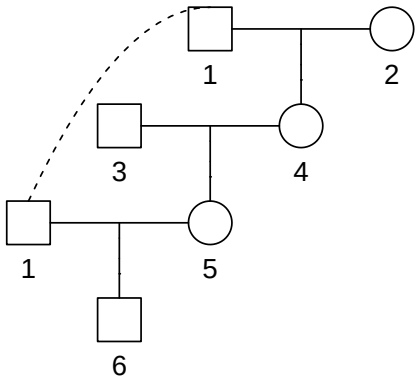
Kinship and pedigree analysis: Methods and applications

Magnus Dehli Vigeland and Thore Egeland

Solutions for exercise set V: Realised relatedness

Exercise V-1

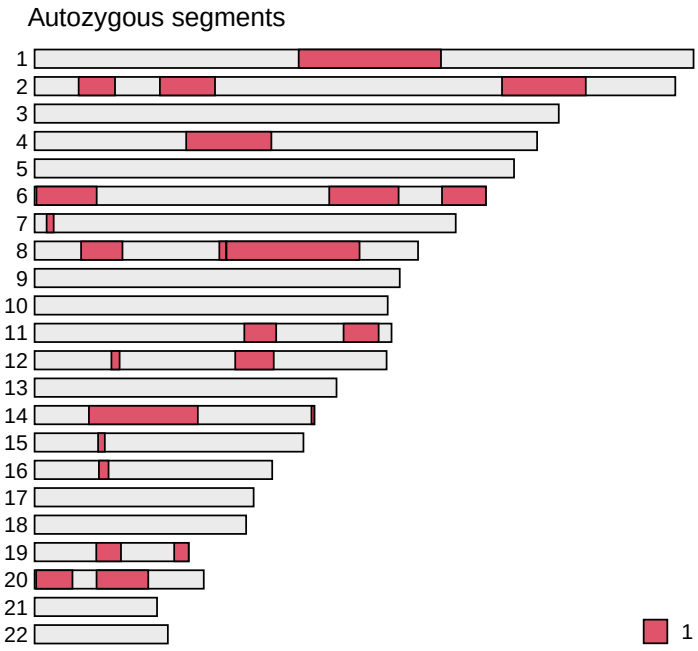
a) Pedigree:



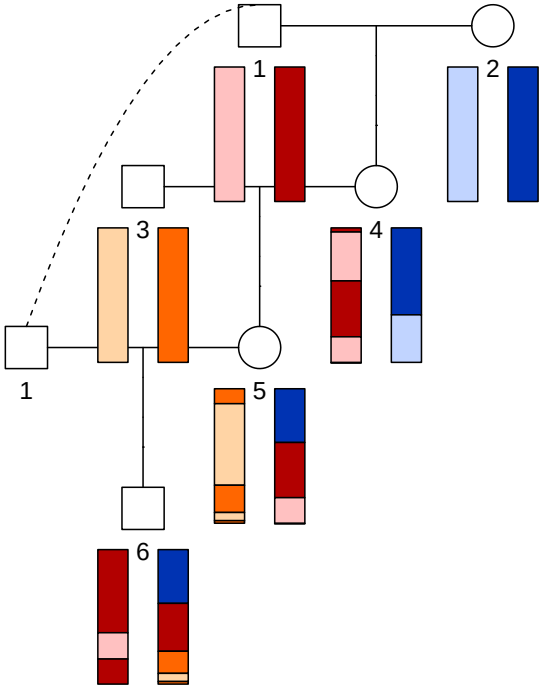
b) $f = \frac{1}{8} = 0.125$, obtained e.g. by the command `inbreeding(x, 6)`.

c) (Answer given in the exercise.)

