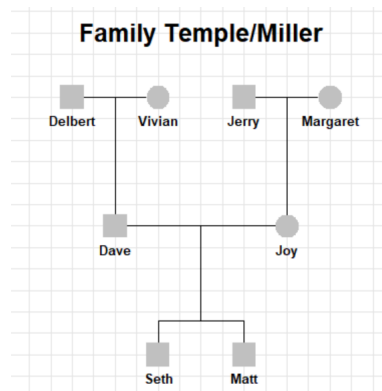


Given a pedigree (family tree), we can compute kinship, which is one half the expected proportion of the genome that relatives share identical by descent. (That is, kinship values range from 0 to $1/2$.) These computations come from the path-counting formula of American population geneticist [Sewall Wright](#). In this project, we will learn about the PED file format to encode family trees in data, and we will write some code to automate these path-counting calculations. The objectives are to (1) gain more practice in programming and to (2) learn through experience the challenges associated with computing statistics on trees and graphs. Below are small exercises to guide the mentee through the project. The mentor will be available to help over email and in-person for 2 meetings.

PED file format

1. Read about the PED file format on this page: [PLINK](#).
2. What the six required columns? Note that columns are separated by tabs.
3. Open the stat550less.ped file in a text editor.
4. Open the temple.ped file in a text editor.
5. Download the HaploPainter software from Source Forge ([link](#)). Unzip (Windows) the folder. Double-click the HaploPainter application in the File Explorer (Windows) to run the program.
6. In HaploPainter application, File → Import Pedigree → Linkage → Select a PED file.
7. Open a text editor. Make your own family PED file. Include siblings, parents, maternal and paternal grandparents, and one set of uncle/aunt + first cousins. Visualize the family tree in the HaploPainter application.



Inbreeding in pedigrees

Here we want to (1) calculate inbreeding and kinship values by hand and (2) determine how these impact the prevalence of recessive conditions. These exercises are based on the following pedigree.

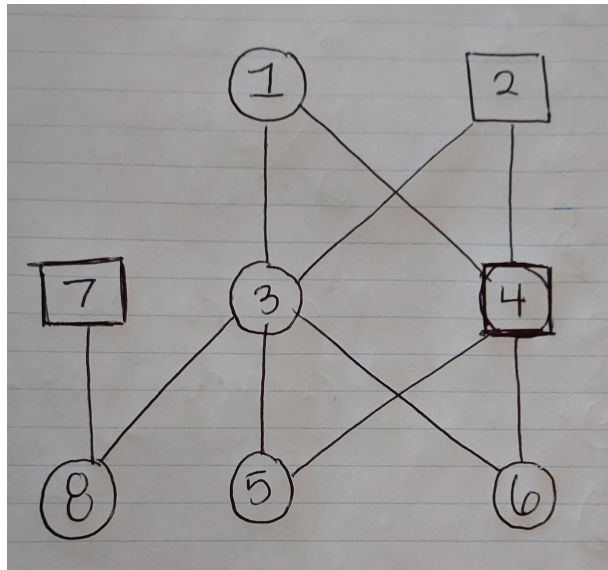


Figure 1: Family tree with inbreeding

1. PED files: as before, use the HaploPainter software to construct a digitized pedigree.
2. Offspring of sib-mating
 - (a) Compute the kinship value between 3 and 4. What is the inbreeding value of 5?
 - (b) Individual 5 has a recessive genotype aa . The recessive allele a is present in the population at 1% frequency. Calculate the probability individual 6 has the homozygous recessive genotype.
 - (c) Individual 5 has a recessive genotype aa . The recessive allele a is present in the population at 1% frequency. Calculate the probability individual 8 has the homozygous recessive genotype.
 - (d) Calculate the probability that two siblings in a family without inbreeding are both homozygous recessive if it is known that one of the siblings is.
 - (e) Explain the law of total probability. Juxtapose the formulas for 2(b-d).
 - (f) (Optional) Verify the probabilities of the condensed ibd states for 5 and 6. See the table here: [Jacquard coefficients](#).