Lecture 4: Coefficients of relatedness
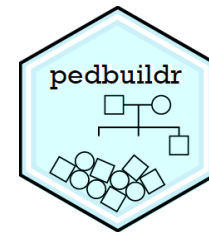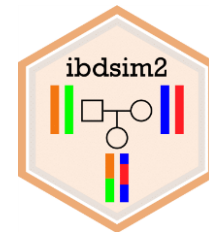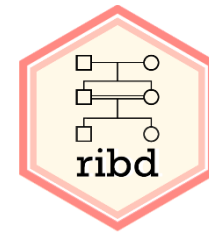
Pedigree analysis in R

ISFG Summer School - Virtual Edition 2021

Magnus Dehli Vigeland

Plan

- Part 1: Measures of relatedness

- Part 2: Realised relatedness

- Part 3: Inference of pairwise relatedness

- Part 4: Pedigree reconstruction

Plan: Today
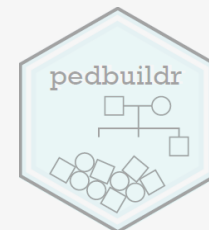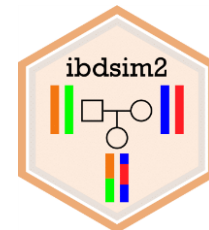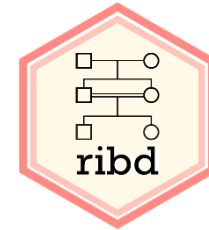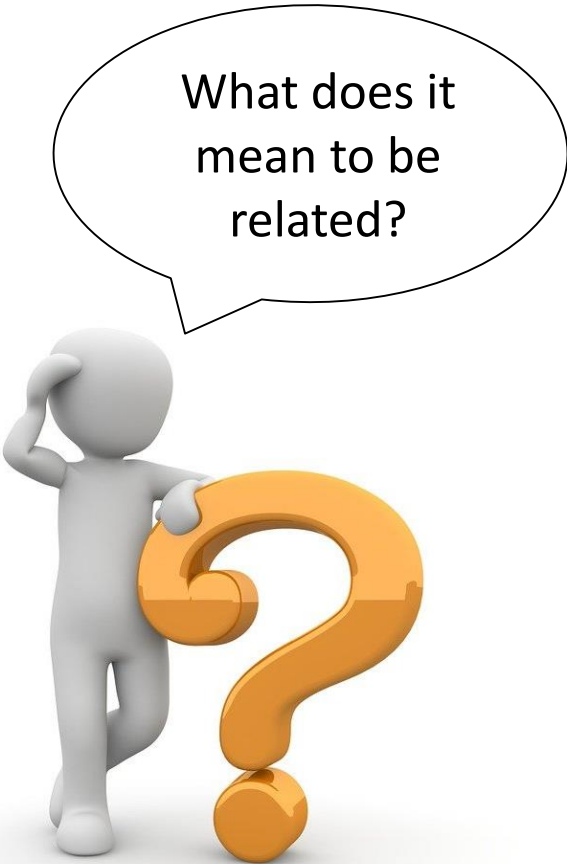
- Part 1: Measures of relatedness

- Part 2: Realised relatedness

- Part 3: Inference of pairwise relatedness

- Part 4: Pedigree reconstruction

Relatedness part 1:
Measures of relatedness

What does it mean to be related?

- Attempt 1
  - being connected through a pedigree
  - having a common ancestor...not too far back

- Attempt 2 - genetic
  - sharing DNA?
  - (more than unrelated people)

- To make this precise, we need some terminology!
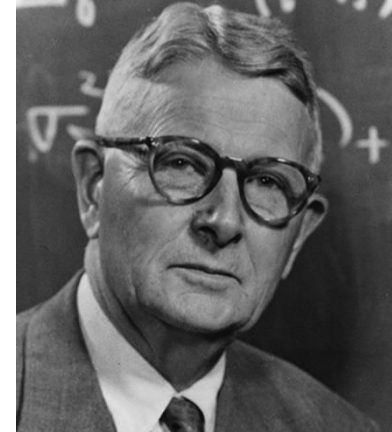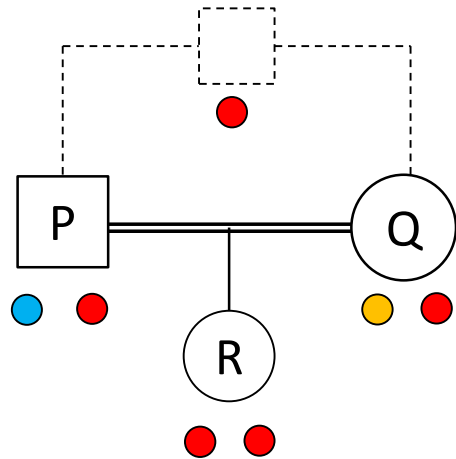
IBD and autozygosity

- IBD = Identical by descent

    = identical alleles with a
        common origin **in the
        given pedigree**

- autozygous = homozygous + IBD
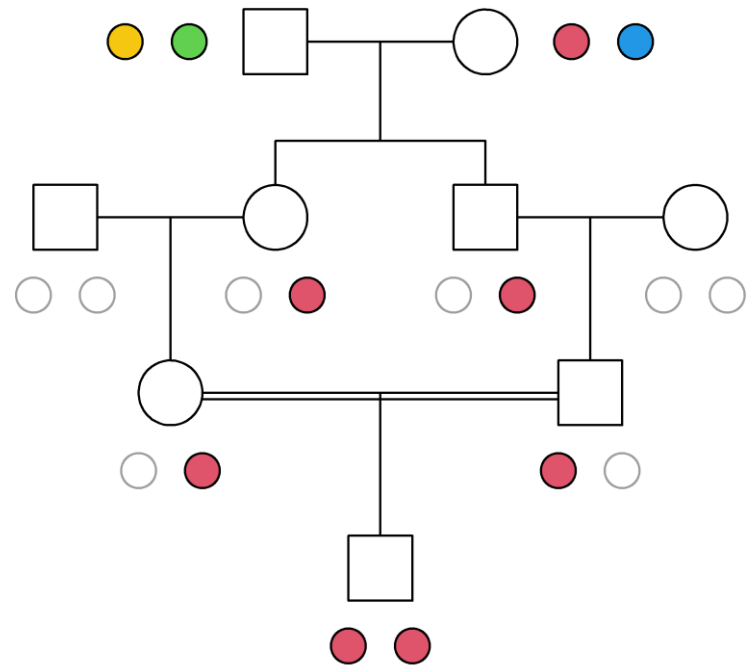
Inbreeding coefficient
$f = P(autozygosity)$

Coefficient of kinship/inbreeding

Sewall Wright
(1889 - 1988)

- Wright (1921): The kinship coefficient φ between P and Q

  $\varphi_{P,Q}$ = P(random allele of P is IBD with random allele of Q)

  = P(R receive IBD alleles from her parents)

  = P(R is autozygous)

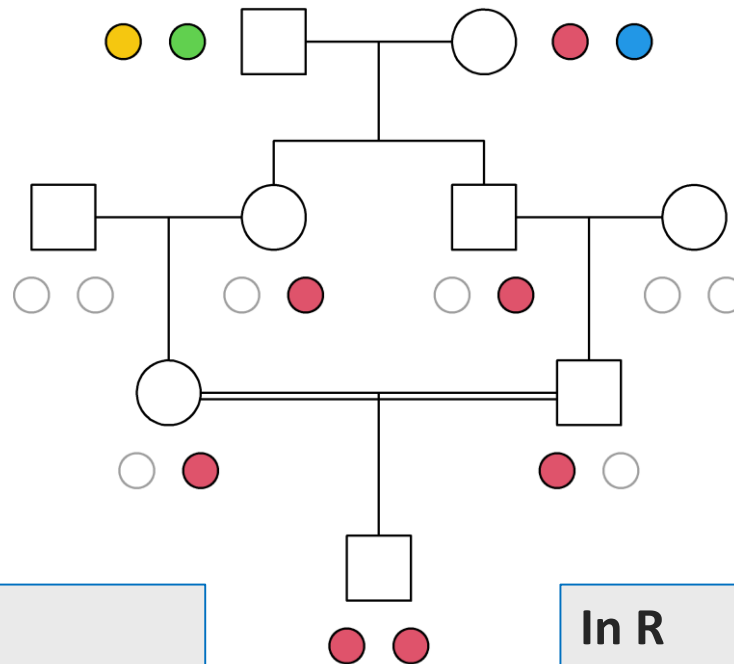 = f_R the inbreeding coefficient of R

P and Q related

$\updownarrow$

$\varphi_{P,Q} > 0$

Inbreding coefficient: Example

By hand

$$f = P(\text{\textcolor{red}{\bullet}} \text{ autozygous}) \cdot 4$$
$$= 0.5^6 \cdot 4$$
$$= 1/16$$

other colors

In R

```
>   library(ribd)

>   x = cousinPed(1, child = T)
>   inbreeding(x, ids = 9)
[1] 0.0625
```

