Relatedness part 2:
Why are some siblings more alike than others?

Magnus Dehli Vigeland

Statistical methods in genetic relatedness and pedigree analysis

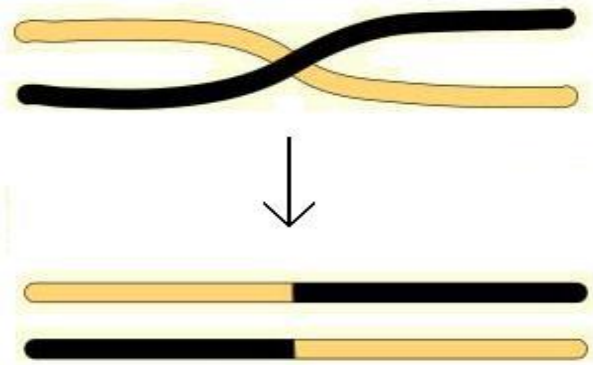
NORBIS course, 6th – 10th of January 2020, Oslo

Plan

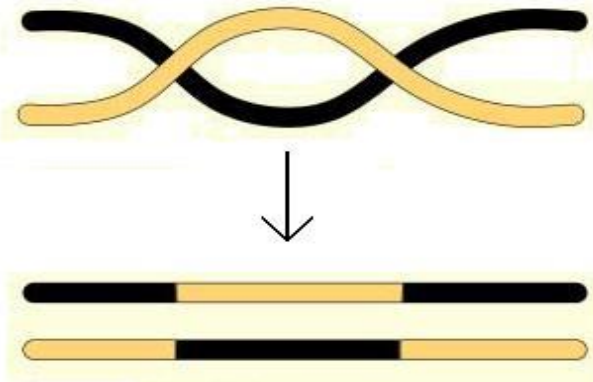
- Revisiting the meiosis
  - Recombination and crossovers
  - Morgan and centiMorgan
- Back to the coefficients: Variation!
- Summary

Recombination

- Crossovers:



Crossover



Double crossover

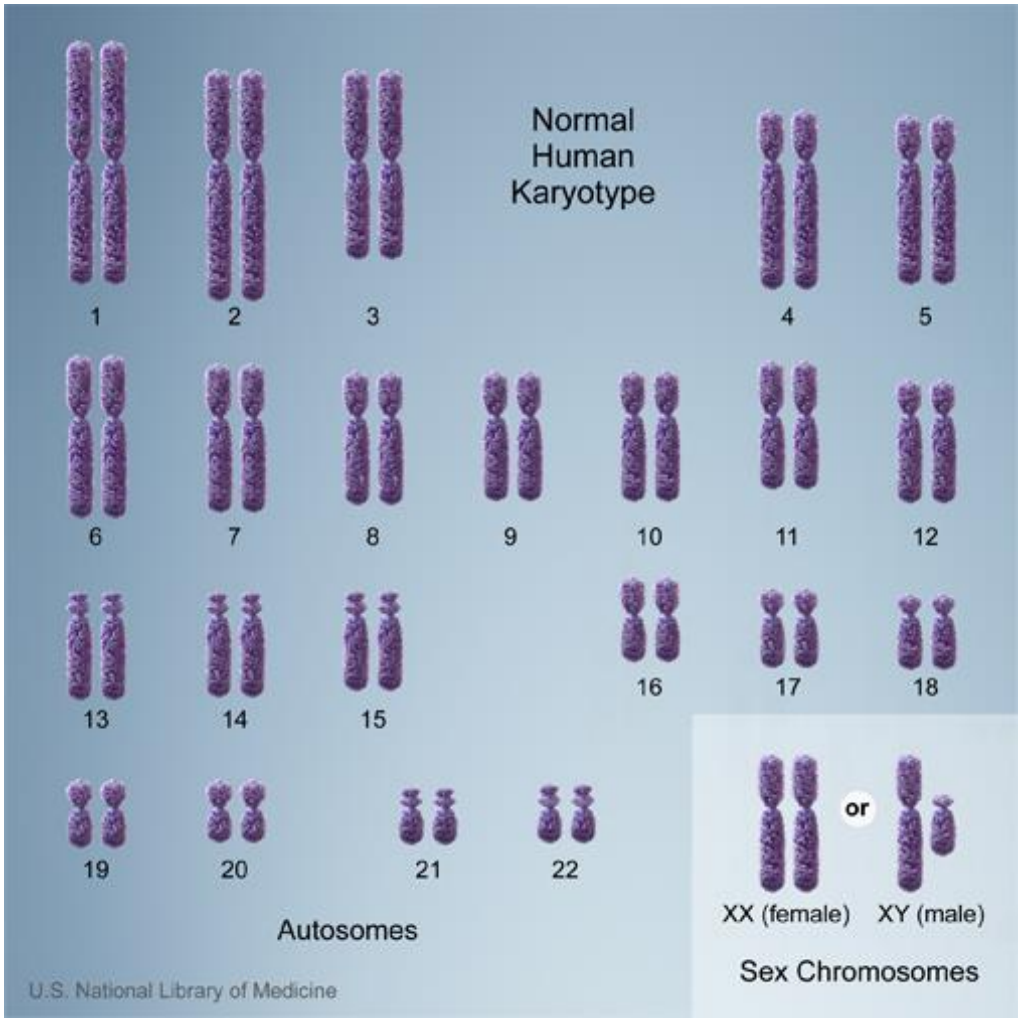
- The genetic distance between two loci:

