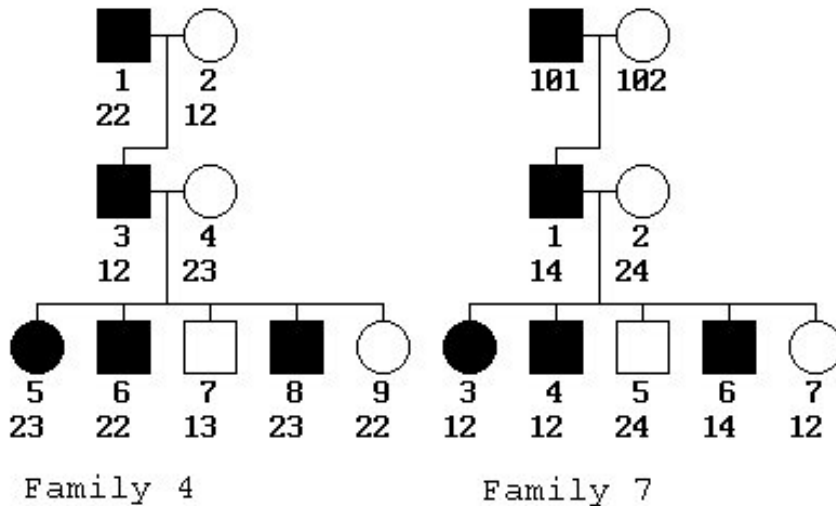


## Homework #5

Due Tuesday Feb. 19 at the beginning of class. Assignments turned in more than 5 minutes after the beginning of class will be penalized 10 points, with an additional 10 points every 24 hours thereafter. You may discuss the homework assignment with other students, but do not share your work. **All Python programs should be run before being turned in. Even experienced programmers can seldom write a program perfectly on the first try.**



- For pedigree 4, how many non-recombinant and recombinant offspring are there?
- For pedigree 7, if grandfather (individual 101) had genotype 13, how many non-recombinant and recombinant offspring are there?
- For pedigree 7, if grandfather (individual 101) had genotype 44, how many non-recombinant and recombinant offspring are there?
- Brief answer: Why do pedigree researchers try to collect grandparents?
- Write a Python program to search a file of DNA sequences for a recognition site which follows this rule:
  - First two bases must be A, T
  - Next base either C or G
  - The next 2 or 3 positions can be any base
  - Next base either C or G
  - Last two bases must be T, A

You may use the `re` module if you wish but are not required to do so.

- Write a Python program to compute the lod score for a specified number of recombinant and non-recombinant offspring and a specified  $\theta$ . Read the recombinant and non-recombinant counts and  $\theta$  value from the command line. Hint:

$$Lod = \log_{10} \frac{(1 - \theta)^{NR} \times \theta^R}{0.5^{(NR+R)}}$$

- Transform your program into a function. Import it into a new program which computes the lod score for all values of  $\theta$  from 0 to 0.5 in increments of 0.05. (There should be 11 values computed.) Print a table of these lod scores, with labels. Also print a line indicating the maximum lod encountered and the corresponding value of  $\theta$ .