Oslo
universitetssykehus

More kinship & inbreeding coefficients

Relationship	Kinship	f (of child)
Parent-child	1/4	1/4
Full siblings	1/4	1/4
Half siblings	1/8	1/8
Grandparent-grandchild	1/8	1/8
Avuncular (uncle/aunt)	1/8	1/8
1st cousins	1/16	1/16
2nd cousins	1/64	1/64
3rd cousins	1/256	1/256

father/daughter incest:
$f = 0.25$
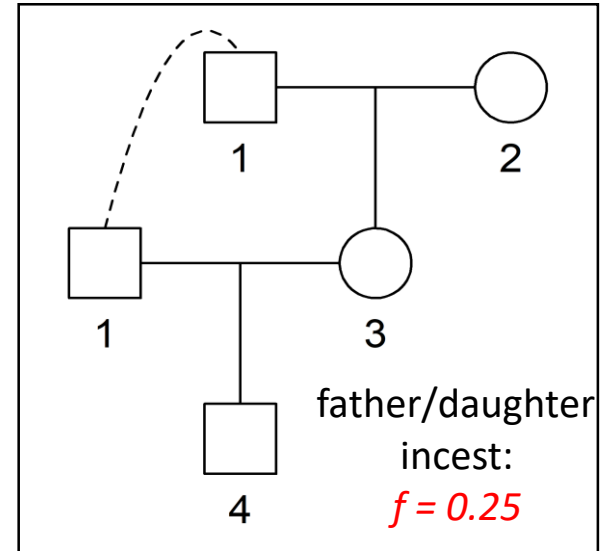
```
>   x = nuclearPed(1, sex = 2)
>   kinship(x, ids = c(1, 3))
[1] 0.25

>   x = addchildren(x, 1, 3, nch = 1)
>   inbreeding(x, 4)
[1] 0.25
```
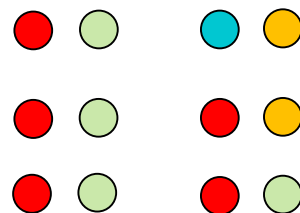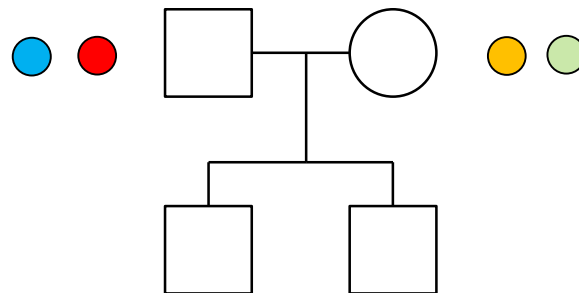
The IBD triangle
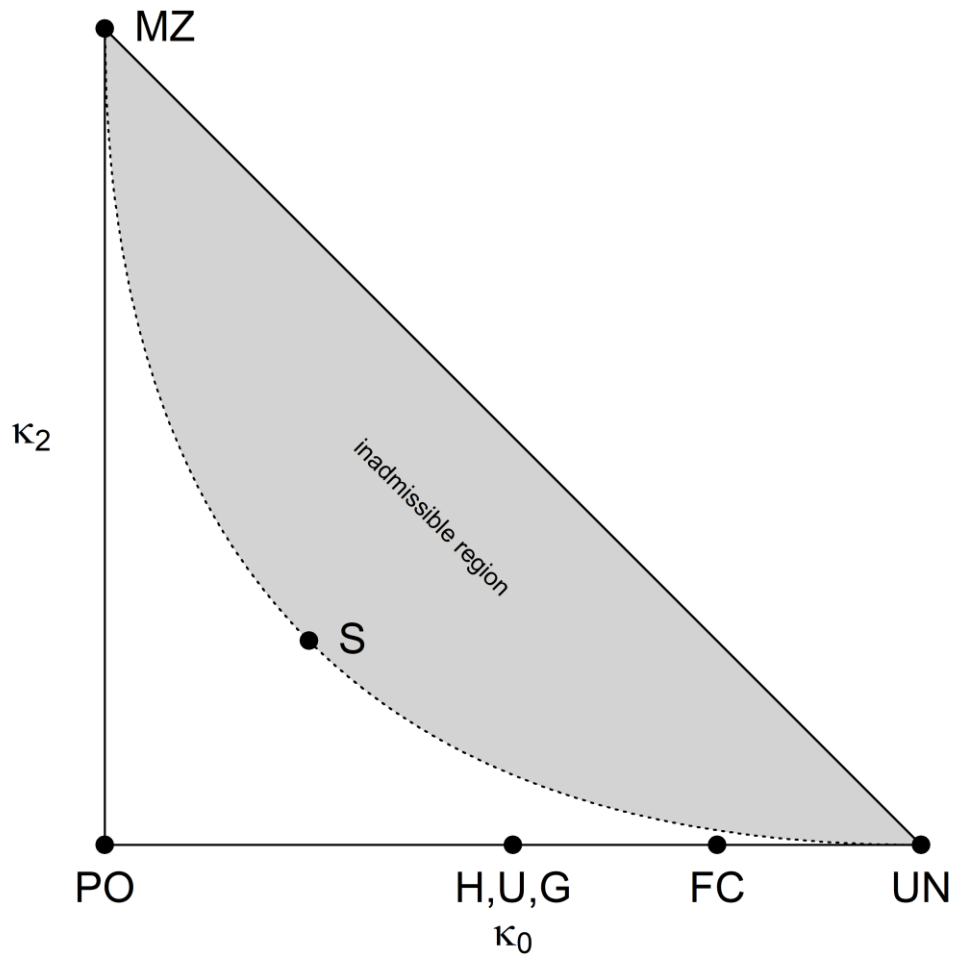
Charles Cotterman
(1914-1989)

Elisabeth Thompson
(1949 -)

IBD coefficients: Warm-up

- Summary so far:
  - Two individuals are related if they can have IBD alleles
  - Their kinship coefficient meassures the amount of IBD sharing
- Natural generalisation:
  - How *many* alleles are IBD in each locus?

