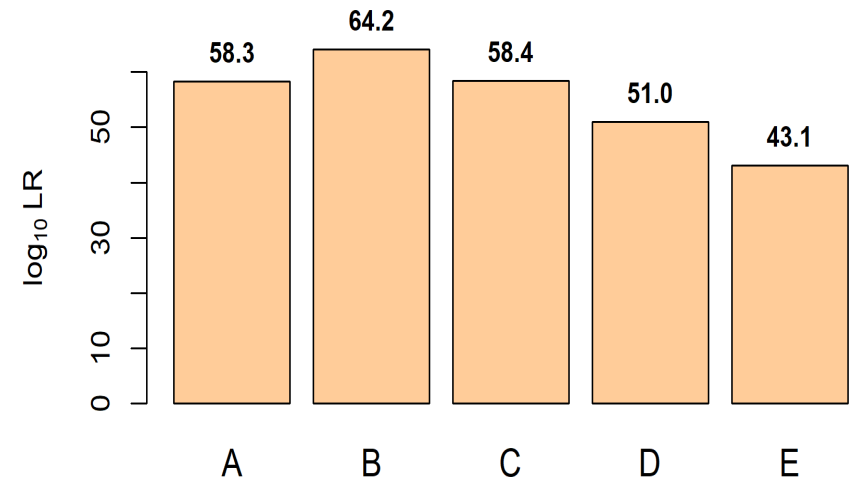
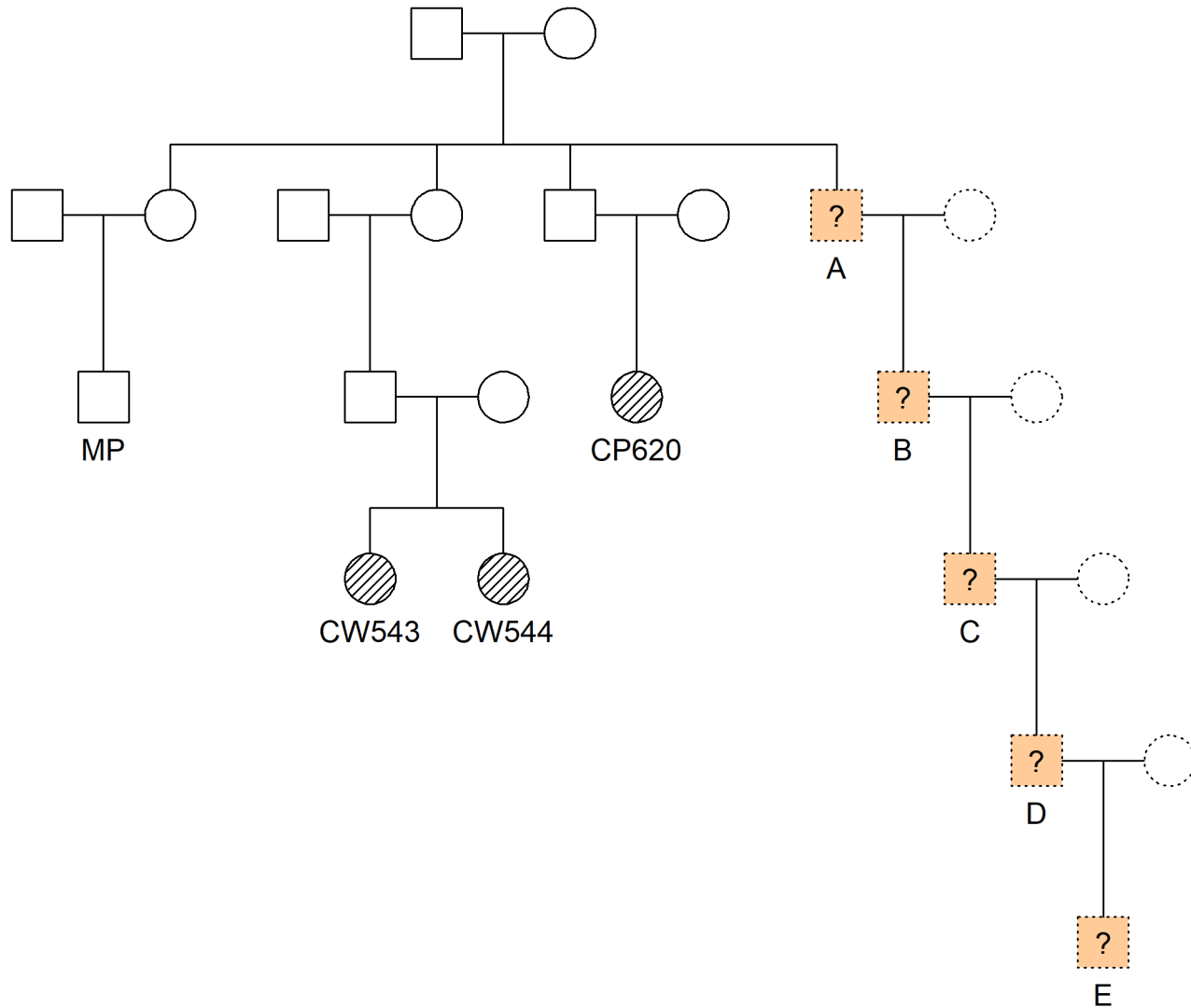