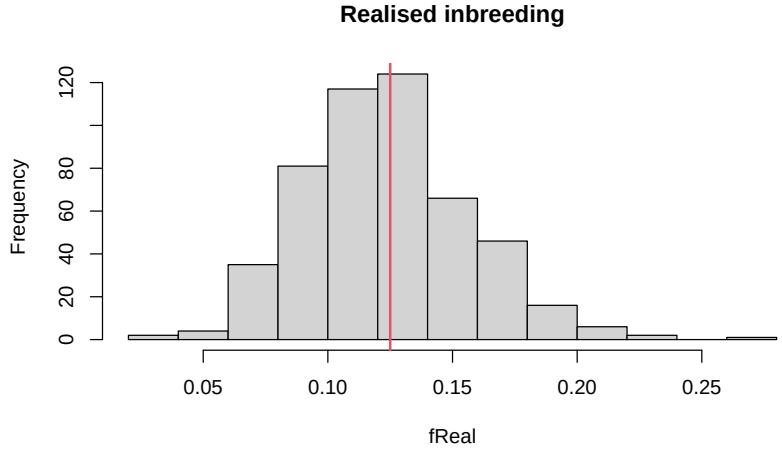
d) The code given in the exercise produces the following karyogram:



e) Haplotype plot:



f) The histogram shows that there is substantial natural variation around the expected value $f = 1/8$.



g) The standard deviation is 3.3%. R command: `sd(fReal)`.

h) The number of segments in each simulation is contained in the `nSeg` column of `r$perSimulation`. A summary of these numbers show that the numbers range from 8 to 32, with an average of 20 segments:

```
summary(r$perSimulation$nSeg)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    8.00   17.00   20.00   19.72   23.00   32.00
```

Exercise V-2

a) (Omitted.)
b) (Omitted.)
c) The distribution of number of segments peaks around 15-20 segments.
d) The second distribution peaks around 20-25 segments. The difference occurs because the second pedigree has more female meioses, which on average recombine more and therefore give more (but shorter) segments.

Exercise V-3

a)
```
x = linearPed(10)
sims = ibdsim(x, N = 500, ids = c(1, 21))
z = zeroIBD(sims)
z
```

```
## $zeroprob
## [1] 0.622
##
## $stErr
## [1] 0.02168483
```

```
1 - z$zeroprob
```

```
## [1] 0.378
```

The simulation estimates a probability of 0.38 that some of of Napoleon's DNA survived down to you. *Comment.* Your answers may differ slightly because of the random number generation. To ensure reproducible results, you should set an explicit random number seed in the simulation command, e.g., `ibdsim(..., seed = 1234)`.

b) To simulate a female line, we add `sex = 2` to the `linearPed()` call. Note also that we use the female founder (2) instead of the male (1) in the `ids` argument.

```
x = linearPed(10, sex = 2)
sims2 = ibdsim(x, N = 500, ids = c(2, 21))
zz = zeroIBD(sims2)
1 - zz$zeroprob
```

```
## [1] 0.514
```

This gives an estimate of 0.51 of some surviving IBD in the female line. The probability is higher in this case because of the higher recombination rate in females.

Exercise V-4 Bonus exercise: Fly forensics

Complete R code:

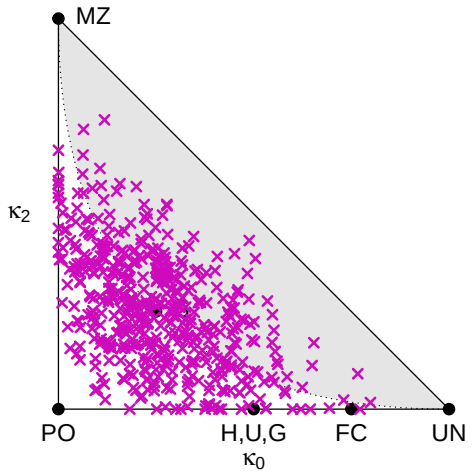
```
# Recombination map
mapFly = list(uniformMap(cM = 107, chrom = 1),
              uniformMap(cM = 110, chrom = 2))

# Pedigree
x = nuclearPed(2)

# Simulation
simsFly = ibdsim(x, N = 500, ids = 3:4, map = mapFly, seed = 123)

# Realised kappa for each sim
kFly = realisedKappa(simsFly)

# Plot
showInTriangle(kFly$perSimulation)
```



The points cover almost the entire triangle! The short fly genome gives much more variation than for humans. Thus, among flies, siblings are much more difficult to recognize as such.