Humans are diploid

IBD = 0, 1 or 2

← no IBD sharing

← share ●

← share ● and ○

Oslo
universitetssykehus

IBD coefficients: Definition
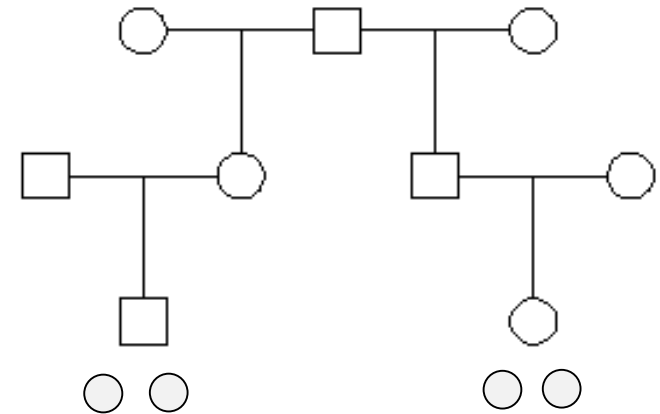

- Given two (non-inbred) individuals

- For a random autosomal locus

  $\kappa_0 = P(0\ \text{alleles IBD})$
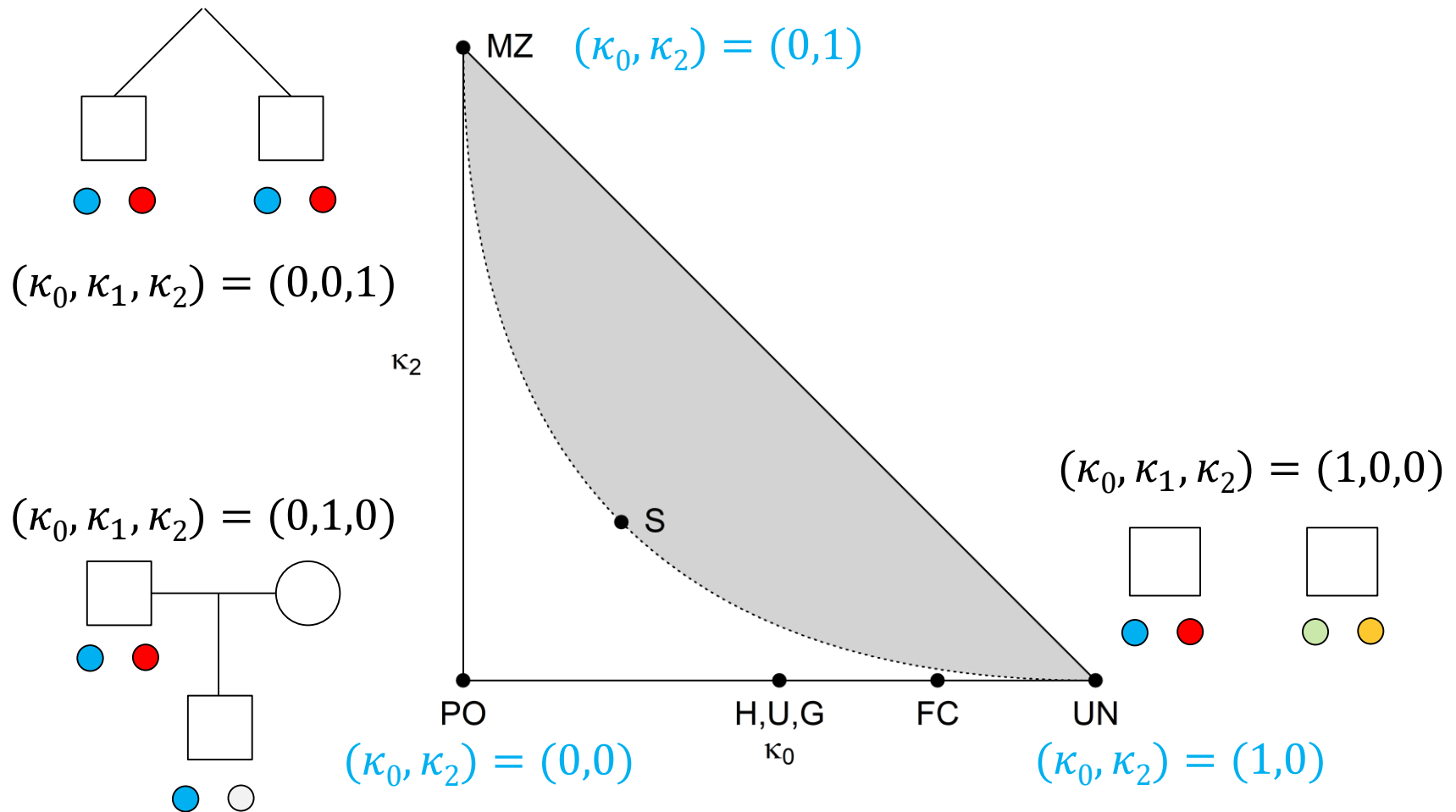
  $\kappa_1 = P(1\ \text{alleles IBD})$

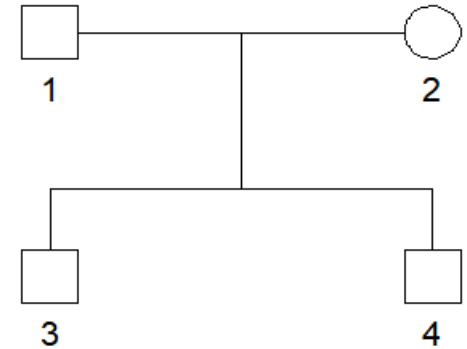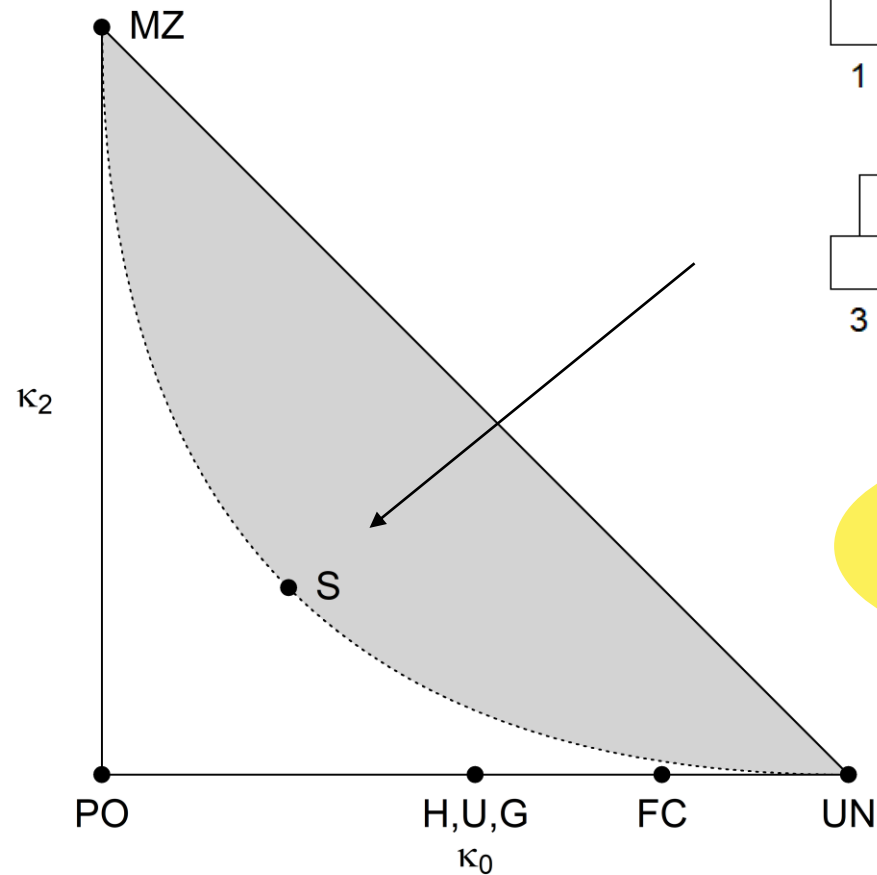 $\kappa_2 = P(2\ \text{alleles IBD})$

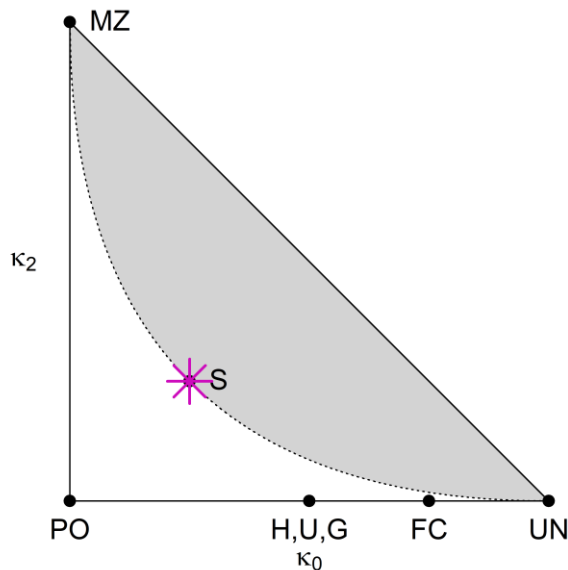- We always have: $\kappa_0 + \kappa_1 + \kappa_2 = 1$
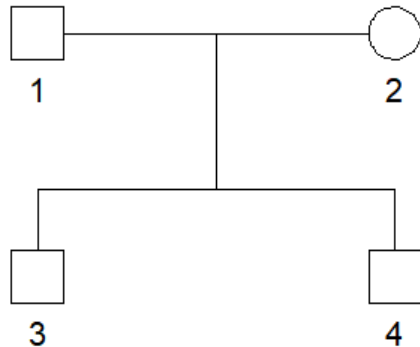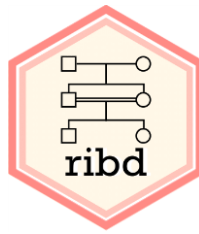
The relatedness triangle

$(\kappa_0, \kappa_1, \kappa_2) = (0,0,1)$

$(\kappa_0, \kappa_2) = (0,1)$

MZ

κ_2

S

$(\kappa_0, \kappa_1, \kappa_2) = (0,1,0)$

$(\kappa_0, \kappa_1, \kappa_2) = (1,0,0)$

PO

H,U,G

κ_0

FC

UN

$(\kappa_0, \kappa_2) = (0,0)$

$(\kappa_0, \kappa_2) = (1,0)$

Oslo
universitetssykehus

What are the coefficients of full sibs

Let's do this in R!