Merlin:  $LR = 1.47 \cdot 10^{64}$





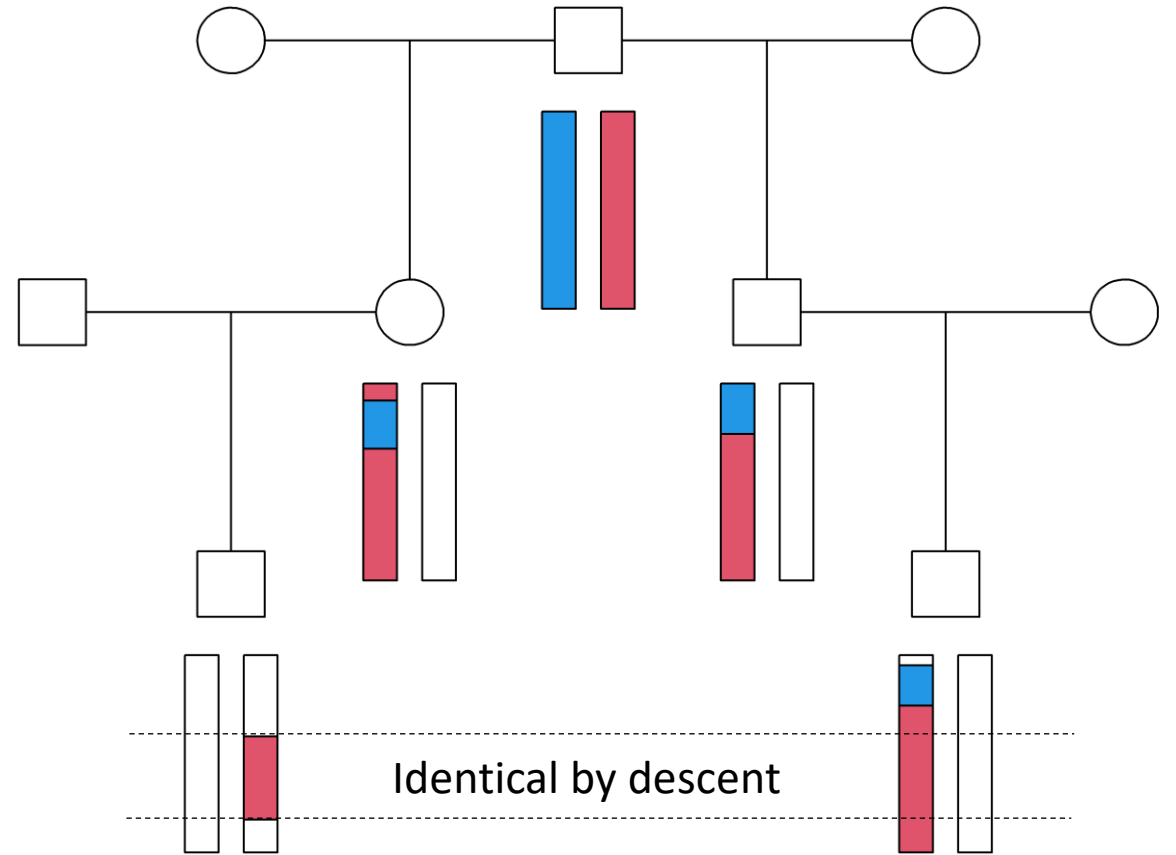
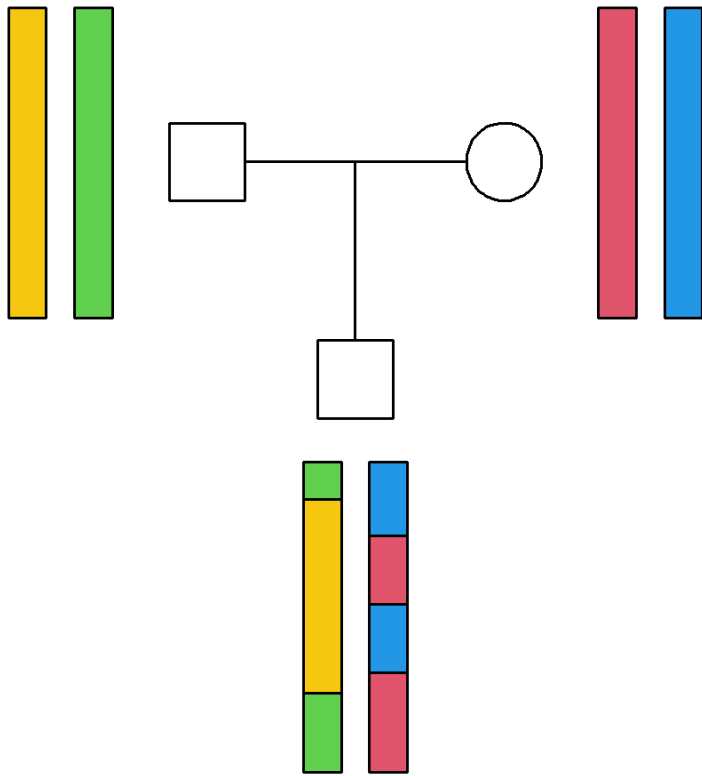
# LR approach 2: H2 = distant relative





IBD segment analysis

# Recap: Recombination and IBD

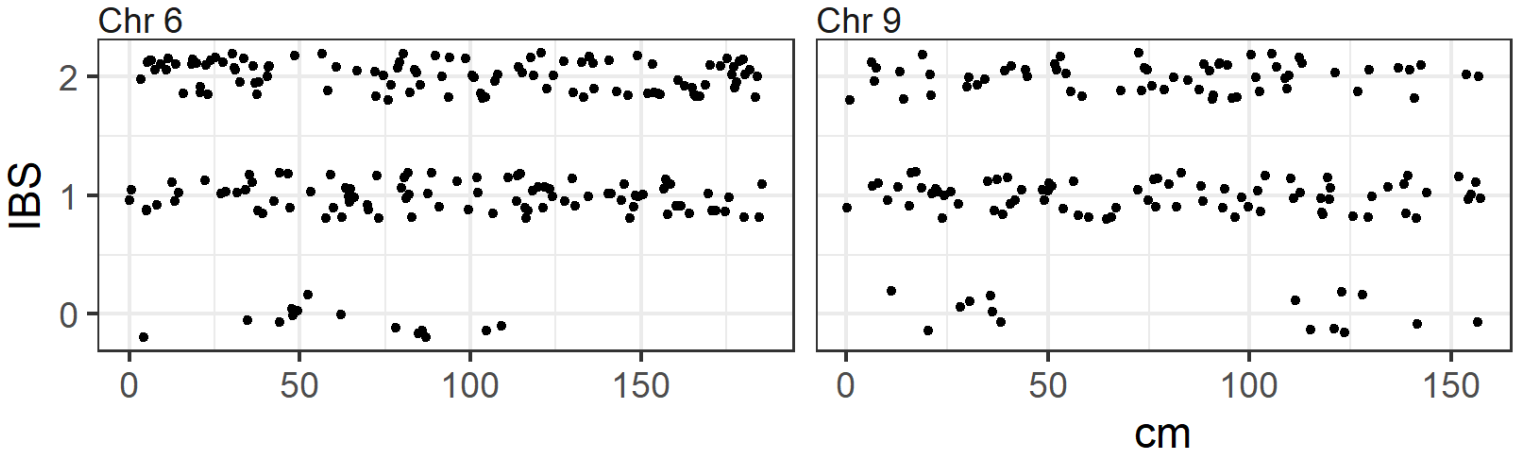


IBD = identity by descent  
IBS = identity by state

# SNPs → IBD segments ??

SNP	Sample1	Sample2	IBS	IBD
snp1	A/A	G/G	0	No
snp2	A/G	A/A	1	?
snp3	A/T	A/T	2	?
...	...	...	...	...

CW063 vs. CP620



# Software for IBD detection

- More than 30 published programs!
- But - some of them ...
  - require  $> N$  markers ( $N =$  huge number)
  - require  $> N$  samples
  - require phased data
  - require massive installation+setup
  - don't work
  - don't exist anymore