  = average number of crossovers between them per meiosis

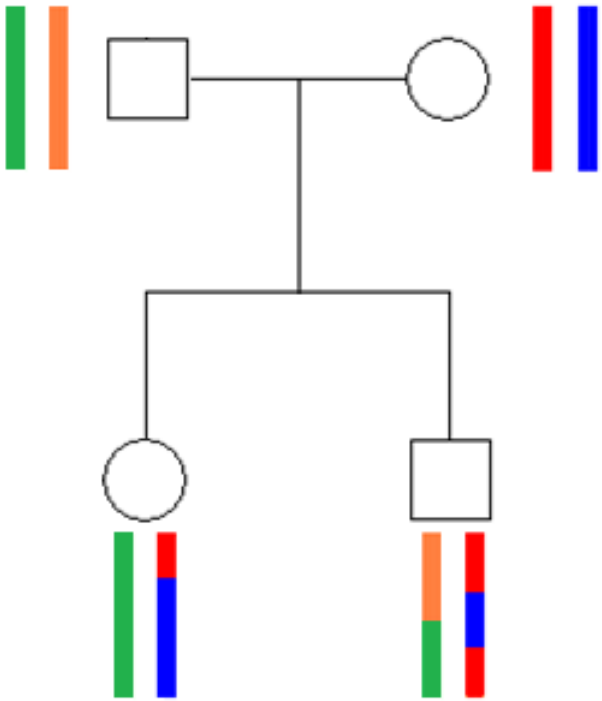
- Units:
  - 1 Morgan (M) = 1 crossover per meiosis (on average)
  - 1 centiMorgan (cM) = 0.01 M

- The human genome: Ca 30 Morgan

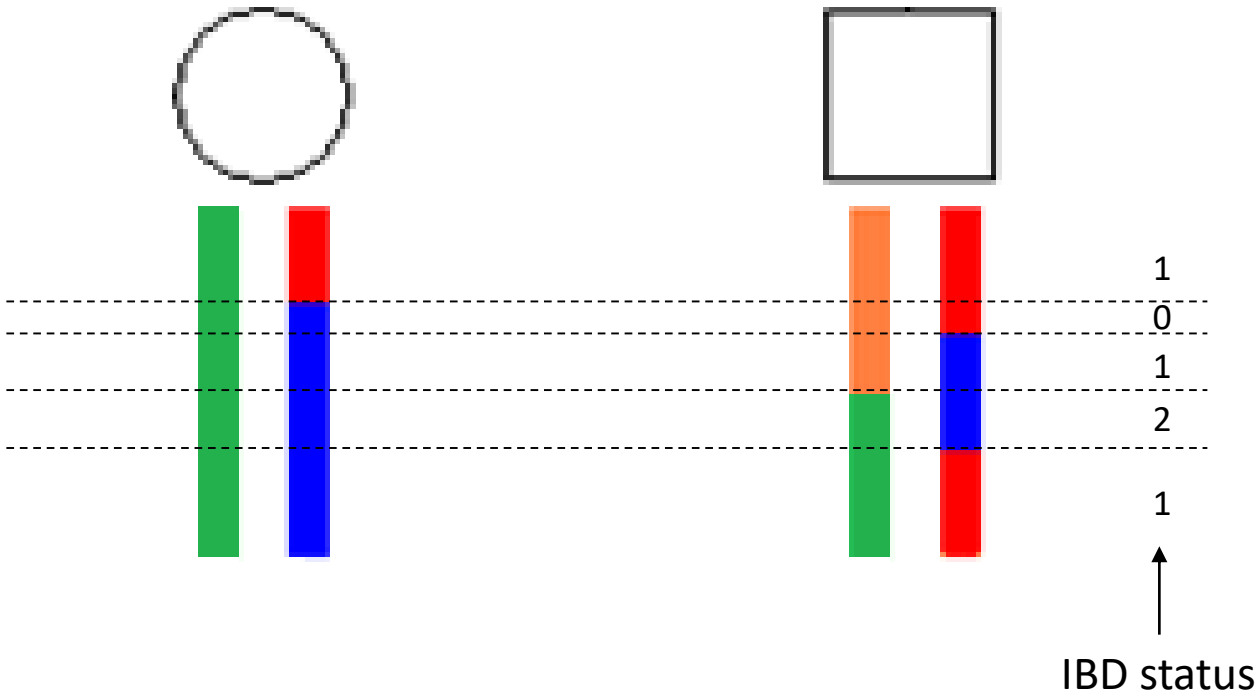
Rule of thumb: One crossover per chromosome arm



How does this work in practice?

