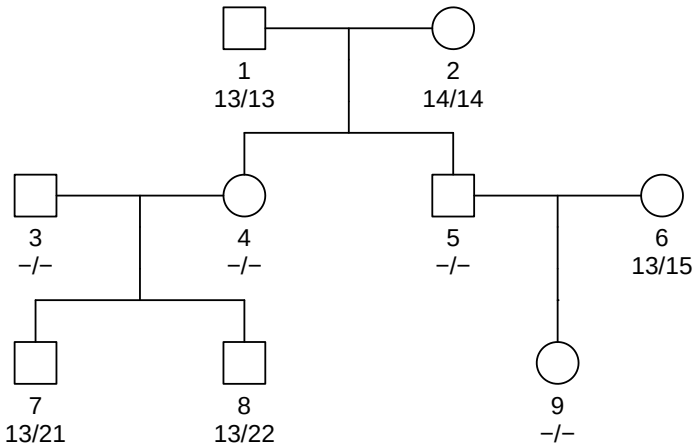
Pedigrees and kinship analysis in forensic genetics

Magnus Dehli Vigeland and Thore Egeland

Exercise set II. Kinship testing in Familias and R

Exercise II-1

Consider the following pedigree, in which some members have been typed with a single DNA marker.



a) What kind of marker is this: SNP or STR? Autosomal or X-linked? How do you know?
b) How many different alleles are observed in the family. What do the allele labels (e.g. `13`) mean?
c) What are the genotypes of individuals 4 and 5?
d) Can you determine the genotype of individual 3?
e) What are the possible genotypes for individual 9, and how likely is each of them?

Exercise II-2 (Kinship analysis in Familias)

This exercise is designed for those who would like to use **Familias**, and prepare input fam-files for **KLINK** (next session).

We consider the following hypotheses regarding the relationship between two males, NN1 and NN2:

- H1: NN1 and NN2 are paternal half sibs.
- H2: NN1 and NN2 are unrelated.

Data for two markers are available. For vWA, the observed alleles are 16 and 17, with frequencies 0.2 and 0.3, respectively. For D12S391, the observed alleles are 20 and 21, with frequencies 0.01 and 0.09, respectively. At vWA, NN1 has genotype 16/17 while NN2 has 17/17. At D12S391 their genotypes are 20/21 and 20/20, respectively.

a) Draw the pedigree corresponding to H1.

b) Calculate the LR in Familias. What are the LRs for the individual markers?

Since manual input is tedious, and not practical for real cases, we next explain how input can be read from files.

c) There are several formats for the marker data. We will use the the format of the file available here: https://familias.name/alcala2026/db.txt. Import the markers from the 'General DNA data' window.

d) There are also several formats for the genotype data. We will use the the format of the file available here: https://familias.name/alcala2026/case.txt. Import the genotypes from the 'Case DNA data' window.

e) Create the pedigree corresponding to hypothesis H1 above in QuickPed and save as a ped-file. Remember to tick of 'Family ID' in the lower right corner of QuickPed. Import the pedigree in the 'Pedigrees' window of Familias.

f) Generate hypothesis H2 in Familias and calculate the LR.

g) Save the Familias project to a file named *halfsib_case.fam*. (This will be used in the next session, where we will open it in KLINK and analyse the case further.)

Exercise II-3 (Kinship analysis in R)

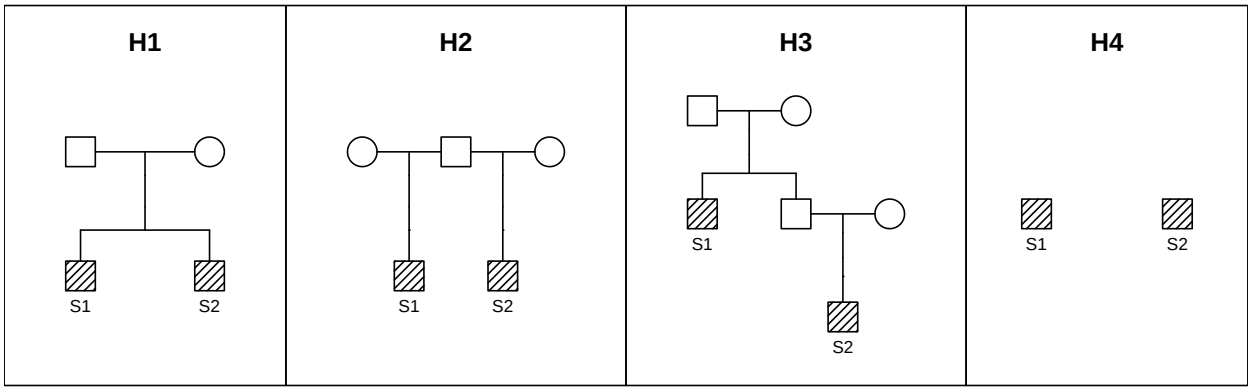
This exercise is intended for those who want to use R and the pedsuite. To get started, open RStudio and load these packages:

```
library(pedsuite)
library(pedFamilias) # not a core package, so must be loaded separately
```

You must also run the following commands to download some example datasets:

```
download.file("https://magnusdv.github.io/pedinr/datasets/data.zip", dest = "data.zip")
unzip("data.zip")
```

The files are put in a subfolder `data` in the current working directory. (To see where you are, run `getwd()`.)