Not a single successful run so far



Henrik Nordtorp

	Software	Reference	Comments
1	Software	Reference	Comments
2	PLINK	<a href="#">Purcell et al., 2007</a>	Probabilistic, no
3	BEAGLE HMM	<a href="#">Browning and Browning, 2007</a>	Probabilistic, LD
4	GERMLINE	<a href="#">Gusev et al., 2009</a>	Not probabilist
5	Relate	<a href="#">Albrechtsen et al., 2009</a>	Probabilistic, t
6	KING	<a href="#">Manichaikul et al., 2010</a>	
7	BEAGLE IBD	<a href="#">Browning and Browning, 2010</a>	Probabilistic, LD
8	fastIBD	<a href="#">Browning and Browning, 2011</a>	
9	DASH	<a href="#">Gusev et al., 2011</a>	
10	MCMC_IBDfinder	<a href="#">Moltke et al., 2011</a>	
11	IBDL	<a href="#">Han and Abney, 2011</a>	Probabilistic, LD
12	ERSA	<a href="#">Huff et al., 2011</a>	
13	IBD_Haplo	<a href="#">Brown et al., 2012</a>	Probabilistic, no
14	Refined IBD	<a href="#">Browning and Browning, 2013a</a>	
15	IBDseq	<a href="#">Browning and Browning, 2013b</a>	
16	IBD-Groupon	<a href="#">He, 2013</a>	
17	HapFABIA	<a href="#">Hochreiter, 2013</a>	
18	PREST-plus	<a href="#">Sun and Dimitromanolakis, 2014</a>	Likelihood
19	Parente2	<a href="#">Rodriguez et al., 2015</a>	
20	FISHR	<a href="#">Bjelland et al., 2017</a>	
21	hmmIBD	<a href="#">Schaffner et al., 2018</a>	Hidden Markov
22	DRUID	<a href="#">Ramstetter et al., 2018</a>	
23	TRUFFLE	<a href="#">Dimitromanolakis et al., 2019</a>	
24	RaPID	<a href="#">Naseri et al., 2019</a>	PBWT on phase
25	TRIBES	<a href="#">Twine et al., 2019</a>	
26	FastSMC	<a href="#">Nait Saada et al., 2020</a>	Hash/Extend pl
27	IBIS	<a href="#">Seidman et al., 2020</a>	Sliding Window
28	Hap-IBD	<a href="#">Zhou et al., 2020</a>	Error adjusted I
29	iLash	<a href="#">Shemirani et al., 2021</a>	Locality sensitiv
30	Phaseibd/TPBWT	<a href="#">Freyman et al., 2021</a>	TPBWT
31	ancIBD	<a href="#">Ringbauer et al., 2024</a>	
32	COANCESTRY	<a href="#">Wang, 2011</a>	
33	NGSremix		
34	NgsRelate		
35	IBDMAP	<a href="#">Bercovici et al., 2010</a>	HMM

# ibdfindr:

hassle-free

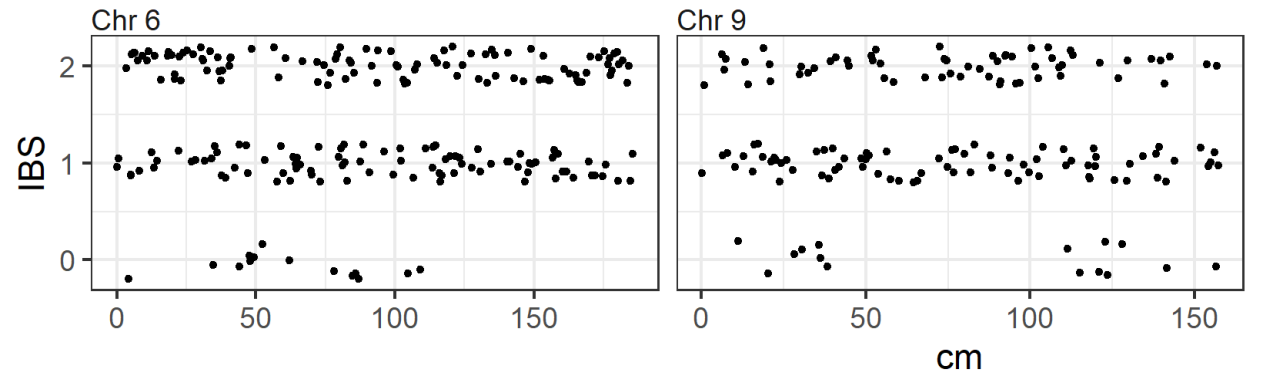
Source code, docs & examples:  
<https://github.com/magnusdv/ibdfindr>

An R package for detecting IBD segments from SNP data

Input data (N = 3836):

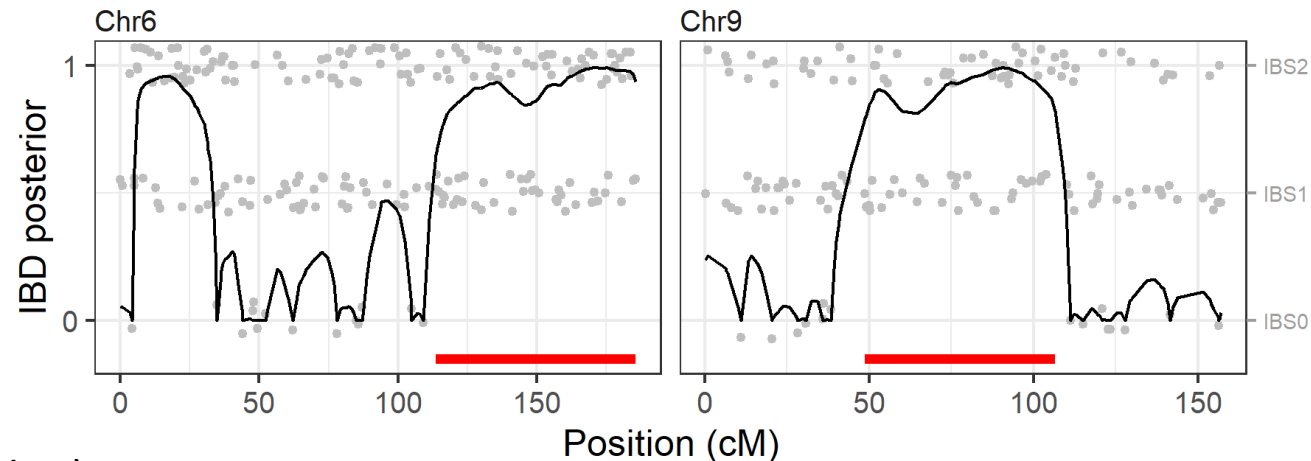
CHR	MARKER	CM	A1	FREQ1	CW063	CP620
1	rs944..	0.00	G	0.61	AG	AG
1	rs464..	0.33	C	0.64	AC	AC
1	rs1091..	1.24	T	0.63	CC	TT
1	rs669..	2.00	C	0.57	CC	CC
1	rs376..	4.81	T	0.59	GG	GT
1	rs736..	5.86	C	0.67	CT	CT
1	rs667..	6.77	G	0.56	AA	GG