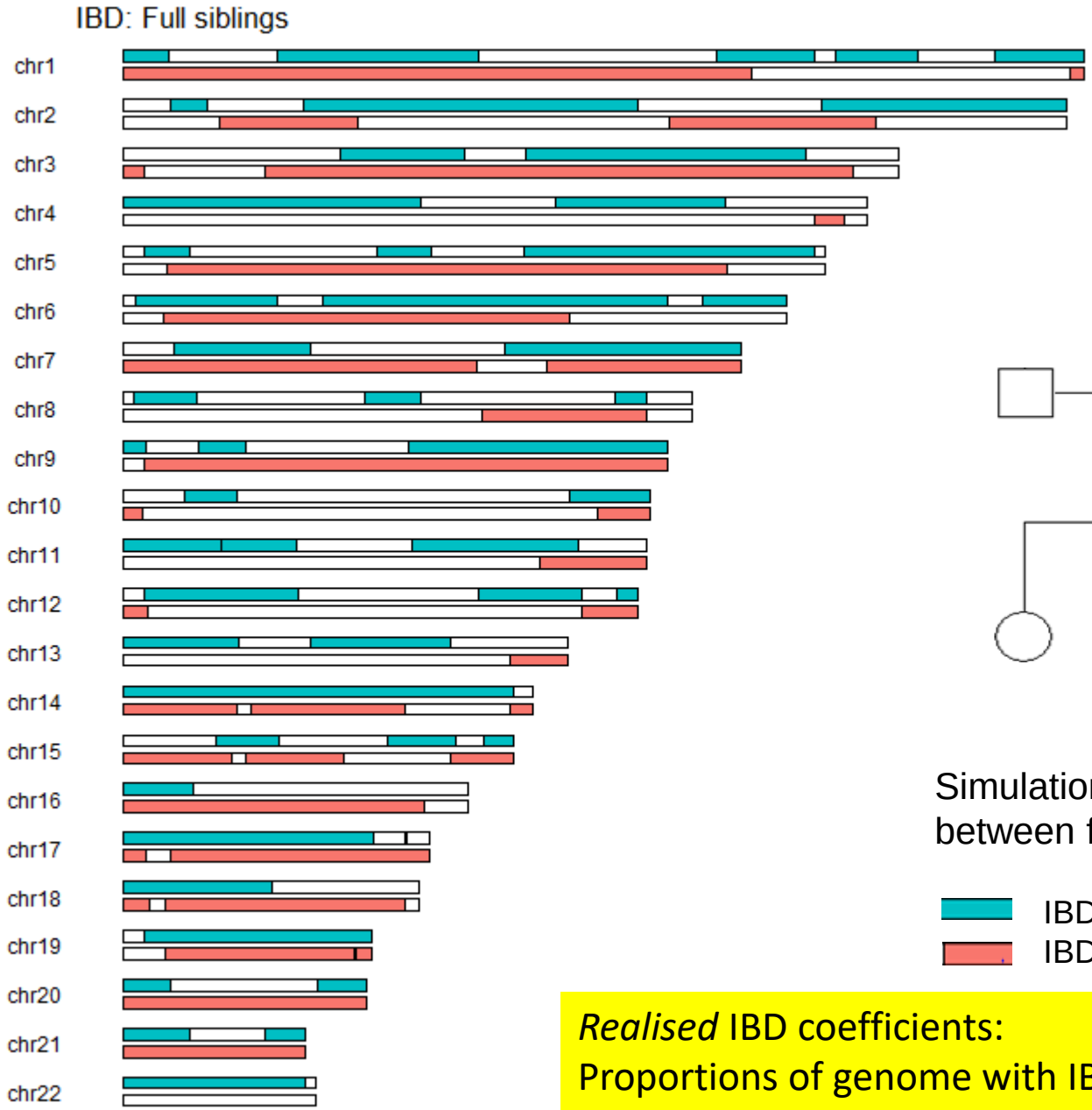


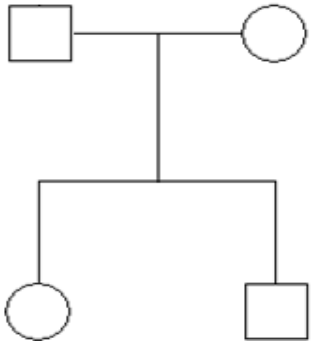
How does this work in practice?



Take home message:

- IBD is not a pointwise concept!
- IBD in a particular locus → IBD in the surrounding region (to closest crossovers)





Simulation of IBD sharing between full siblings

IBD from father
IBD from mother

Realised IBD coefficients:
Proportions of genome with IBD = 0, 1, 2

Realised inbreeding coefficient

- Recall the definition:

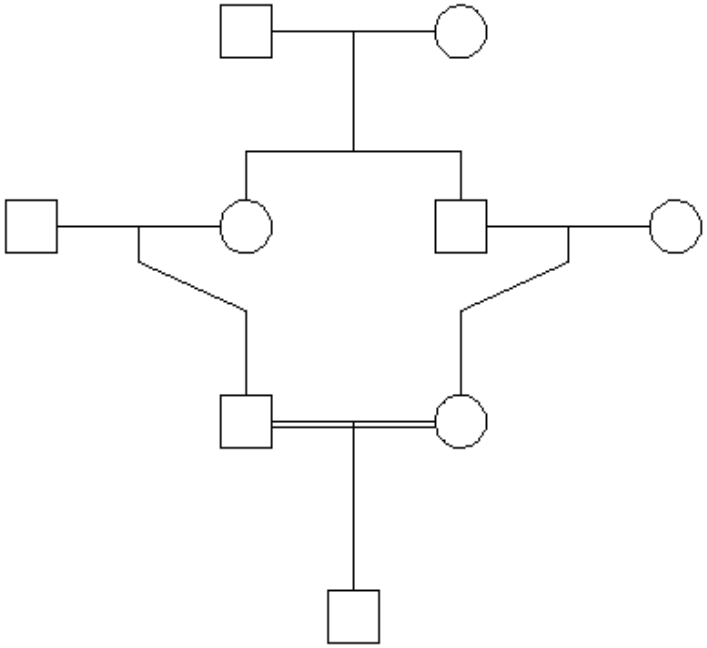
  $f = P($ autozygosity in random locus $)$