ribd: Pedigree-based relatedness coefficients
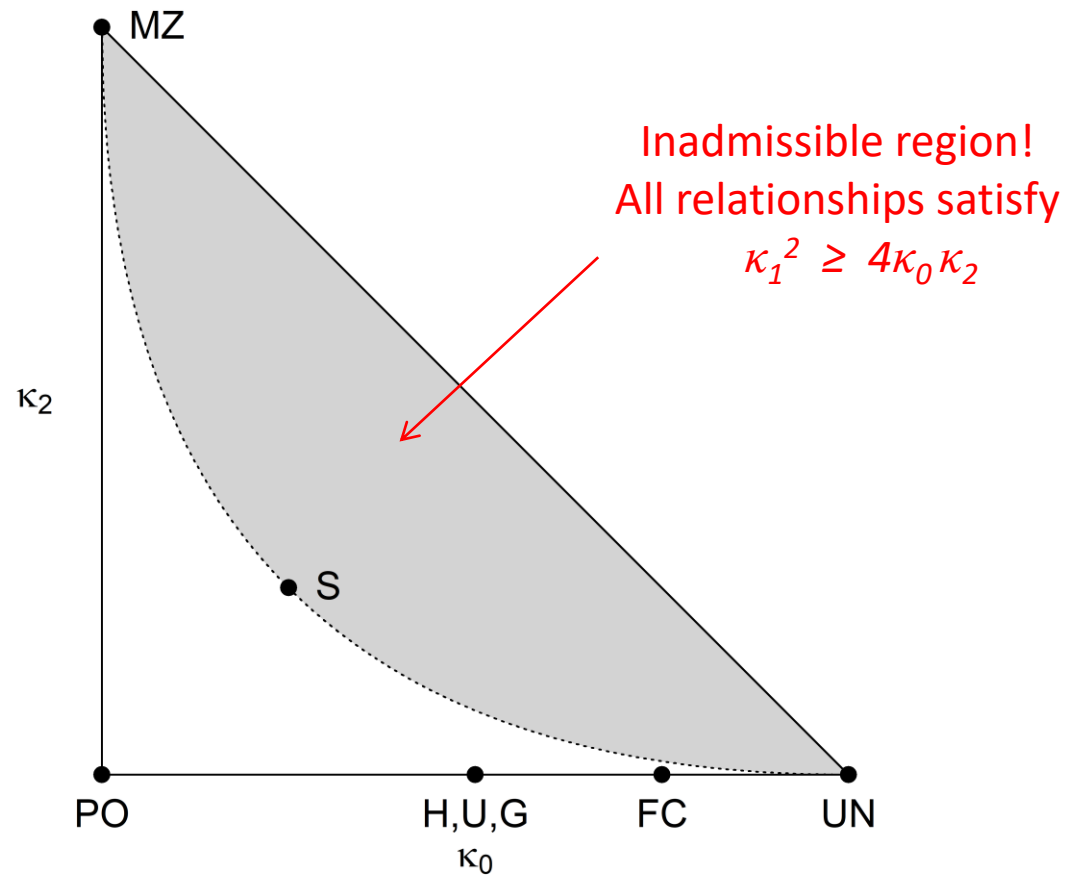


```
>   library(ribd)
>   x = nuclearPed(2)

>   kinship(x, ids = 3:4)
[1] 0.25

>   kappaIBD(x)
id1 id2 kappa0 kappa1 kappa2
  1   2   1.00    0.0    0.00
  1   3   0.00    1.0    0.00
  1   4   0.00    1.0    0.00
  2   3   0.00    1.0    0.00
  2   4   0.00    1.0    0.00
  3   4   0.25    0.5    0.25

>   k = kappaIBD(x, ids = 3:4)
>   showInTriangle(k, pch = 8, cex = 2)
```
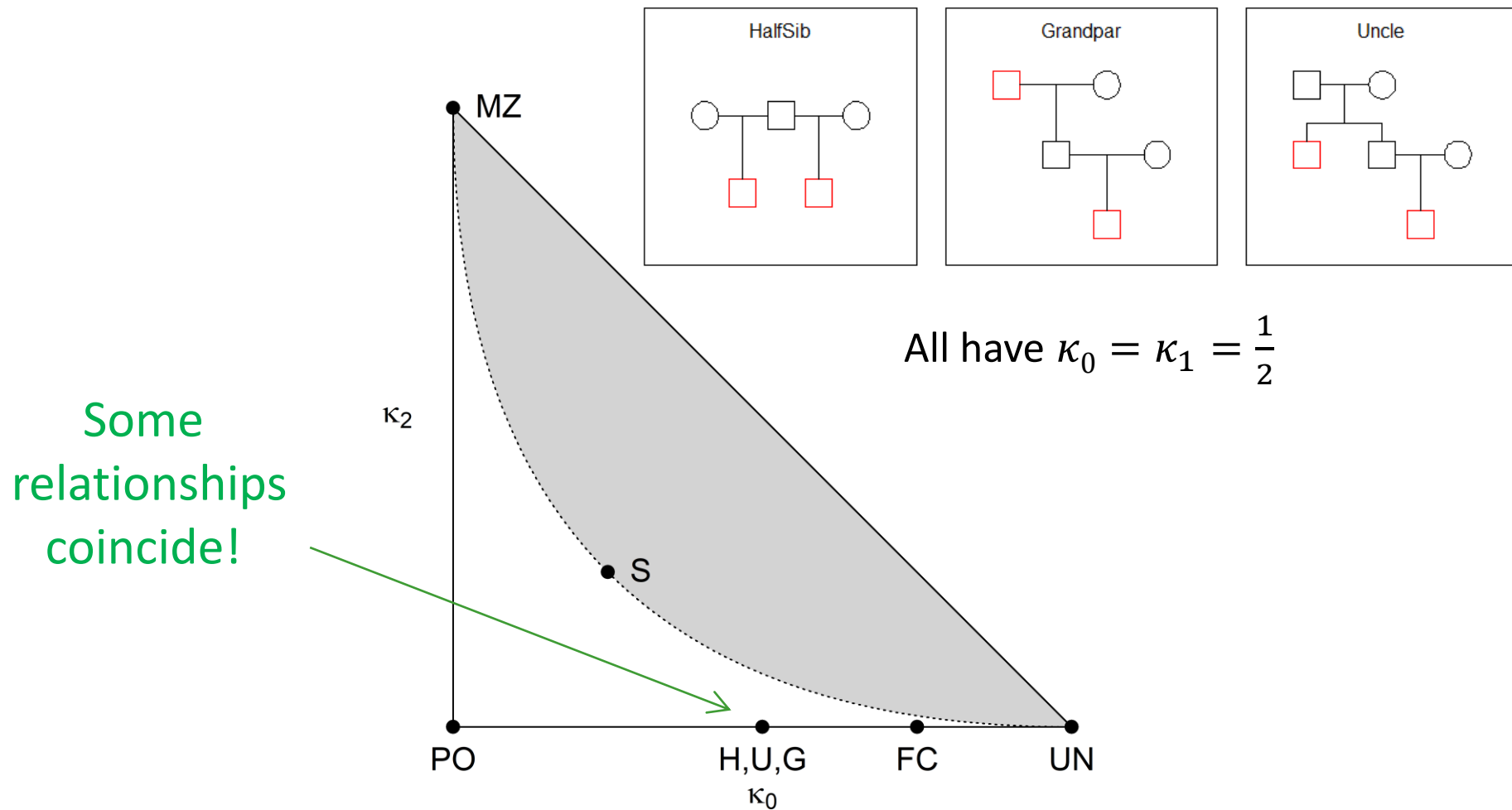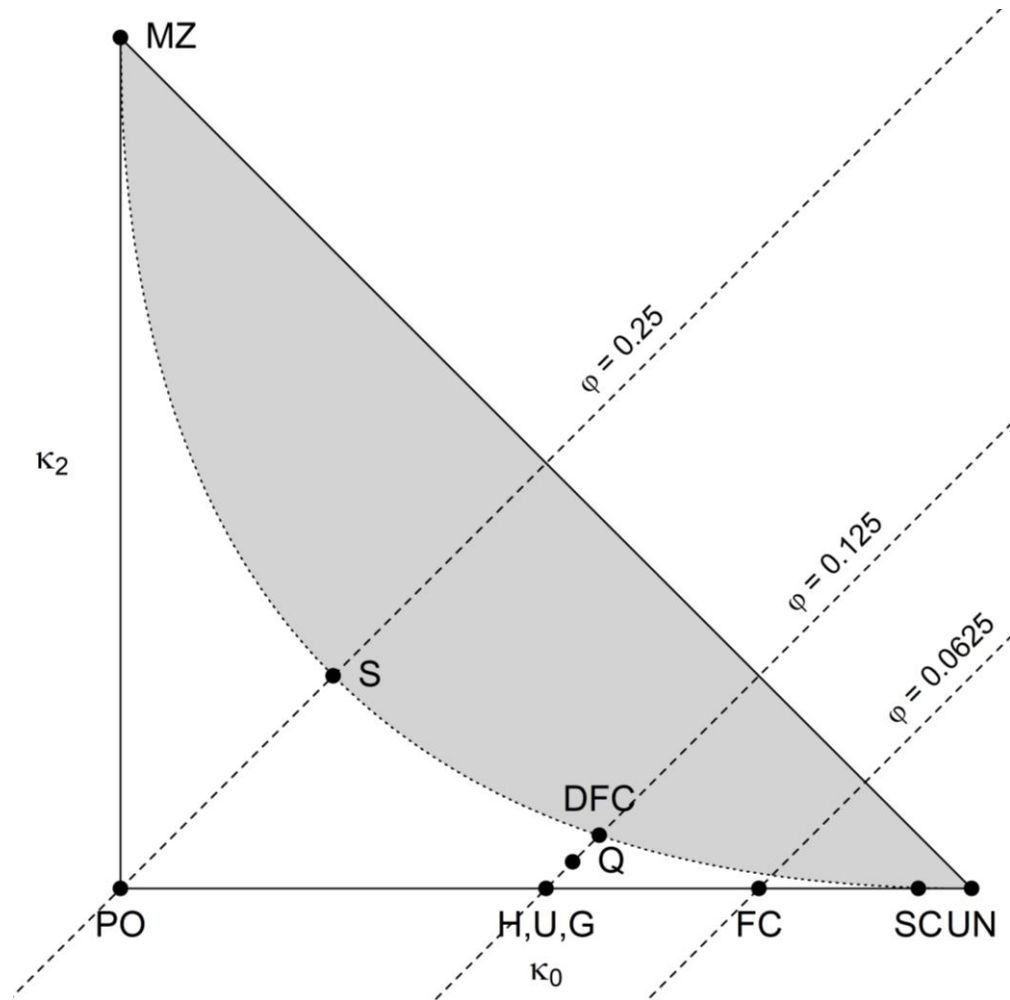
The relatedness triangle

Inadmissible region!
All relationships satisfy
$\kappa_1^2 \geq 4\kappa_0\kappa_2$

The relatedness triangle

All have $\kappa_0 = \kappa_1 = \frac{1}{2}$

Some relationships coincide!