CW063 vs. CP620



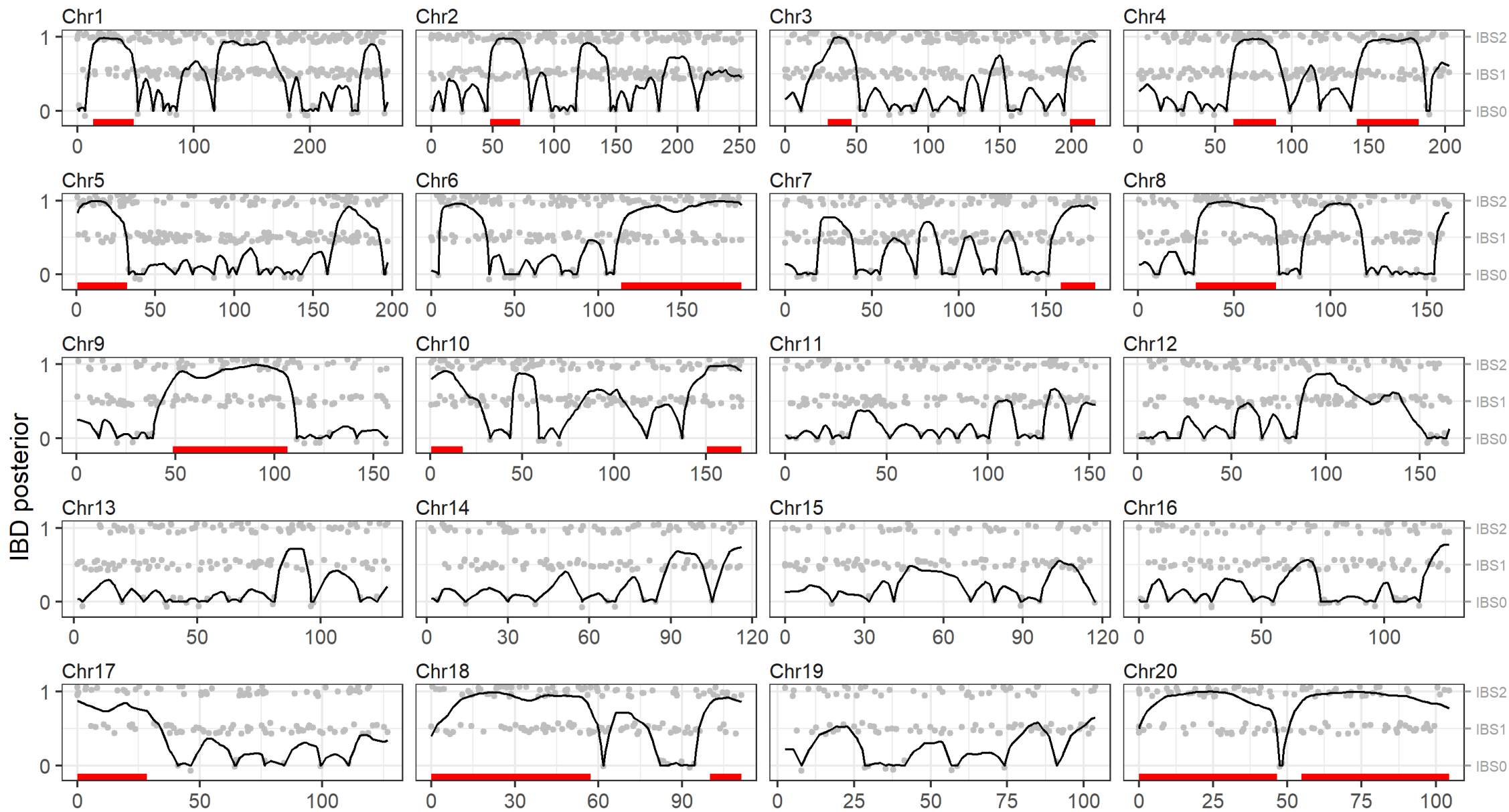
## R code

```
library(ibdfindr)
ibd = findIBD(data)
plotIBD(ibd, chrom = c(6,9))
```

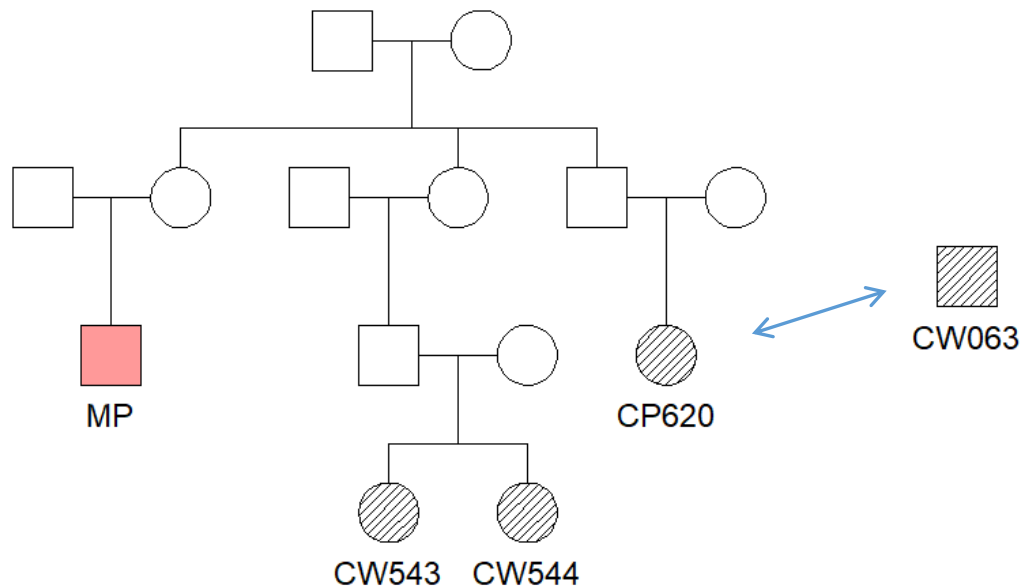


- Fits a continuous-time HMM
- Predicts IBD segments (Viterbi algorithm)
- Finds posterior IBD probs (forward-backward algorithm)

# IBD segments and posteriors for CW063 vs. CP620



+ chrom 21, 22



Summary for CW063 vs. CP620:

- 22 segments
- 744 cM total sharing

Q1: Compatible with MP? (1st cousins)

Shiny app **ibdsim2**:

<https://magnusdv.shinyapps.io/ibdsim2-shiny/>

## IBD sharing by family members

Pedigree 1

Built-in pedigree

Choose

Load ped file

marcinelle-F10

Individuals

MP, CP620

Label

F10

PARAMETERS

Chromosome

1 - 22  X

Crossover model

$\chi^2$   Haldane

Sex-specific map

On  Off

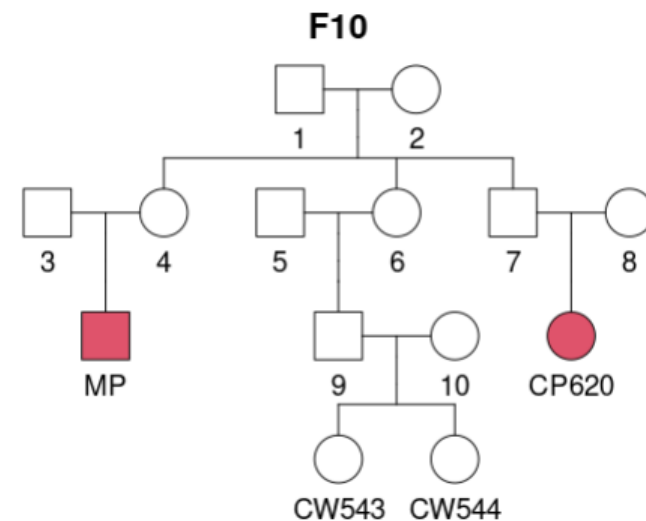
Cutoff

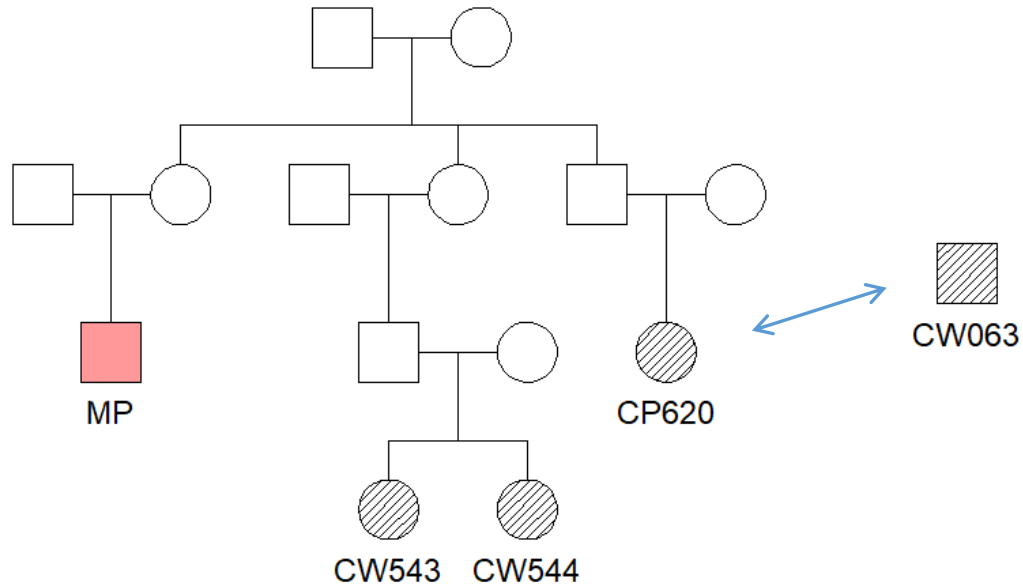
10

Seed

123

Simulate!





