Statistical methods in genetic relatedness and pedigree analysis

NORBIS course, Oslo, June 2022
Magnus Dehli Vigeland and Thore Egeland

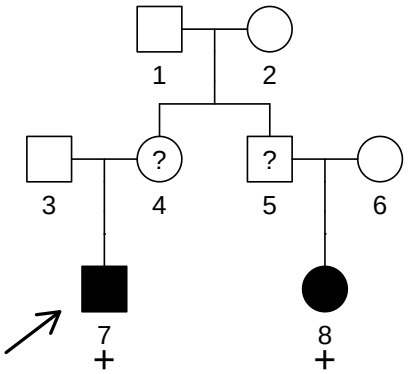
Exercise set VIII. Segregation analysis for variant interpretation

Before starting, load the core **pedsuite** packages and the specialised package **segregatr**.

```
library(pedsuite)
library(segregatr)
```

Exercise VIII-1

Consider the following pedigree, affected with an autosomal dominant disease with reduced penetrance. A potentially causal variant is detected in some members. The disease allele has frequency 0.001.



a) Who are affected? Who are carriers of the variant? What does the arrow mean? What do the question marks mean?

b) Create the pedigree in R with the following code:

```
x = cousinPed(1) |>
  swapSex(c(3,8)) |>
  relabel("asPlot")
```

(Bonus exercise: Use `plotSegregation()` to reproduce the plot shown above.)

c) Initially we assume that the penetrance vector is $(f_0, f_1, f_2) = (0.05, 0.7, 0.7)$. What do the numbers mean? What is the phenocopy rate and penetrance for this disease?

d) Compute the full-likelihood Bayes factor (FLB) by completing and running this command:

```
FLB(x,
    affrected = ,
    carriers = ,
    proband = ,
    unknown = ,
    freq = ,
    penetrances = )
```

What is the strength of the segregation evidence?

e) Suppose that after a clinical examination, individuals 4 and 5 are found to be affected. Change the arguments to `FLB()` accordingly and re-run the analysis. What is the FLB now?

f) Suppose further that individuals 4 and 5 are carriers of the rare allele in question. Compute the FLB with the new information.

g) Adjust the penetrance values to those of a fully penetrant dominant disease with no phenocopies, and compute FLB again. Explain why the answer is undefined (`NaN`) in this case.

h) Set individuals 1 and 2 to have unknown affection status. What is FLB then?

Exercise VIII-2

In this exercise we go back to the original situation of the previous exercise, and examine the effect of including additional affected relatives.

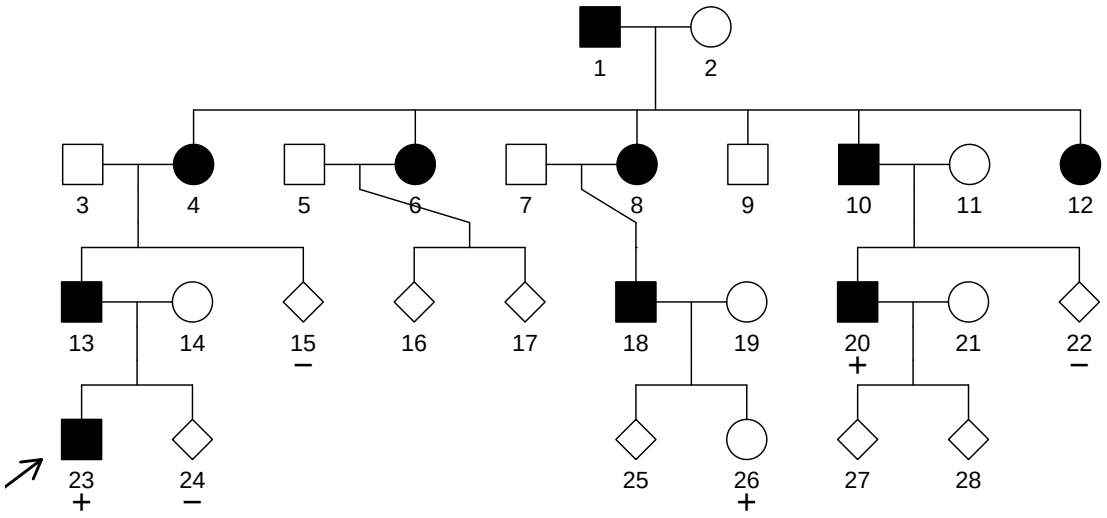
a) It is brought to light that individual 8 has a brother, who is also affected - but not yet genotyped. How do you think this affects FLB. Modify the pedigree and check your intuition.
Hint: Recreate the pedigree as in part a of Exercise VIII-1, and add `x = addSon(x, parents = 5:6)`.

b) The brother is subsequently genotyped. Find the new FLB in both scenarios: (i) if he carries the variant, and (ii) if he does not carry the variant.

c) Individuals 4 and 5 have a half sister, who is also affected, and who also carries the variant. Compute the FLB and classify the evidence according to the thresholds given in class.

Exercise VIII-3

This exercise is based on a real case. The family shown below, slightly modified for anonymity purposes, is affected with a rare, dominant form of hereditary spastic paraplegia (HSP).



A potential genetic cause was identified in the family. The plus and minus signs indicate known carriers/non-carriers.

a) Explain that the variant does not segregate perfectly with the disease.

b) Load the pedigree:

```
y = readPed("data/segregation-hsp.ped")
```

c) Assume that the disease has penetrance of 90% and no phenocopies, and allele frequency 0.00001. Compute the FLB and give a conclusion.