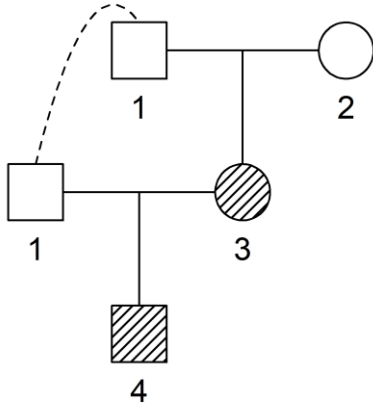κ_2

MZ

S

PO   H,U,G   FC   UN

κ_0

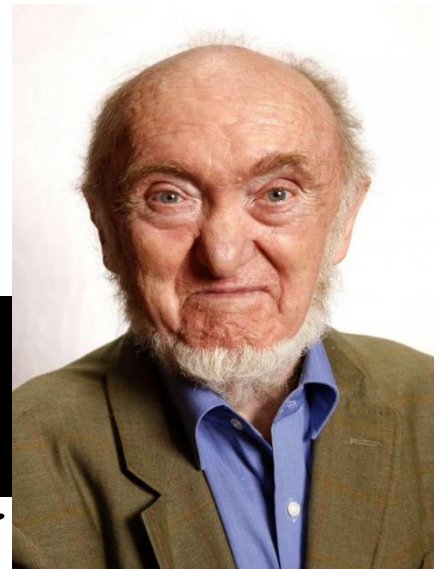# An important identity: $\varphi = \frac{1}{4}\kappa_1 + \frac{1}{2}\kappa_2$

Reminder

κ-coefficients are only defined for non-inbred individuals.
For the whole story, we need 9 coefficients!

Jacquard's identity coefficients

Only for the brave!

Covered in Chapter 3 of
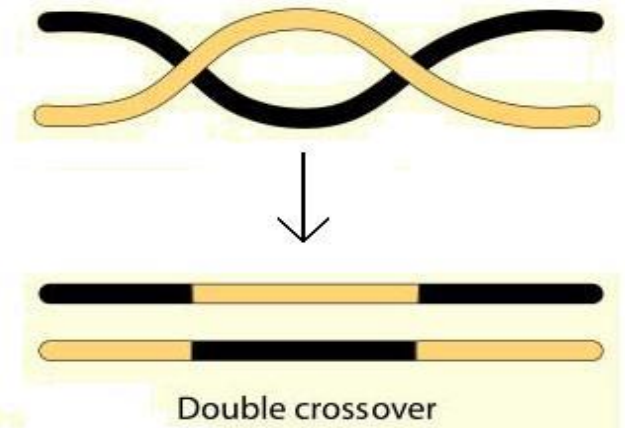*Pedigree analysis in R*

Albert Jacquard
(1925 - 2013)
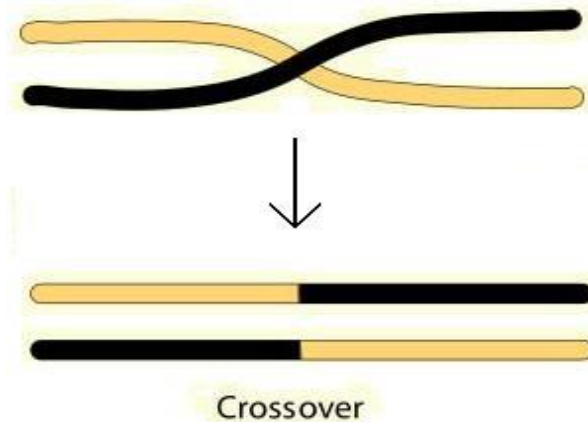
Relatedness part 2:

Realised relatedness
or
Why are some siblings more alike than others?

Recombination

- Crossovers:

Crossover

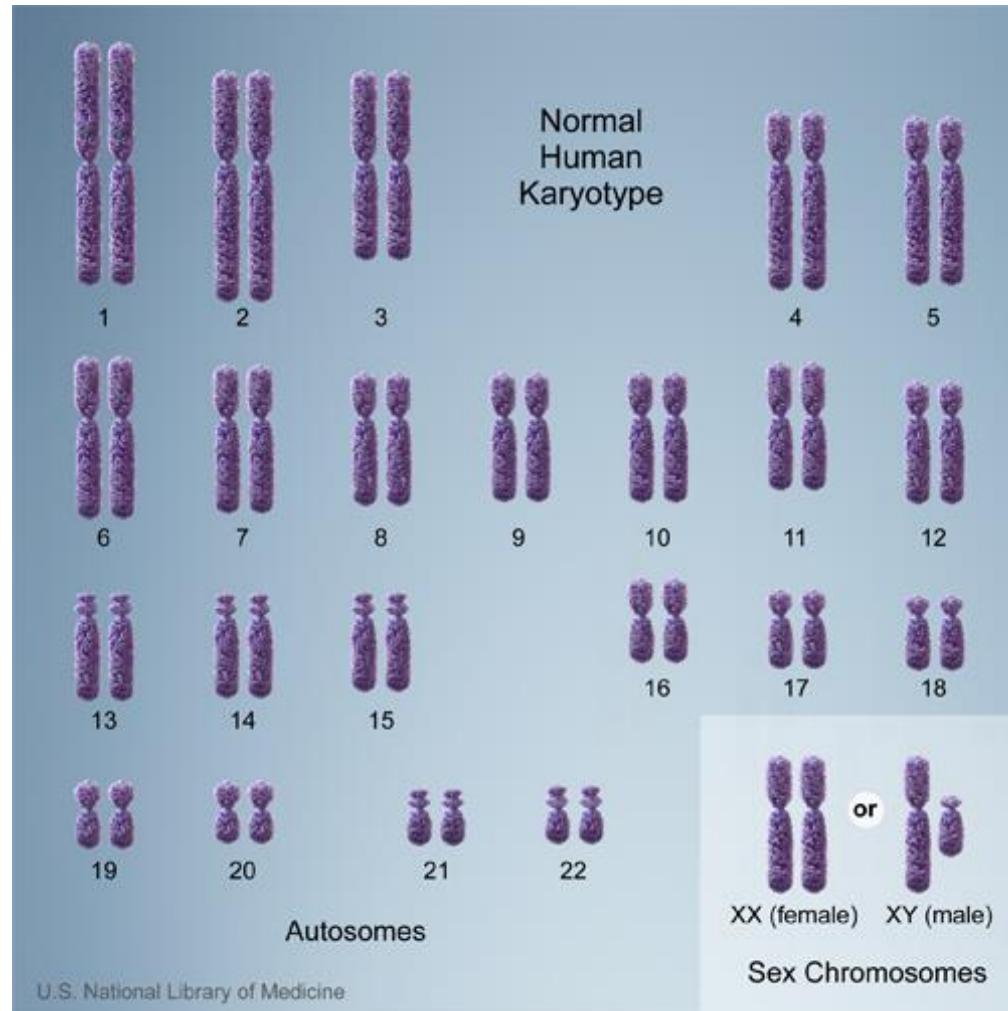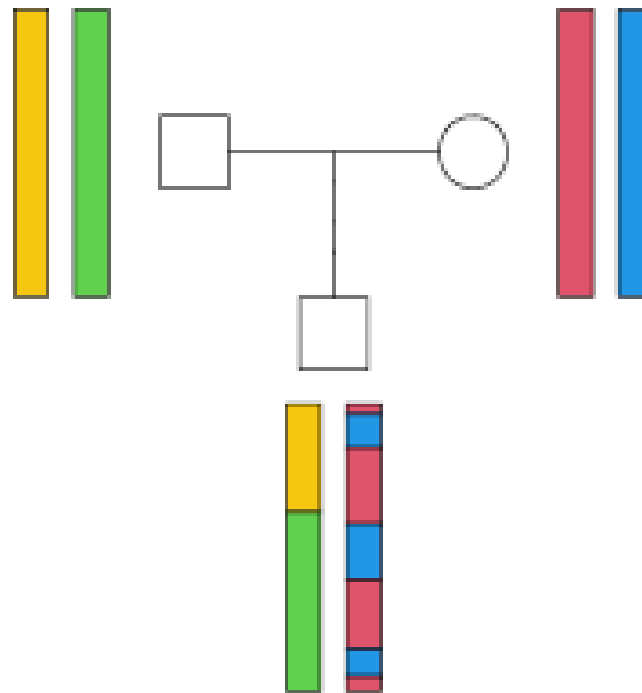Double crossover