Summary for CW063 vs. CP620:

- 22 segments
- 744 cM total sharing

Q1: Compatible with MP? (1st cousins)



Shiny app **ibdsim2**:

<https://magnusdv.shinyapps.io/ibdsim2-shiny/>

Q2: Are there more likely alternatives?

## IBD sharing by family members

Pedigree 1

Built-in pedigree

Choose

Load ped file

marcinelle-F10

Individuals

MP, CP620

Label

F10

PARAMETERS

Chromosome

1 - 22  X

Crossover model

$\chi^2$   Haldane

Sex-specific map

On  Off

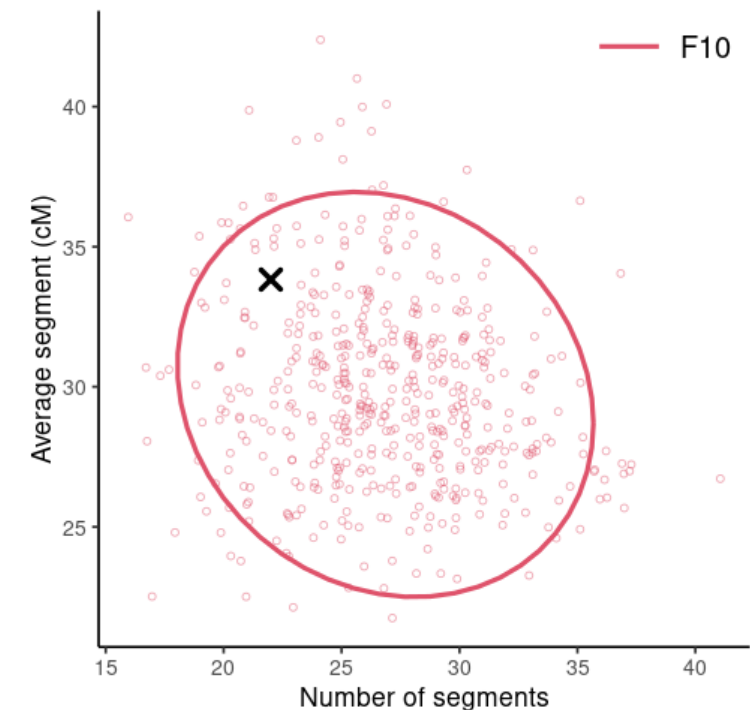
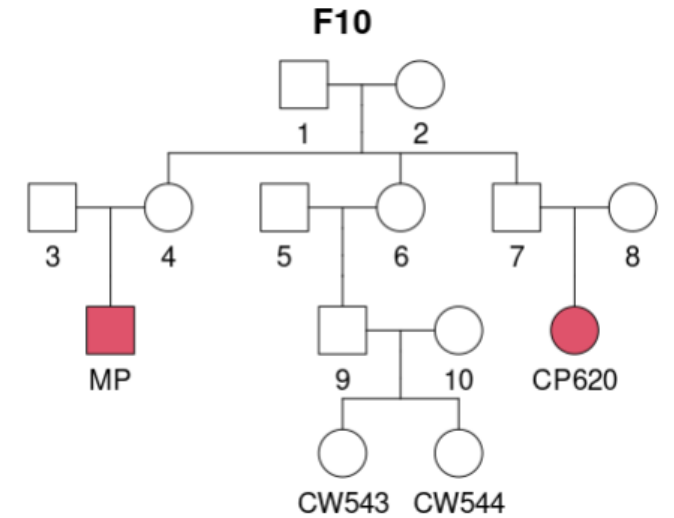
Cutoff

10

Seed

123

Simulate!



# ibdRel: Classifying relationships from IBD segments

<https://github.com/hnordto/>



Henrik Nordtorp

ibdRel Analysis Settings

**Input**

Segment lengths

- 34.6
- 48.9
- 24.3
- 16.6
- 17.7
- 27.8
- 40.4
- 31.6
- 27.4
- 72

Cutoff

10

Classification resolution

Pedigree  Equivalence class  Kappa

Kinship  Degree

Classify

**Classification** Pairwise test

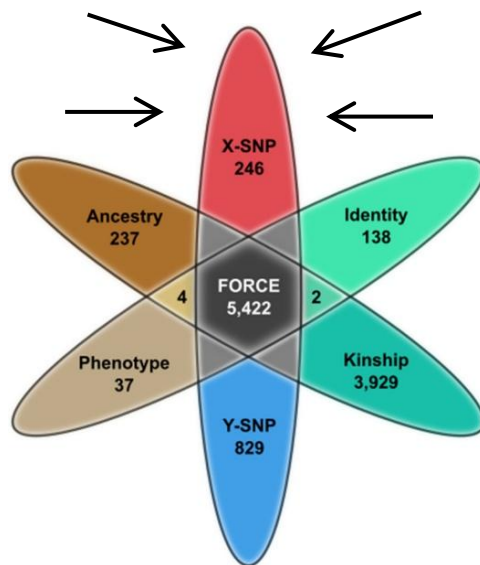
RELATIONSHIP	LIKELIHOOD	OUTLIER	LOF
<a href="#">g-grandfather--g-grandson-m</a>	0.3491	FALSE	1.0155277
<b><a href="#">1st cousins-p</a></b>	<b>0.2162</b>	<b>FALSE</b>	<b>0.9953897</b>
<a href="#">granduncle--grandnephew-p</a>	0.1799	FALSE	1.0108339
<a href="#">half-uncle--half-nephew-m</a>	0.1289	FALSE	1.0810152
<a href="#">granduncle--grandnephew-m</a>	0.0860	FALSE	1.1189264
<a href="#">half-uncle--half-nephew-p</a>	0.0210	FALSE	1.1927262
<a href="#">1st cousins-m</a>	0.0178	TRUE	1.3384601
<a href="#">1st cousins 1r-p</a>	0.0004	TRUE	1.4597961
<a href="#">g-granduncle--g-grandnephew-p</a>	0.0003	TRUE	1.4204110
<a href="#">gg-grandfather--gg-grandson-m</a>	0.0002	TRUE	1.5981151
<a href="#">g-grandfather--g-grandson-p</a>	0.0002	TRUE	2.2419124
<a href="#">half 1st cousins-m</a>	0.0001	TRUE	2.2007680