  = expected fraction that is autozygous

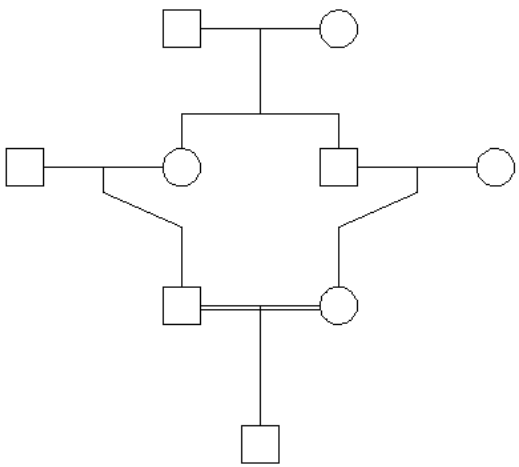
- Empirical/realised/observed inbreeding:

  f_{real} = fraction which is actually autozygous

- Considerable variation
  - between chromosomes
  - between individuals
  - between species (depends on genome length!)

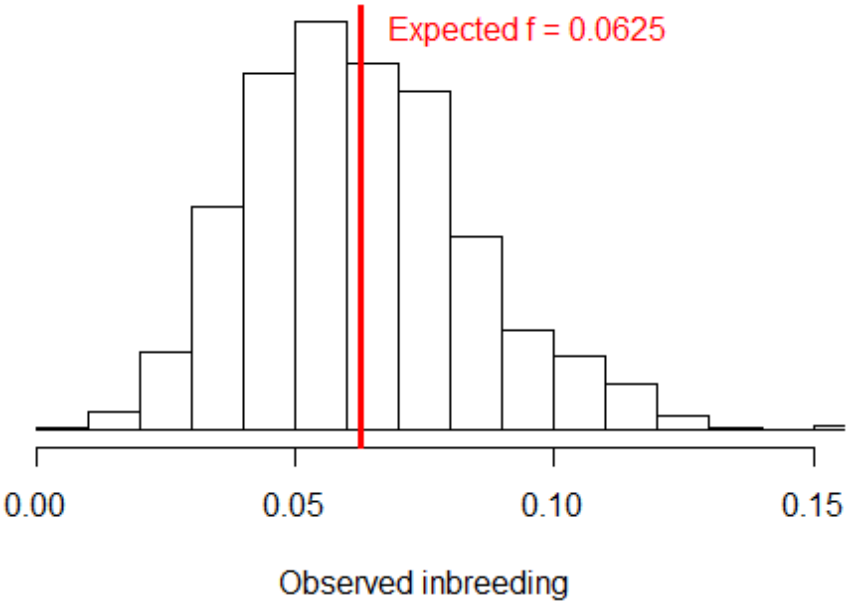


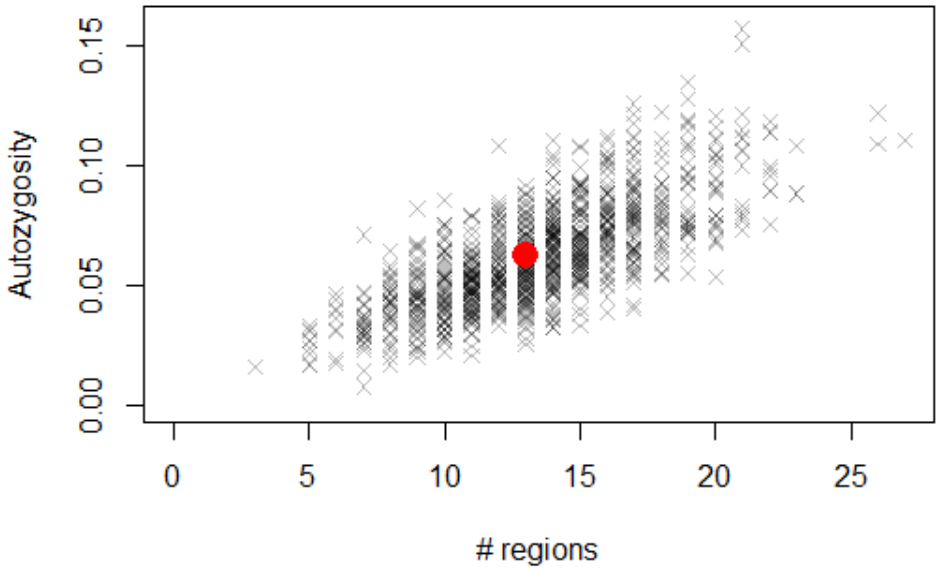
Realised inbreeding: natural variation



1000 simulations

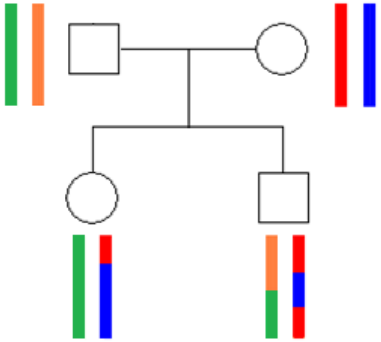
- R-package: **ibdsim2**
- Decode recombination map
- All 22 human autosomes





