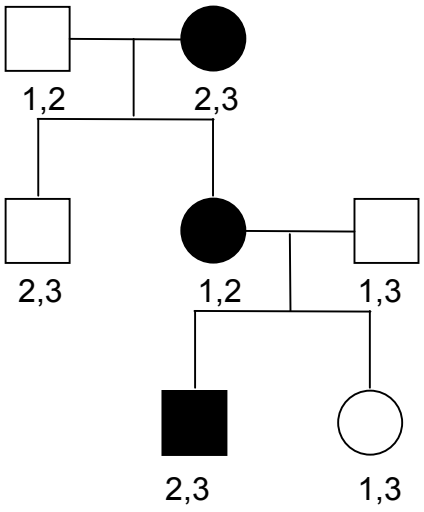
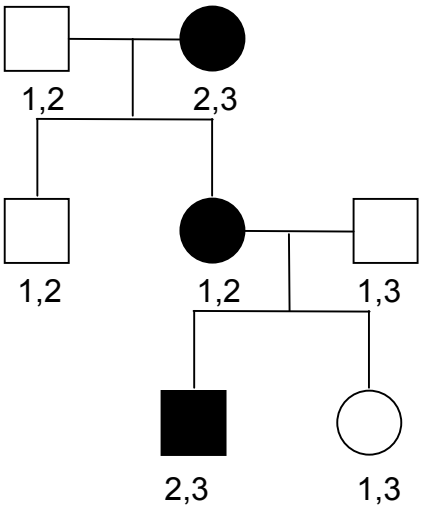
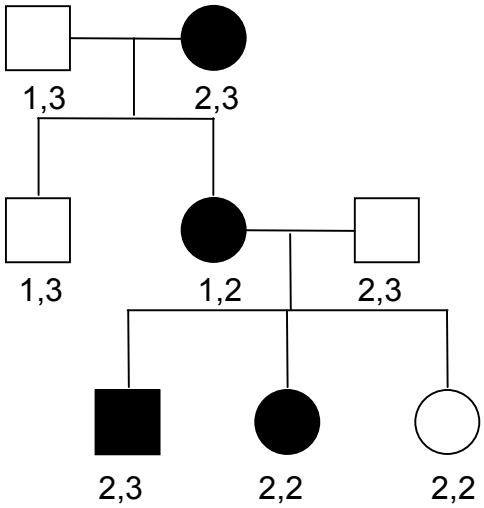
1. Consider these examples of a dominant autosomal disease segregating in families caused by a mutation in a single gene. The numbers below the individuals in the pedigrees represent genotypes at a particular DNA marker. Assume that in each family, individual I-2 is heterozygous at the disease locus and, in this individual, the disease allele is on the same homolog as the marker allele 2.



Family I



Family II



Family III

(a) A "founder" is any individual without a parent in the pedigree. For each family, identify the number of non-recombinants and recombinants among the non-founders only. Use only these non-founders to compute the odds (or LOD=log of the odds, whichever you prefer) that the disease-causing mutation is located 1 map unit from the marker, r = 0.01.

Family I: No recombinants. Odds = 15.4, LOD = 1.19.

Family II: The son in the second generation is a recombinant; the other three children are non-recombinants. Odds = 0.16, LOD = -0.8.

Family III: The daughter in the third generation is a recombinant; the other four children are non-recombinants. Odds = 0.31, LOD = -0.51.

(b) Given the data from all three families taken together, calculate the overall odds ratio for the model in which the disease locus and the marker are linked at 1 map unit distance.

*You are calculating the likelihood of observing the data in Family I AND Family II AND Family III under the hypothesis that the disease locus and marker are linked at r=0.01, relative to the likelihood of observing all these data under the null hypothesis of no linkage. You multiply the odds ratios from the individual families. Thus, the overall odds = 15.4*0.16*0.31 = 0.76, overall LOD = 1.19 - 0.8 - 0.51 = -0.11.*

(c) What does it mean to have an odds ratio less than 1 (a LOD score less than 0)?

The model of linkage between the marker and the disease locus, at the r you consider, is less likely than the model in which they are unlinked.

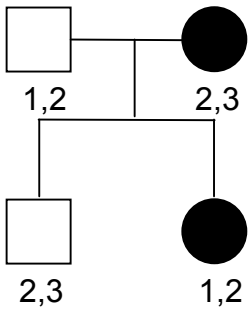
(d) Give an explanation for the difference in odds ratios between the families that you calculated in (a).