Choose pedigree

1st cousins-p



# The X-FORCE!



- 246 X-chromosomal SNPs

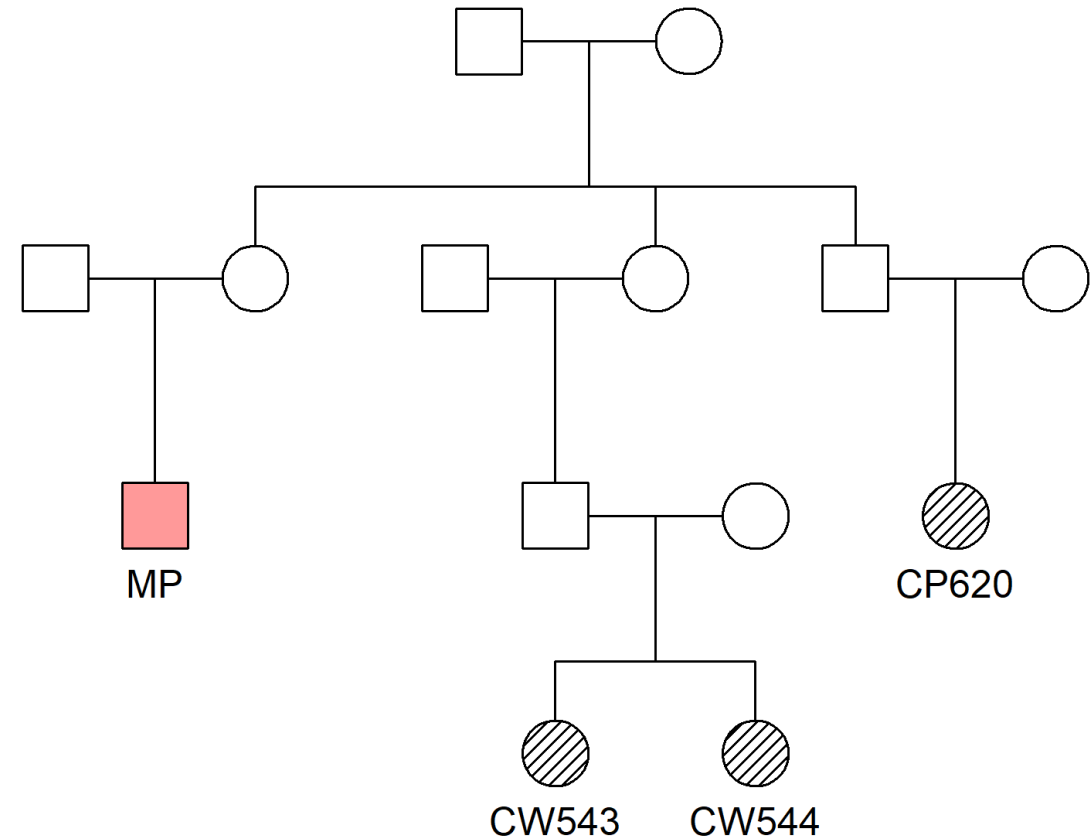
# But...are the individuals related on X?

## Autosomal

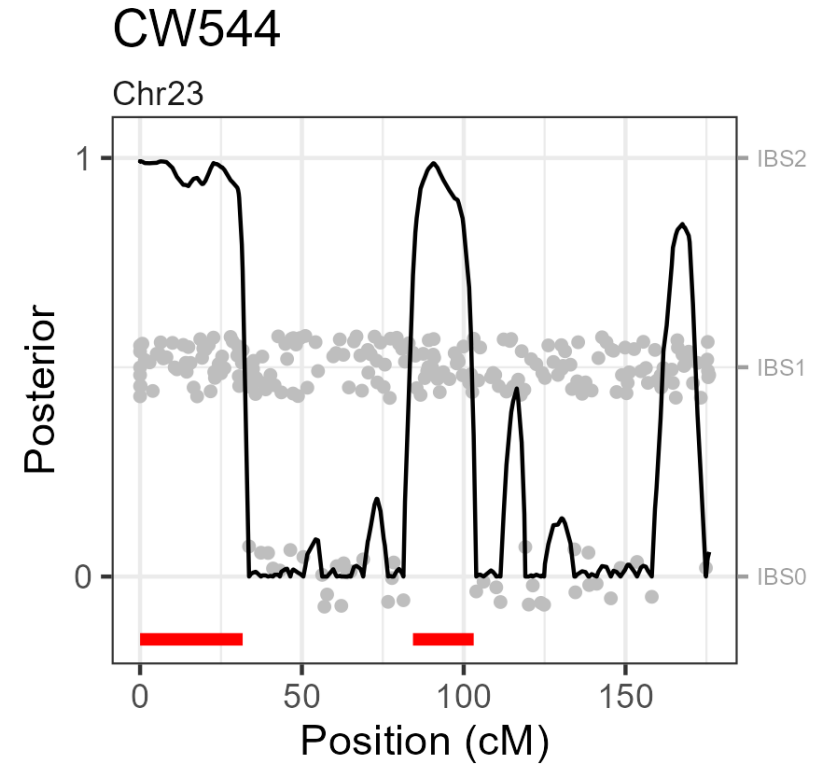
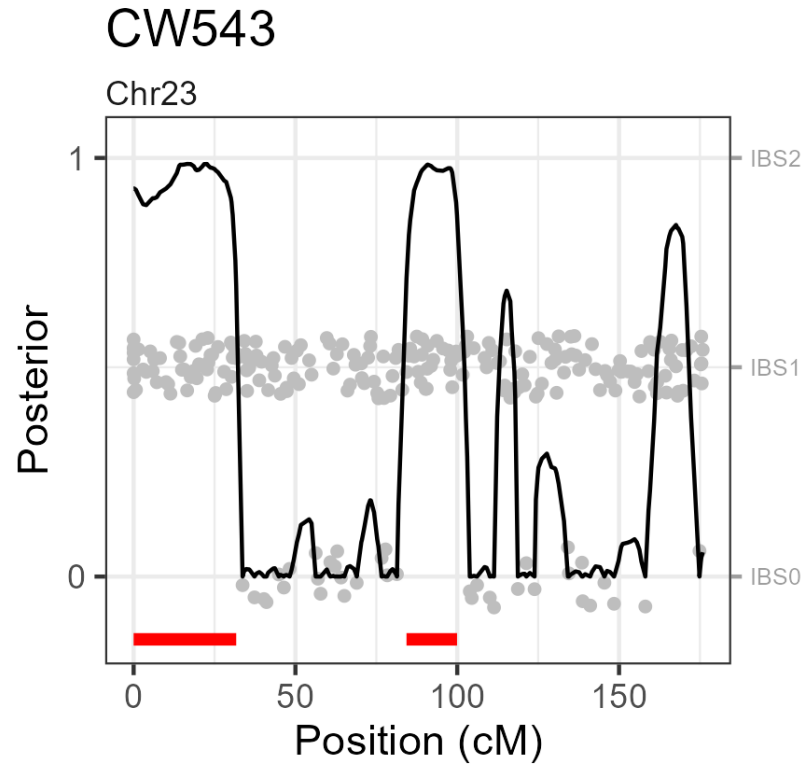
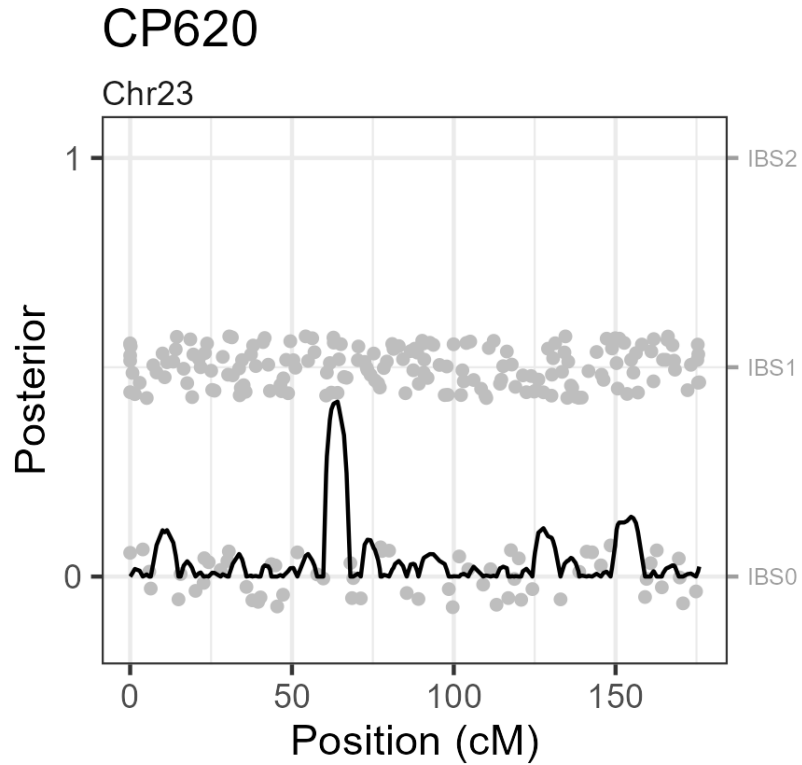
	CP620	CW543	CW544
Kinship ( $\varphi$ )	0.0625	0.03125	0.03125
Expected IBD	25%	12.5%	12.5%