Distribution of realised IBD coefficients

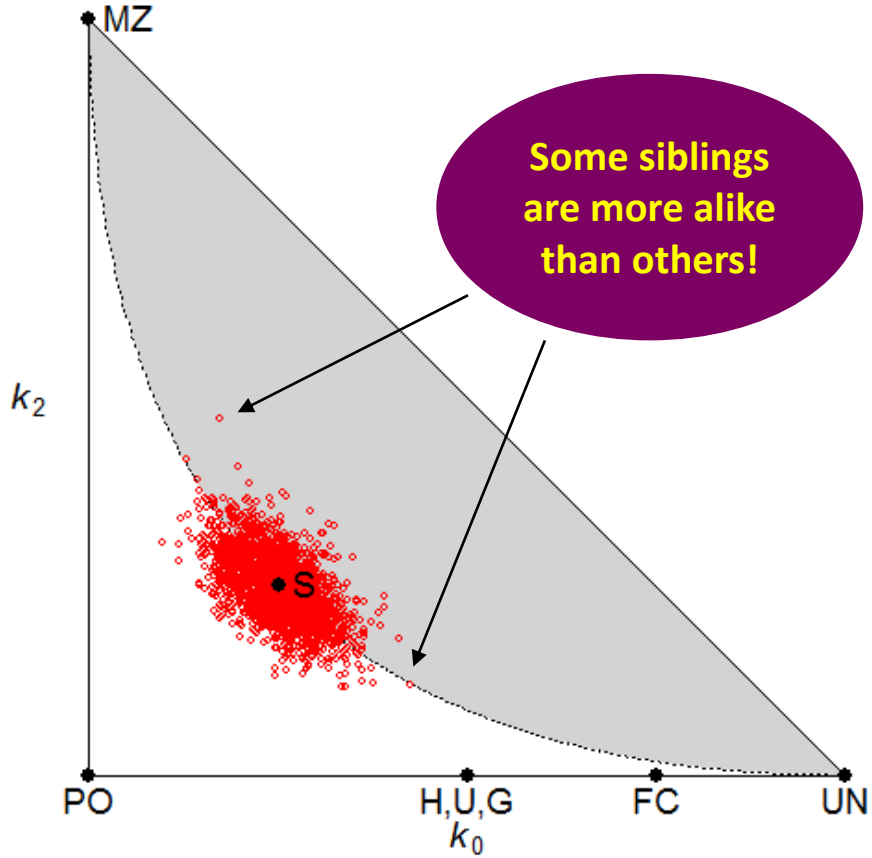
1000 simulations



```
library(ibdsim2)

x = nuclearPed(2)
s = ibdsim(x, sims = 1000)
k = realised_kappa(s, id.pair = 3:4)

forrel::showInTriangle(k, ...)
```



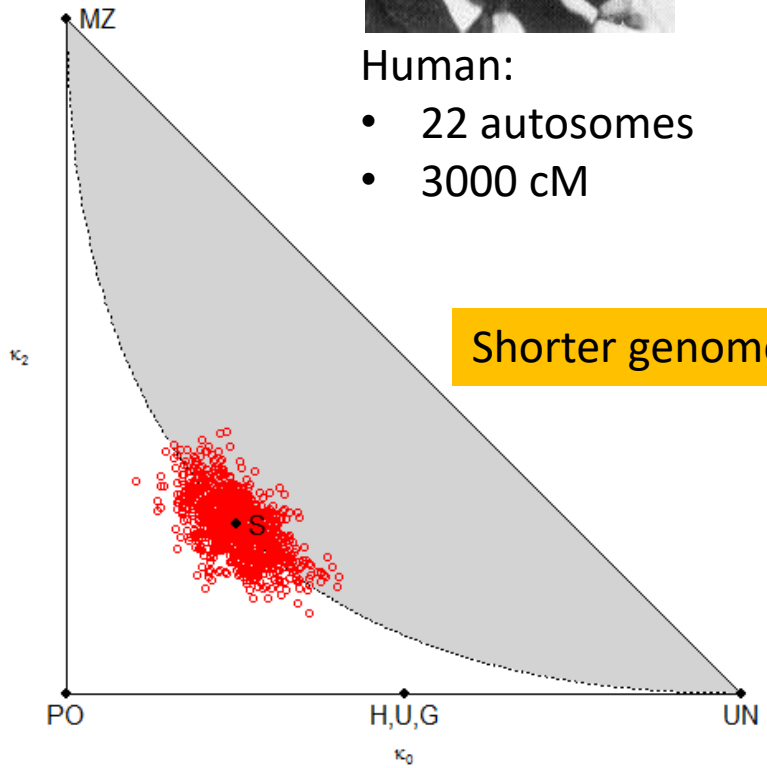
Variation depends on the genome

Human:

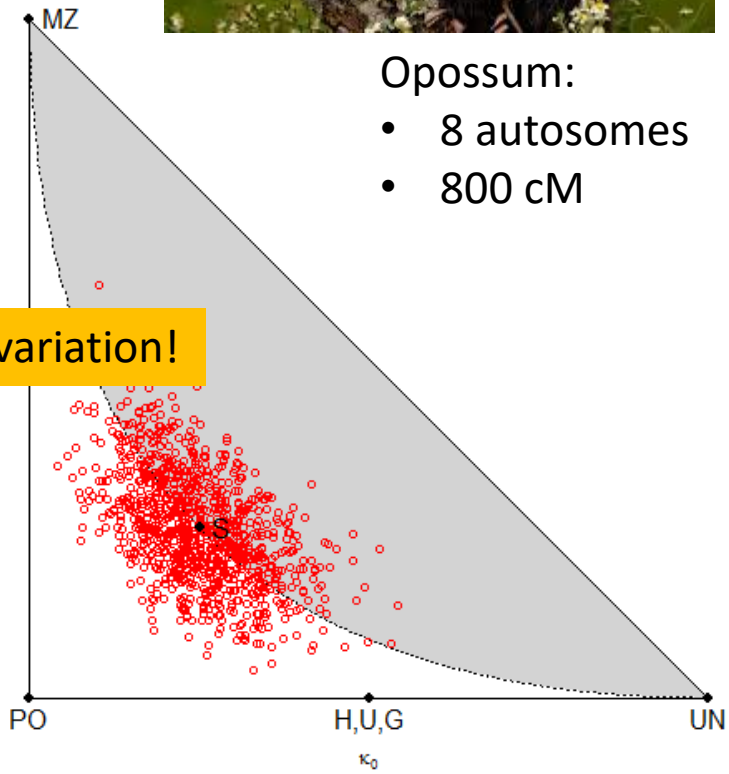
- 22 autosomes
- 3000 cM

Opossum:

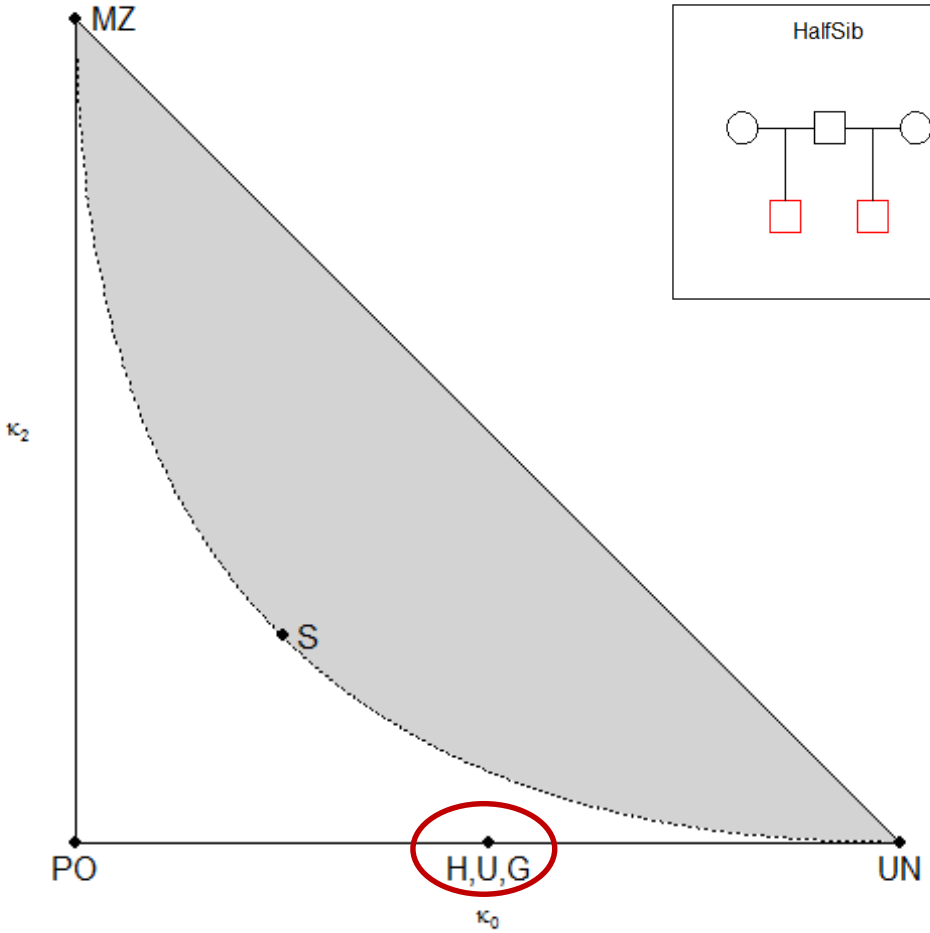
- 8 autosomes
- 800 cM

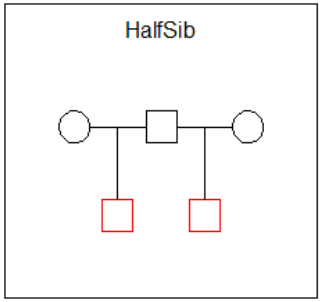


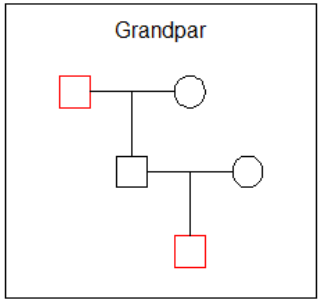
Shorter genome = more variation!