- The genetic distance between two loci:

  = average number of crossovers between them per meiosis

- Units:

  – 1 Morgan (M) = 1 crossover per meiosis (on average)

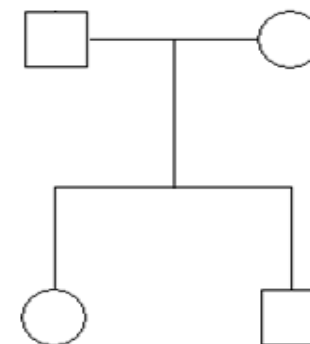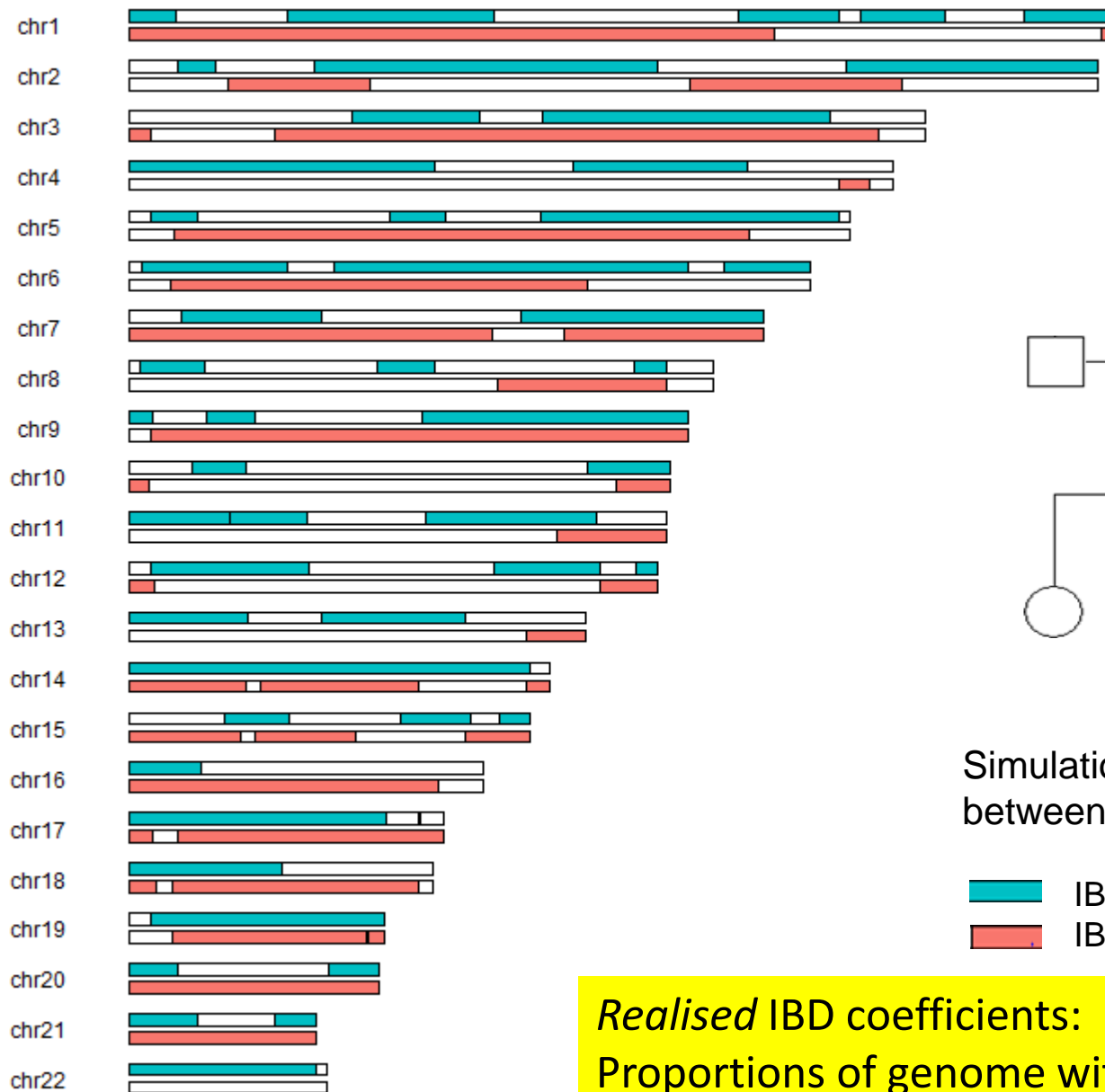 – 1 centiMorgan (cM) = 0.01 M

- The human genome: Ca 30 Morgan

Rule of thumb: One crossover per chromosome arm

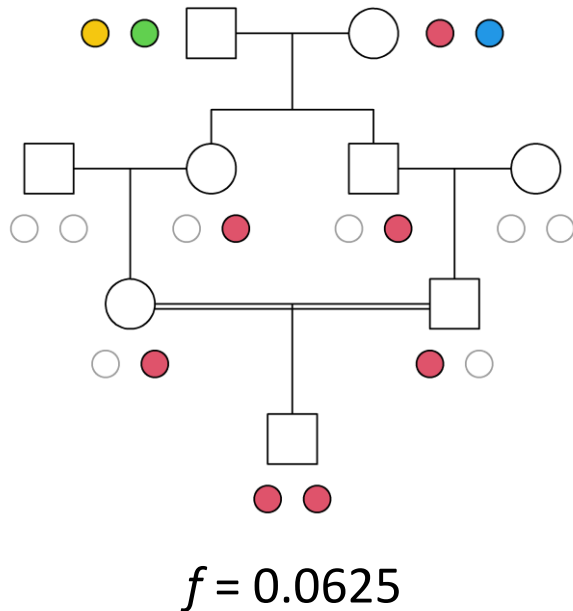# Recombination in practice

IBD: Full siblings

Simulation of IBD sharing between full siblings

IBD from father
IBD from mother

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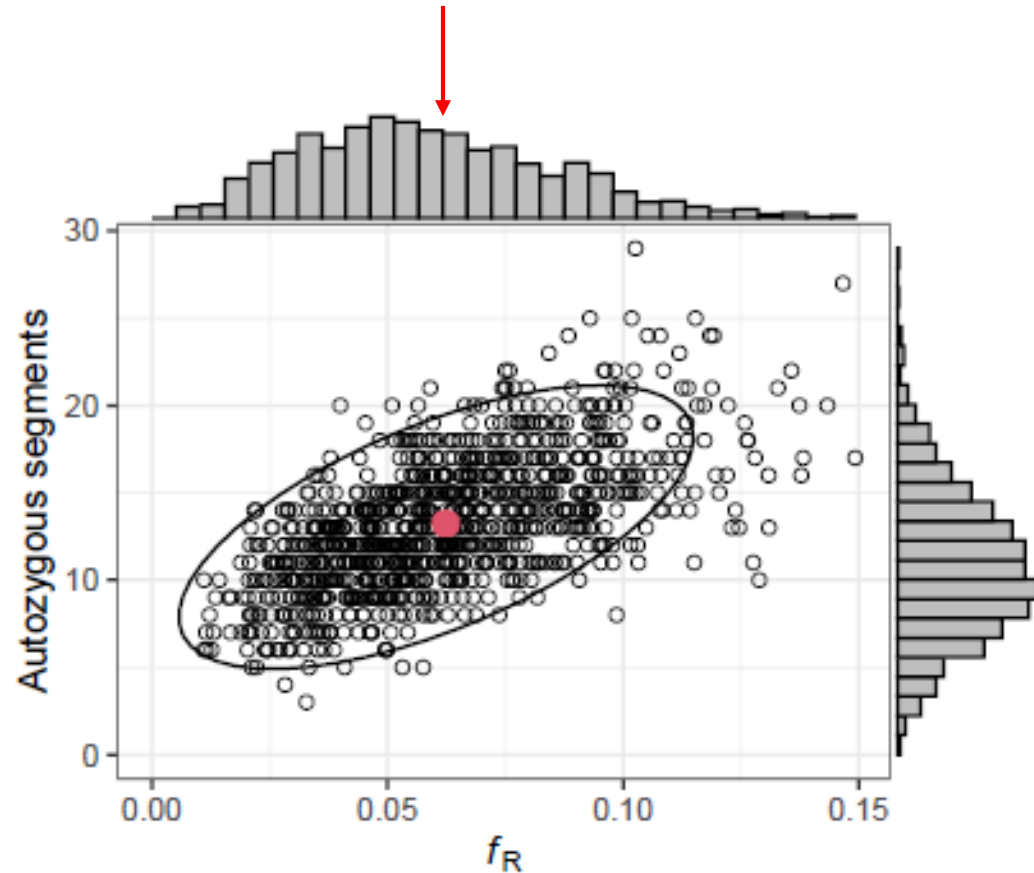Proportions of genome with IBD = 0, 1, 2