## X-chromosomal

	CP620	CW543	CW544
Kinship ( $\varphi$ )	0.125	0.1875	0.1875
Expected IBD	25%	<b>37.5%</b>	<b>37.5%</b>



# ibdfindr: comparing POI to each ref on X



- No IBD sharing on X between POI and CP620!?
- Does this ruin everything?
- **ibdsim2**: No, this is quite normal for this relationship!




Pedigree 1

Built-in pedigree

Choose ▾

Load ped file

 marcinelle-F10.

Individuals

MP, CP620

Label

F10

---

PARAMETERS

Chromosome

1 - 22  X


Crossover model

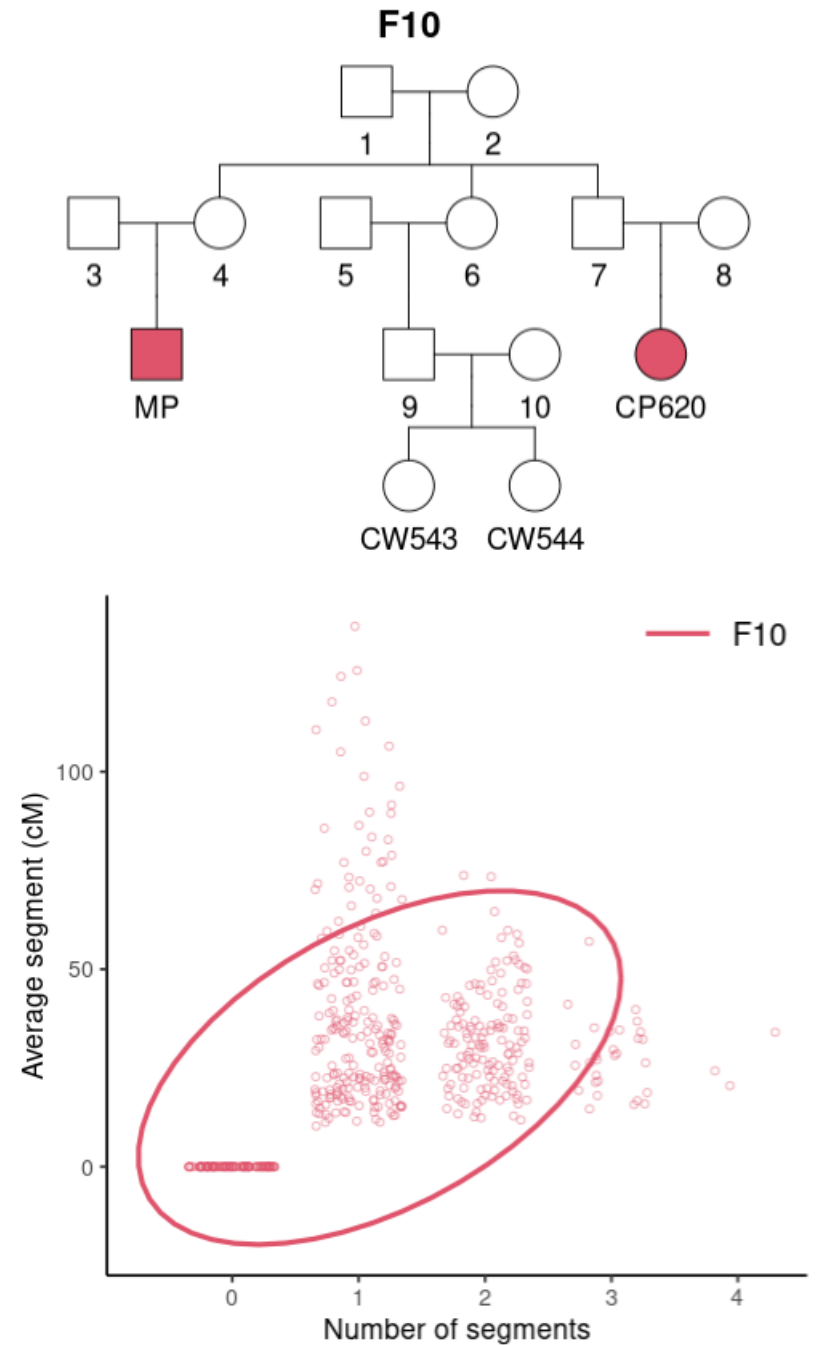
$\chi^2$   Haldane