The causative mutation controlling the disease may be not the same for each family (the disease is genetically heterogeneous).

(e) Imagine that Family III had a totally different set of genotypes for the marker instead of what's given above, such that in analyzing the pedigree you found zero recombinants. For this new data set, calculate the odds or LOD of linkage between the disease mutation and the marker at r = 0.01 in Family III. These odds are different from the result you got in (a) for Family I, yet in both cases there are no recombinants. Why? Which kind of family would a geneticist prefer to work with?

Odds for this new version of Family III = 30.4 (LOD = 1.48), higher than the odds for Family I, because the observation of no recombinants is less likely to happen by chance in a larger family. For this reason, large families are always preferable.

2. Consider the following pedigree representing a dominant disease caused by a mutation in a single autosomal gene, and genotypes at a single DNA sequence marker. Assume that individual I-2 is heterozygous at the disease locus:



(a) Given a guess for the recombination fraction r between the disease mutation and the marker, why can't you calculate an odds of this pedigree using the formula odds = $[(1-r)^n r^k]/[0.5^{r+k}]$?

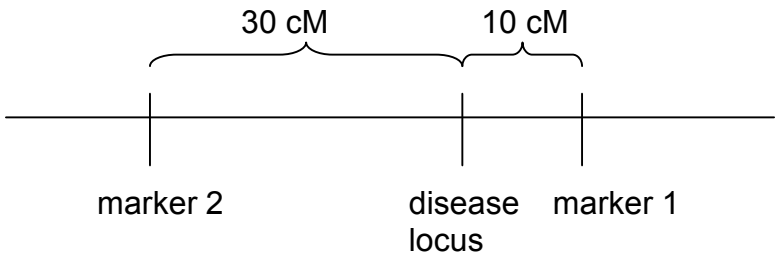
It's clear that the daughter got the disease allele from her mother, but you don't know whether she is a recombinant or not. There are two possible scenarios. Either the mother had the disease allele on her chromosome bearing the 2 allele of the marker, in which case the daughter is not a recombinant, or the mother had the disease allele on her chromosome bearing the 3 allele of the marker, in which case the daughter is a recombinant.

(b) Establish how to resolve this, which is called "the phase problem," and calculate the true odds for r = 0.1.

For each of the two scenarios, compute the odds of observing the pedigree under a model of linkage between the disease mutation and the marker at a genetic distance of r = 0.1. Then factor in that each scenario is equally likely, by multiplying each odds by 0.5 (the probability of each scenario). Finally, add the expressions together, because you are considering the total likelihood of the first OR the second scenario; the total odds of observing the pedigree is the sum of the possibilities.

Odds = $0.5*[(1-r)^0 r^2]/[0.5^2] + 0.5*[(1-r)^2 r^0]/[0.5^2] = 1.64$.

3. Imagine that you are studying a simple Mendelian disease. The true position of the mutation locus responsible for the disease is flanked by two markers, one exactly 10 cM away and the other exactly 30 cM away:



As an experimental geneticist you don't start out knowing the underlying truth about where the locus is; you start out only with data. Imagine your data are as follows: you have a pedigree consisting of 10 non-founder individuals (10 meioses) scored for the disease phenotype and for genotypes at each of the two markers. When tracing the co-inheritance between marker 1 and the disease, you find exactly 1 recombinant individual out of 10. When tracing the co-inheritance between marker 2 and the disease, you find exactly 3 recombinant individuals out of 10. These data give a point estimate for the distance between the markers and the disease locus of 10 cM and 30 cM, respectively. As we discussed in class, these are only estimates inferred from the data. However, in this case each estimate happens to corresponds to the true distance—the absolute, underlying correct answer.

(a) Calculate the odds ratio for r=0.1 for marker 1, and calculate the odds ratio for r=0.3 for marker 2.

For marker 1, odds = 39.67 (LOD = 1.6); for marker 2, odds = 2.27 (LOD = 0.36).

(b) Again, each model you considered in (a), *i.e.* the estimated genetic distance between the markers and the disease locus, represents the correct answer. Given that the model you considered for marker 1 is equally as correct as the model you considered for marker 2, why are the odds ratios different?

Between two unlinked loci the observation of 3/10 recombinations is more likely to happen by chance than is the observation of 1/10 recombinations. This means that the observation of 1/10 recombinants is more convincing: the odds ratio is larger and we can believe with more certainty that the model is correct and the two loci are truly linked under the model we consider.