Variation in realised inbreeding

f = 0.0625

1000 simulations
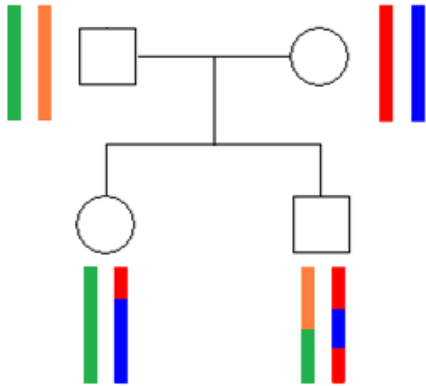- R-package: **ibdsim2**
- Decode recombination map
- All 22 human autosomes

Expected: f = 0.0625

Autozygous segments

f_R

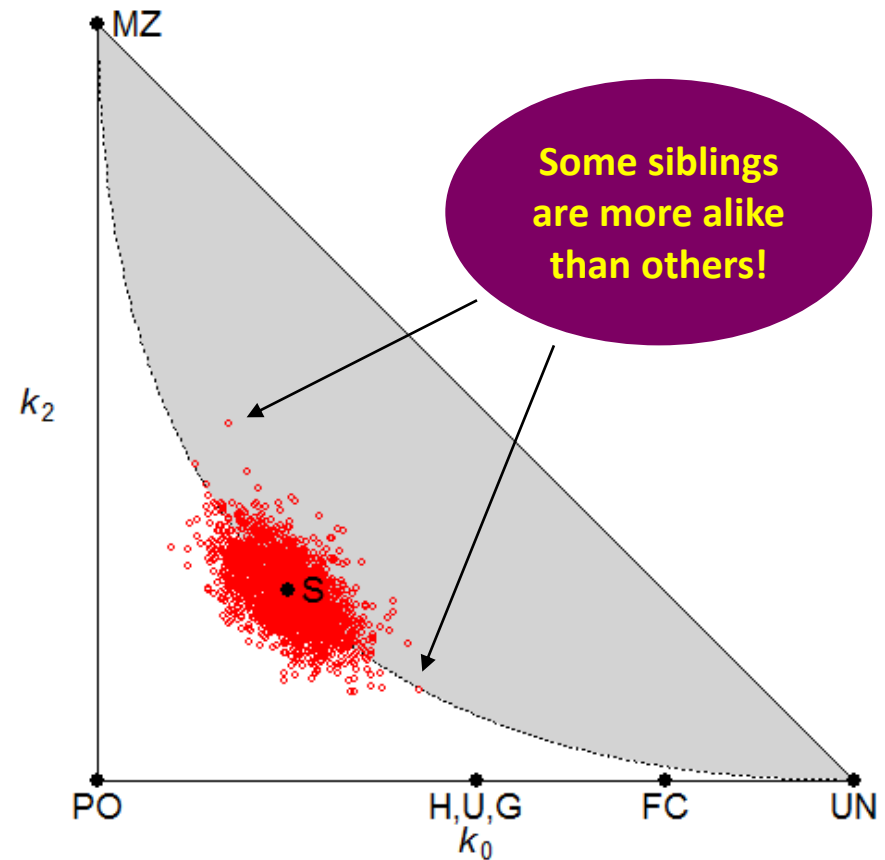# Variation in realised IBD coefficients

1000 simulations


```
>   library(ibdsim2)

>   x = nuclearPed(2)
>   s = ibdsim(x, N = 1000)
>   k = realisedKappa(s, ids = 3:4)

>   ribd::showInTriangle(k)
```
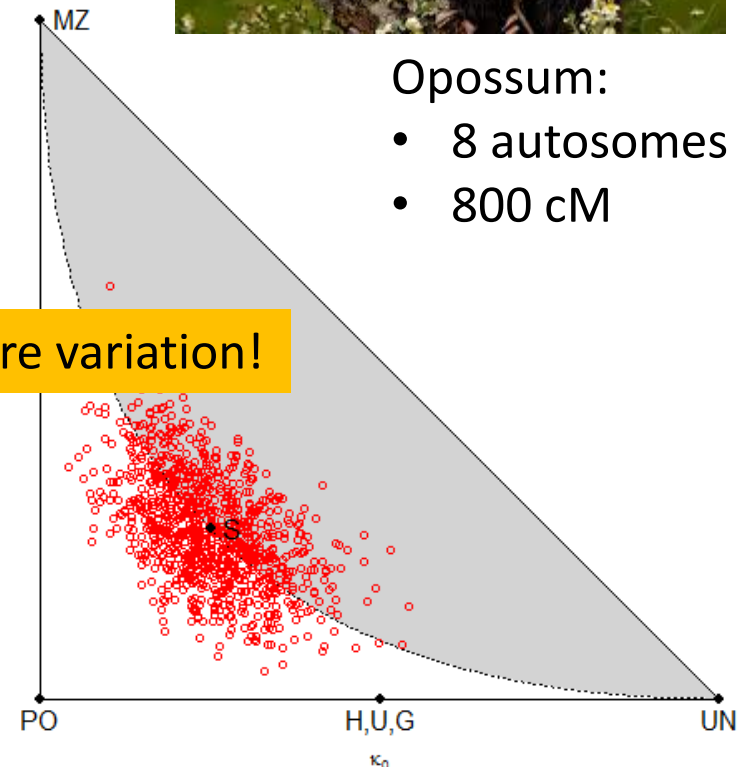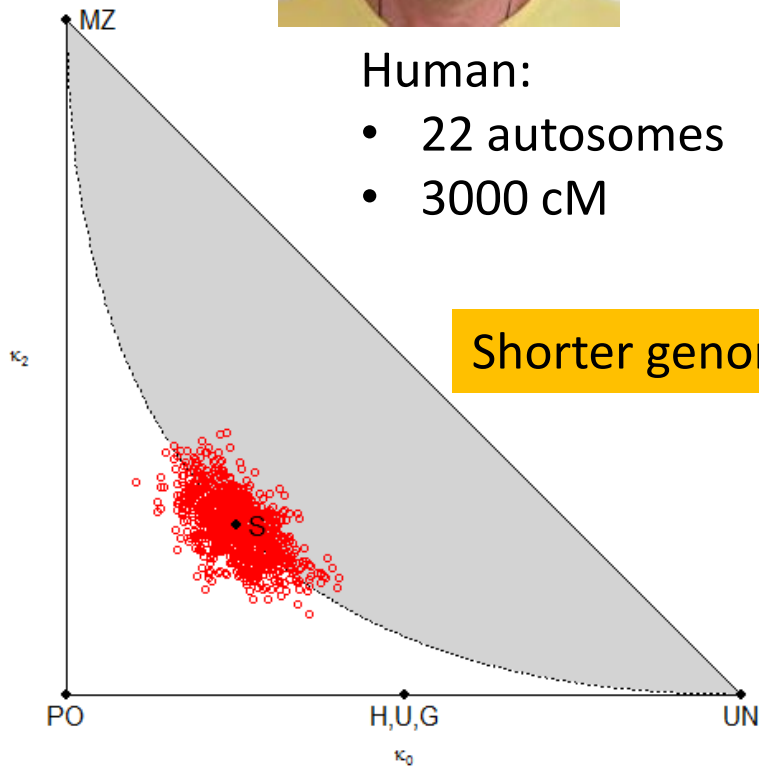
Some siblings are more alike than others!

Variation depends on the genome
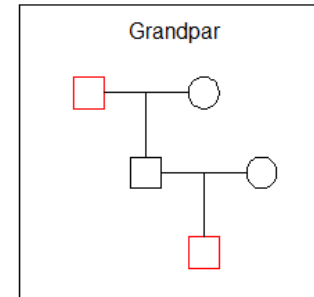

Human:
- 22 autosomes
- 3000 cM

Opossum:
- 8 autosomes
- 800 cM

Shorter genome = more variation!