Sex-specific map

On  Off

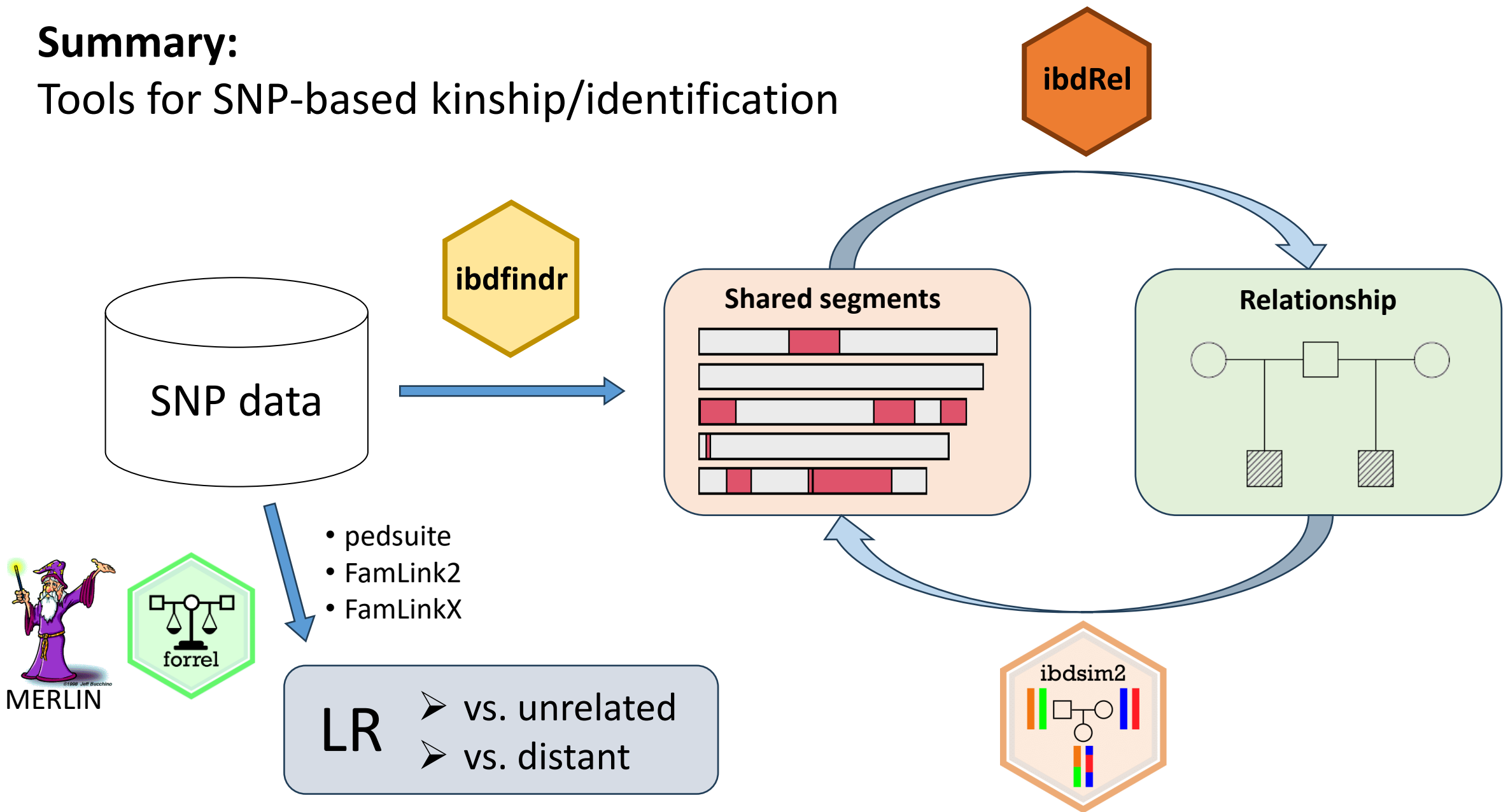
Cutoff  Seed





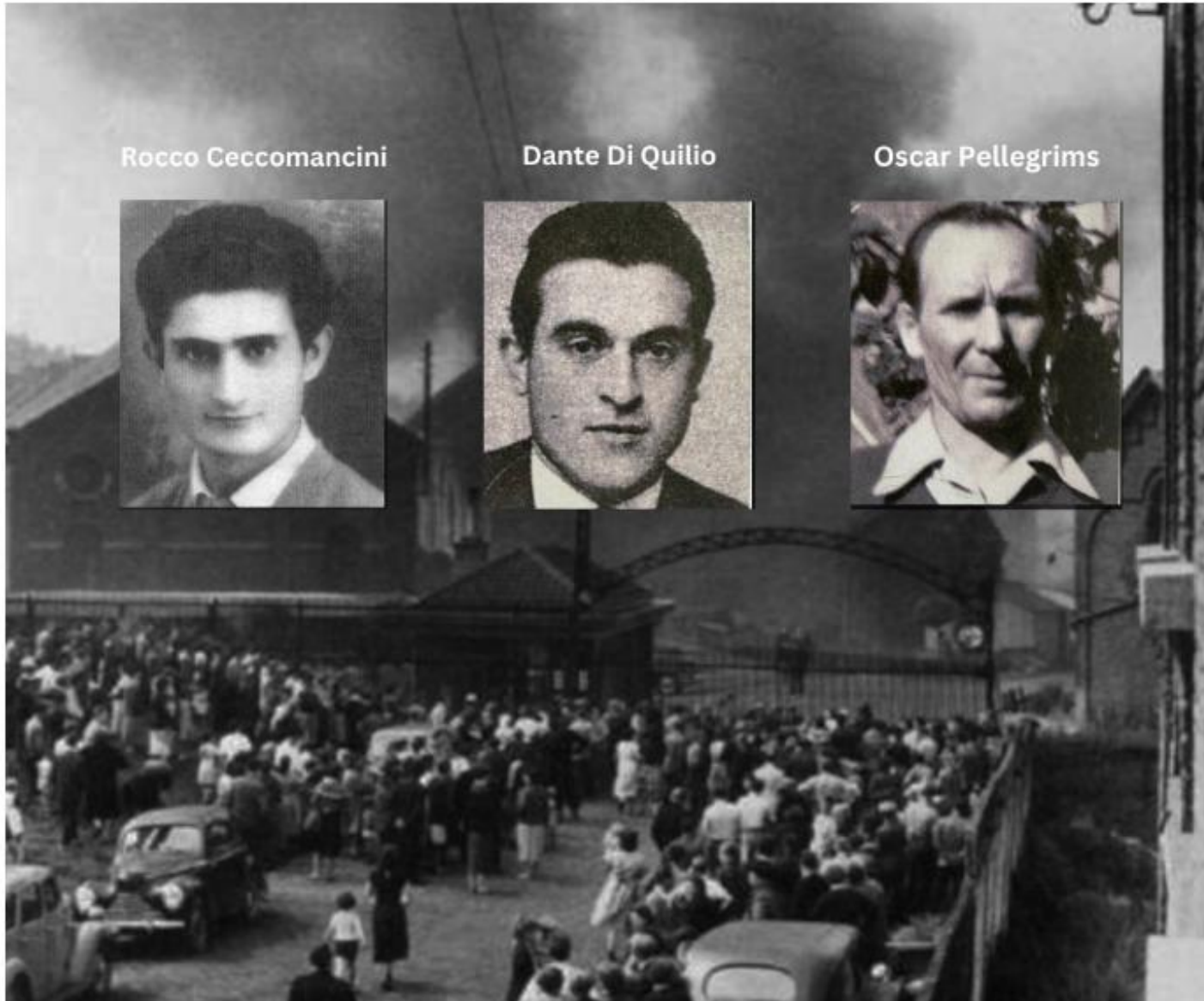
# Summary:

## Tools for SNP-based kinship/identification



27/03/2024

# A third victim from Bois du Cazier identified thanks to DNA analysis



Rocco Ceccomancini



Dante Di Quilio



Oscar Pellegrims



## Collaborators

### Belgium/NICC

- Stijn Desmyter
- Beatrijs Vanhooydonck

### Sweden/Linköping

- Andreas Tillmar
- Adam Staadig
- Pernilla Haage

### Norway/Oslo

- Thore Egeland
- Magnus D. Vigeland
- Henrik Nordtorp