

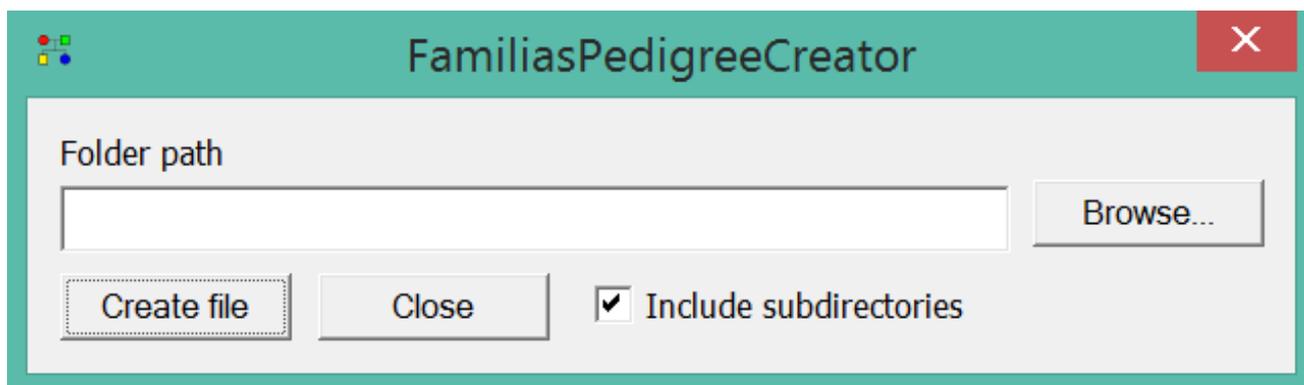
Extra Exercises

Thore Egeland

Input files for Exercises C-2 and C-3 are available from <http://familias.name/cyprus2017/ExtraExercises.zip>.

Exercise C-1. This exercise demonstrates the program `FamiliasPedigreeCreator` made by Daniel Kling. The program generates an R script which plots the pedigrees in `fam`-files in a folder.

- Download <http://familias.name/Ch2Input.zip>, uncompress the files and store them in a folder named `Ch2Input`.
- Open `FamiliasPedigreeCreator` (see figure below), specify `Folder path` `Ch2Input` and tick of `Include subfolders`. Run the program.
- Paste the commands produced by the program into R and look at some of the figures produced; each folder of `Ch2Input` with `fam`-files should have a sub-folder named `Pedigree plots` with plots stored as `png`-files. See <http://familias.name/VideosBook.pdf> for a video-demonstration of this exercise.



Exercise C-2. We will do Exercise 2.3 once more, but now we will read input from files and use the DVI module of `Familias`.

We quote from Exercise 2.3: Hikers come across human skeletal remains in a forest. Evidence around the site provides a clue as to the identity of the individual. You are asked to test a bone to determine if the individual (bone) is related to an alleged mother (AM) and the mother's other daughter, the alleged full sister (AS,) see Figure 1. The hypotheses are

H_1 : The bone belonged to the daughter of AM and sister of AS.

H_2 : The bone belonged to someone unrelated to AM and AS.

- Open `Familias`. Import the marker data from `exerciseC2-markers.txt`.
- Enter `Tools > DVI > Add Unidentified Persons`. Import data from `exerciseC2-bone.txt`. Look at the profile by marking the line with `Bone` and pressing `View`.
- Enter `Add reference families (AM)` for instance by pressing the `Next` button. Read the data for the family members by pressing `Add` and `Import Simple` the file `exerciseC2-family.txt`.
- Next we need to define how `AF`, `AM`, `AS` are related to the `Missing person` (i.e., `Bone`). This is done by pressing `Add` in the right hand side of the window and entering the parent child relationships. Name the pedigree `Missing sister`. Enter `Close` and `Close` followed by `Next` to get into the `Results` module. Enter `Search`. Choose to save the file and press `Update` and use default settings. What is the *LR*? How do you interpret the *LR*?

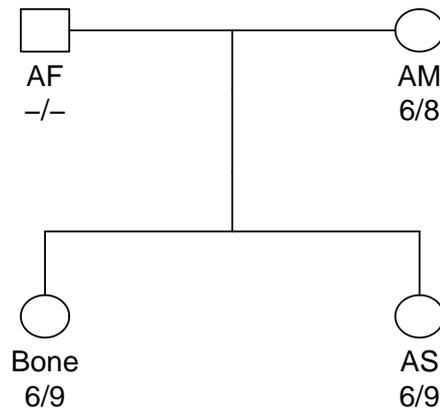


Figure 1: The case of the missing sister. One of eight markers shown.

- e) Do Exercise 2.3 once more and check that you get the same result, the same LR as above.

Exercise C-3. This exercise studies the same case as the previous exercise. The difference is that we imagine that we have no genotype data for the **Bone**, see Figure 2. We will address the question: Will we get a reliable conclusion once we have genotyped the **Bone** for the eight markers available? Or should we type more markers or more individuals?

- Open the file `exerciseC3.fam`. Enter `Add reference families (AM)`
- Open `Prepare pedigree plots`. After entering `OK` you will get a file with `R` commands. Paste these commands into `R`. (If the `R` library `Familias` is not previously installed, it will be installed now if you have internet access.) Note that a plot is produced with genotyped individuals indicated.
- Return to `Familias`. Enter `Evaluate` to get into to `Reference family evaluation tool`. Press `View family` to have a look at the family and the marker data.
- In this window data can be simulated and exclusion probabilities calculated as explained next. Press `Start`. Tick of `Conditional simulations`, choose 100 simulations with seed 123. This implies that the simulations will be *conditional* on the observed marker data. For instance, as mutations are not included now (mutations can be modelled), **Bone** must have allele 6 or 8 for the marker shown in Figure 2. Enter `OK`, `Save` and `No`. Save the file with `R` commands and paste these commands into `R`. Once the simulations are finished, the results can be viewed in `R` or `Familias` (after pressing `OK`). Excerpts of the output is shown Figure 3. The probability of exclusion is 0.998. Interpret! What is the expected LR ? What is the probability that LR will exceed 10000?

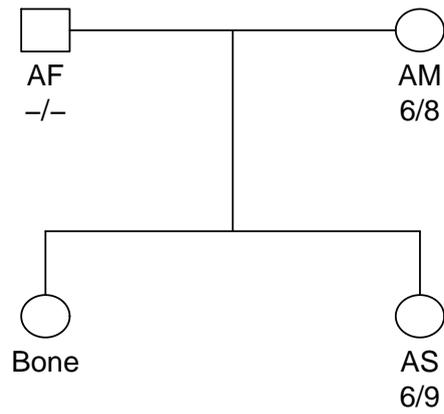


Figure 2: The case of the missing sister. One of eight markers shown. Now, there is no data for Bone.

Refe...	#..	#..	I..	Mean LR	Median LR	P(LR>10000)	Exclusion pr...
Refer...	2	8	0	6.21e+008	2.62e+005	0.770000	0.998437

Figure 3: The case of the missing sister. Simulation output,