We will analyse the relationship between two males, S1 and S2, who are genotyped with 15 STR markers. The hypothesised pedigrees are shown above.

a) Define the first three pedigrees by running:

```
ids = c("S1", "S2")
H1 = nuclearPed(children = ids)
H2 = relabel(halfSibPed(), old = 4:5, new = ids)
H3 = relabel(avuncularPed(), old = c(3,6), new = ids)
```

b) Load the marker data, attach the 'NorwegianFrequencies' of allele frequencies, and check that the result contains the typed individuals S1 and S2 as singletons. *Hint:*

```
H4 = readPed("data/kinship-riddle.ped")
H4 = setFreqDatabase(H4, database = NorwegianFrequencies)
summary(H4)
```

c) Find the LRs using H4 as the reference.

d) Notice that H2 and H3 give identical LRs. Do you think this is a coincidence? Explain!

e) Include a hypothesis specifying that S1 and S2 are first cousins, and find the LR against H4. *Hint*:

```
H5 = relabel(cousinPed(1), old = 7:8, new = ids)
```

f) *Bonus exercise for Familias users:*
Confirm the LRs in Familias, after saving the data as a fam-file as follows:

```
peds = list(H4 = H4, H1 = H1, H2 = H2, H3 = H3, H5 = H5)
writeFam(peds, famfile = "riddle.fam")
```

In the code above, note that H4 is listed first, since this contains the marker data.

Bonus exercises (if you have time)

Exercise II-4

This exercise demonstrates how you can continue projects started in Windows Familias in R. Once the Familias data, i.e., the `.fam` file, has been converted to pedigree objects in R, we have access to all the functionality of the `pedsuite`.

A child (CH) and the alleged father (AF) are genotyped with 21 STR markers. The mother is not disputed. We consider the following two hypotheses:

- H1: The alleged father is the biological father.
- H2: The alleged father and the child are unrelated.

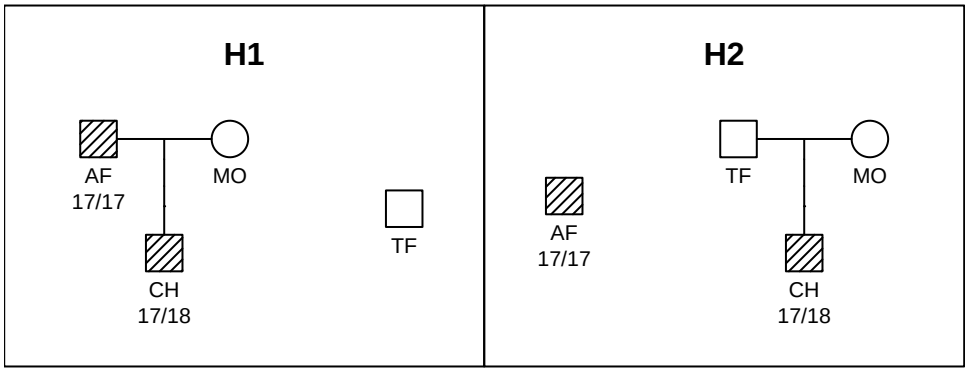
a) Use `readFam()` from the **pedFamilias** package to download the Familias file and convert to R.

```
dat = readFam("http://familias.name/alcala2026/paternityCase.fam")
```

b) The next step is to understand the output of `readFam()`. Run the following commands to explore it:

```
is(dat)
names(dat)
summary(dat)
plotPedList(dat)
```

c) Produce the plot below as accurately as possible, including hatched symbols for the typed members, and genotypes for the first marker.



d) To simplify matters below, extract the two hypotheses as separate objects, `H1` and `H2`:

```
H1 = dat[[1]]
H2 = dat[[2]]
```

e) Show that LR comparing H1 to H2 is 0, by running:

```
res = kinshipLR(H1, H2)
res
```

f) Find the marker with LR = 0.

g) Calculate the LR once more, after first removing the marker giving LR = 0. (Note: This is a practice we advise strongly against!)

Rather than removing incompatible markers, we introduce a mutation model. The possible mutation models include `custom`, `equal`, `proportional`, `stepwise` and `onestep` and are described in the documentation of `pedmut::mutationModel()`. Different models can be used for females and males. Note that `custom` is completely general as you can define the mutation matrix. Below we use the `equal` model with rate 0.001, for both females and males.

h) Apply an equal mutation model to PENTA_E (only) and recalculate LR. *Hint:*

```
H2 = setMutmod(H2, marker = "PENTA_E", model = "equal", rate = 0.001)
lr = kinshipLR(H1, H2, source = 2)
```

What is the total LR now? What is the LR for the marker PENTA_E?

i) What is the total LR when the `equal` model is used for *all* markers?

j) *Bonus exercise for KLINK users:*
Confirm the above LR using KLINK. Comment on the result.