Oslo
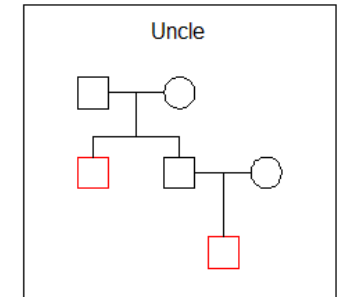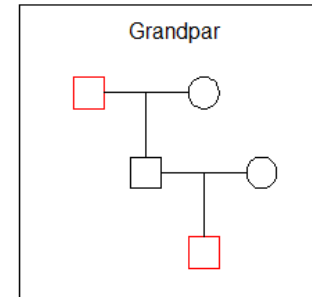universitetssykehus

Indistinguishable relationships?
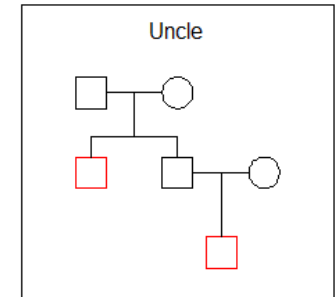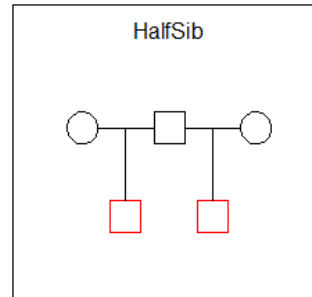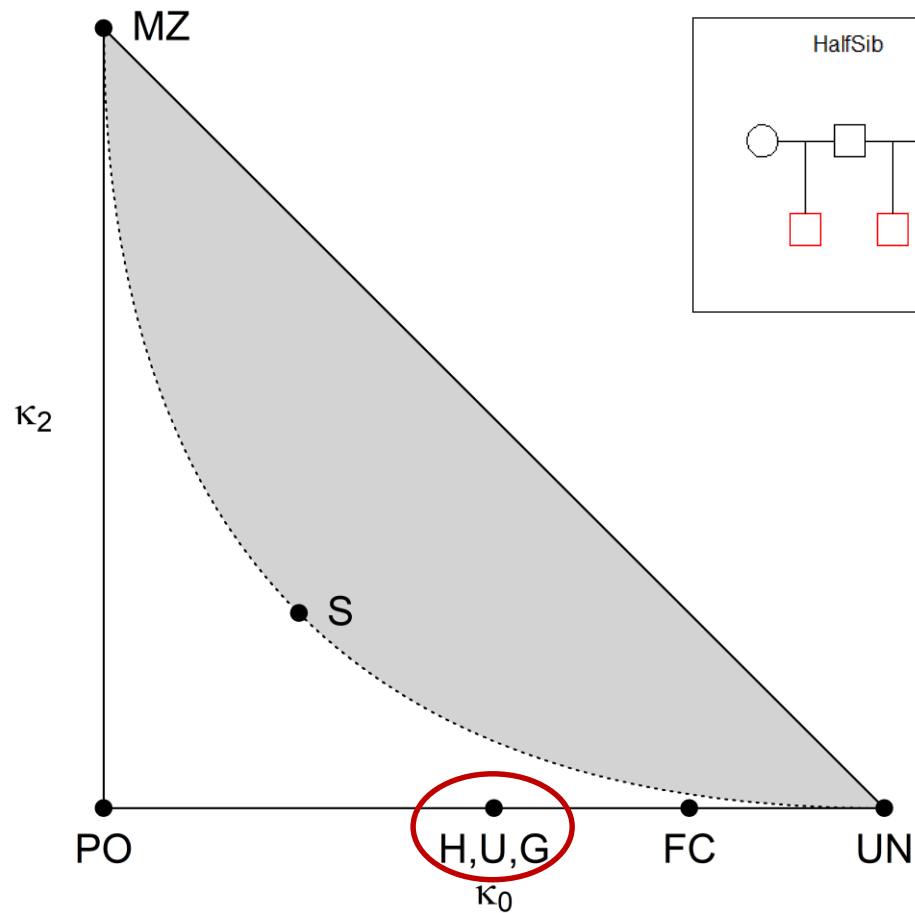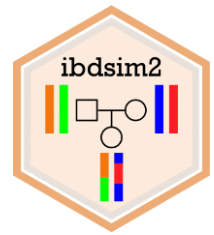
$$\kappa_0 = 0.5$$
$$\kappa_1 = 0.5$$
$$\kappa_2 = 0$$

Simulated IBD distributions
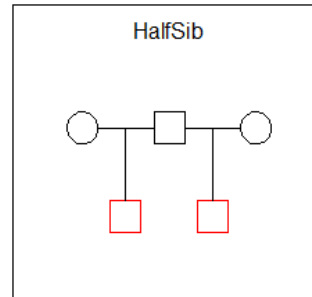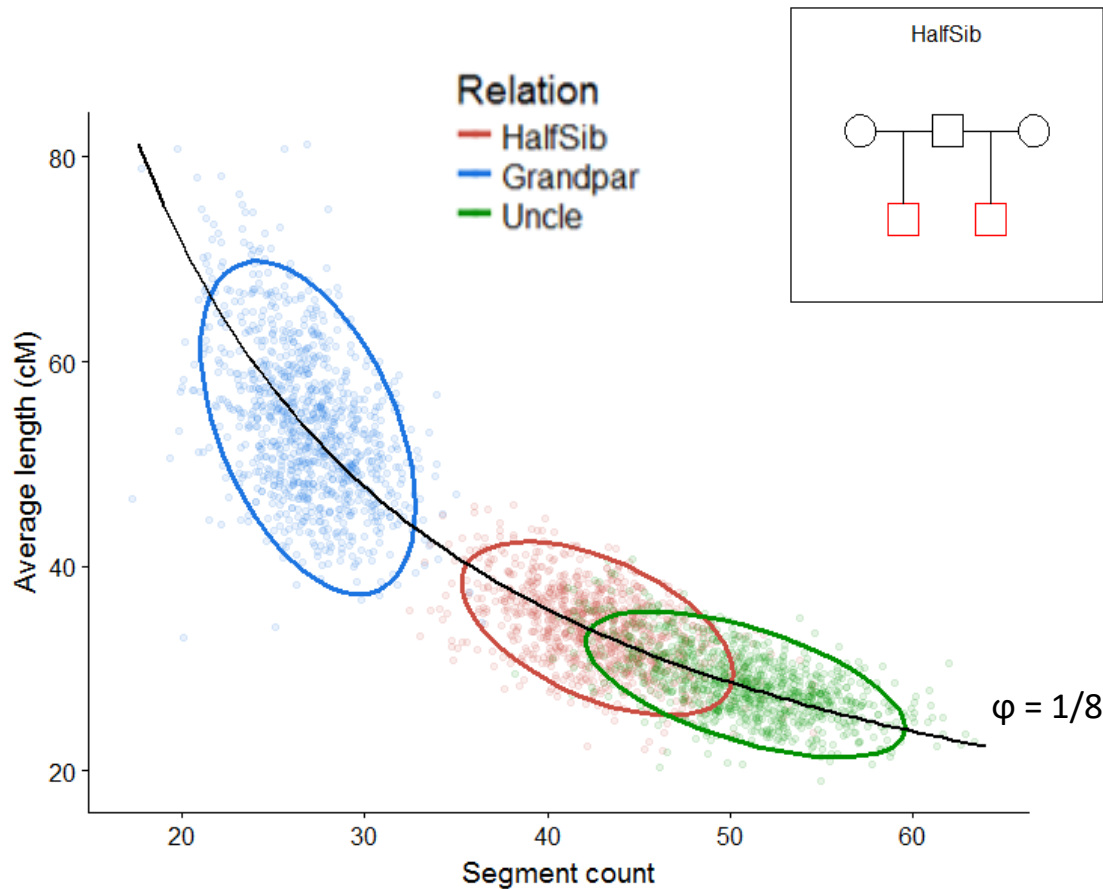
Conclusion
In theory often distinguishable!
In practice quite hard.

The probability of zero IBD

N'th cousins	P(zero IBD)
first	0.0 %
second	0.0 %
third	1.5 %
fourth	28 %
fifth	67 %
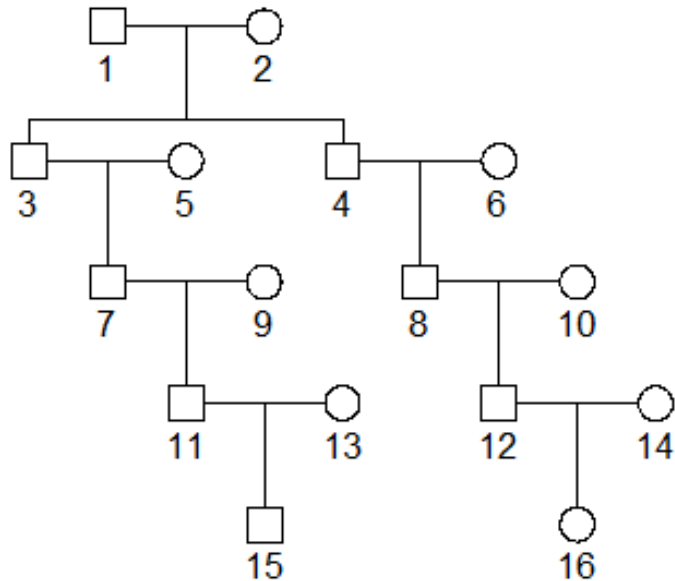
Third cousins: Expected fraction of the genome with IBD = 1:

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

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 $\kappa = (\kappa_0, \kappa_1, \kappa_2)$

- 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*