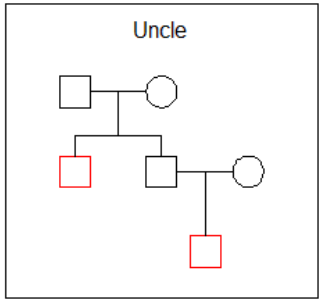


Indistinguishable relationships?



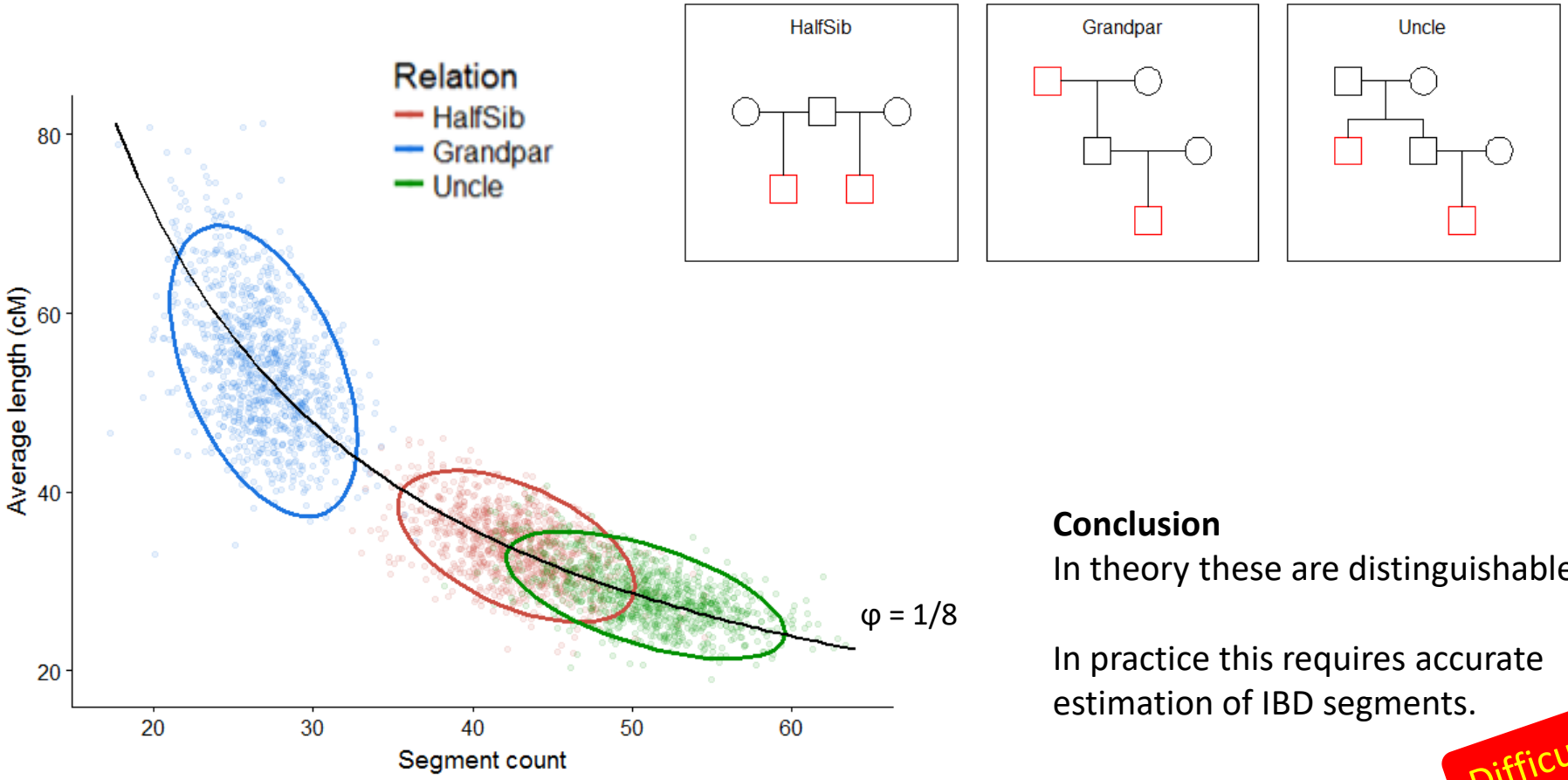






$$\kappa_0 = 0.5$$
$$\kappa_1 = 0.5$$
$$\kappa_2 = 0$$

Simulated IBD distributions



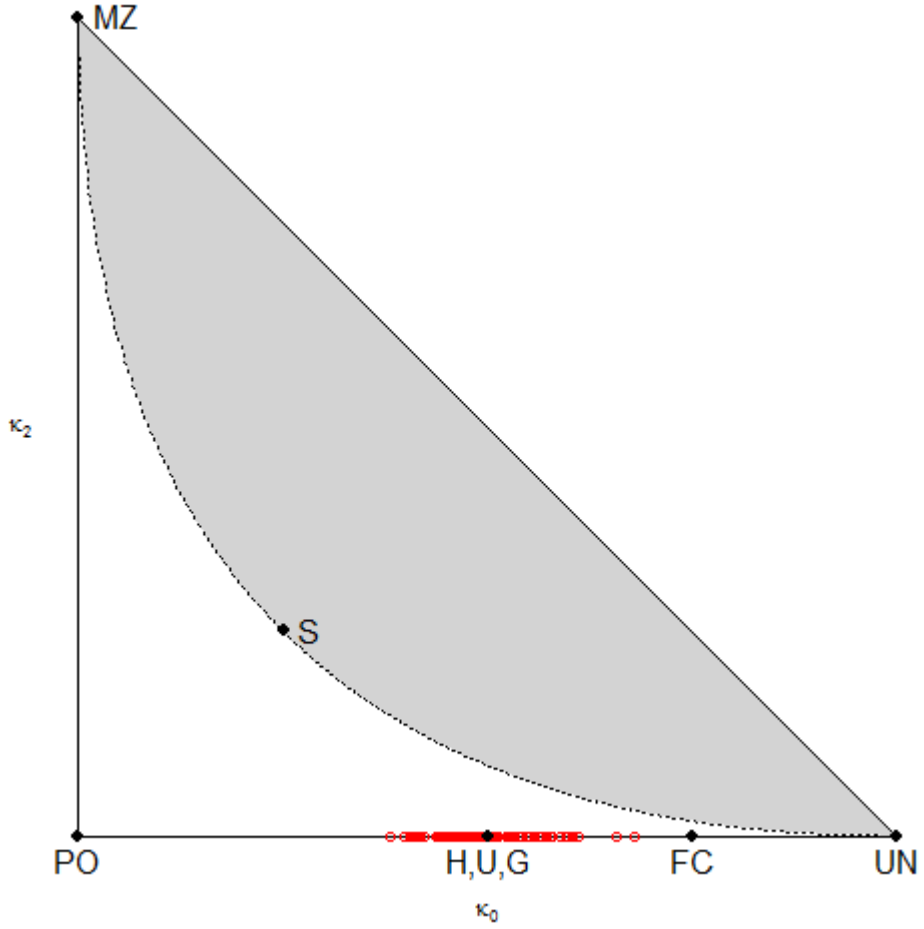
Conclusion

In theory these are distinguishable!

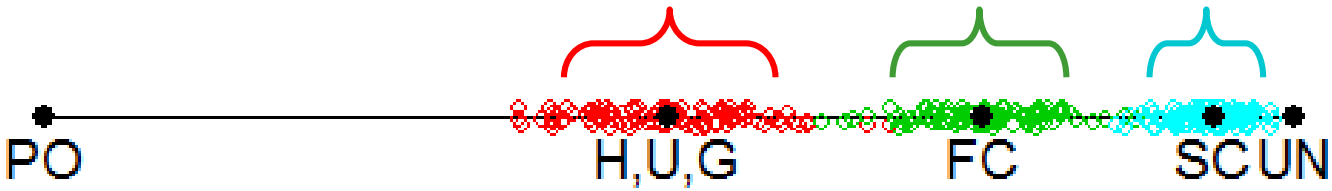
In practice this requires accurate estimation of IBD segments.

Difficult!

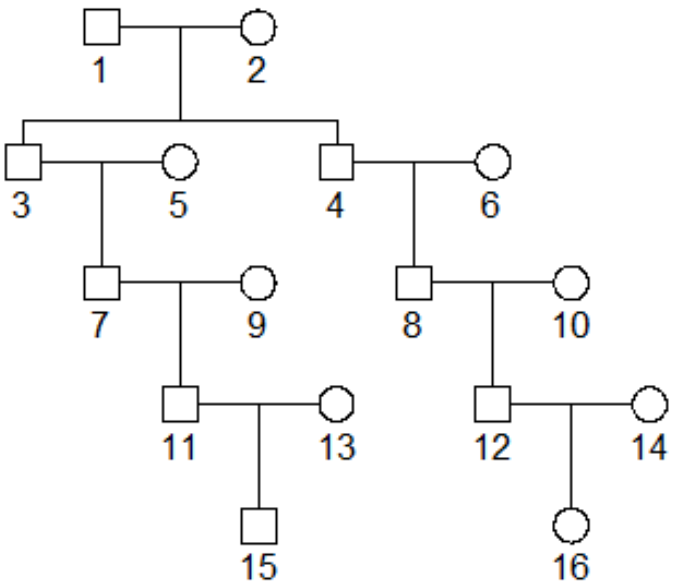
Realised IBD coefficients: Half siblings



Zooming in on the κ_0-axis



The probability of zero IBD



Theoretically possible to have no IBD sharing!

N'th cousins	*P*(zero IBD)
first	0.0 %
second	0.0 %
third	1.5 %
fourth	28 %
fifth	67 %

Third cousins: Expected fraction of the genome with IBD = 1:

$$k_1 = \frac{1}{64}$$

Two individuals can have a common ancestor without being genetically related

Relatedness: Summary

- Measuring relatedness with increasing precision:
  - the kinship/inbreeding coefficient φ
  - the IBD coefficients $\kappa = (\kappa_0, \kappa_1, \kappa_2)$
  - Jacquard's 9 coefficients Δ

- Each coefficient is
  - the probability of observing a certain IBD pattern in a random locus
  - the expected proportion of the genome in this state

- IBD is not a pointwise phenomenon: Always in segments
  - determined by meiotic crossovers
  - consequence: Variation in the *realised* IBD!

- Family relation $\not\Rightarrow$ genetic relation

So...what does it mean to be related?

- Pedigree based definition: $\varphi > 0$

  potentially having alleles IBD

- Genomic definition (**realised** relatedness):

  actually having alleles IBD