Biol 213 Genetics (Fall 1999): Addendum
How to refine calculations of map distance

Some of you have had trouble with results obtained from PC-Fly, finding what appear to be internal inconsistencies in the map positions of the markers and your unknowns. Part of the problem is common to all attempts to map by recombination: Large distances are inaccurate. We talked about this problem a few weeks ago, but now you are prepared to appreciate its implications.

Suppose you are using PC-Fly to find the map distance between two markers: *maimed* and *blind*. You've created a dihybrid and are crossing it with a doubly mutant tester strain. You count up recombinants (wild-type and *maimed/blind*) and calculate a recombinant frequency of 0.33. Does that mean that the two genes are 33 map units apart?

Unfortunately not. The likelihood of recombination is proportional to the true distance between genes. If this distance is very small, then recombination is unlikely, and double and triple recombination is very unlikely, so much that we can ignore them. But if this distance is large, then we have to consider the possibility that multiple recombination events have taken place. Fig. 1 shows three different possible outcomes if recombination occurs between the genes *maimed* and *blind*. Note that recombinants are observed when the number of events is odd, and no recombinants are observed when

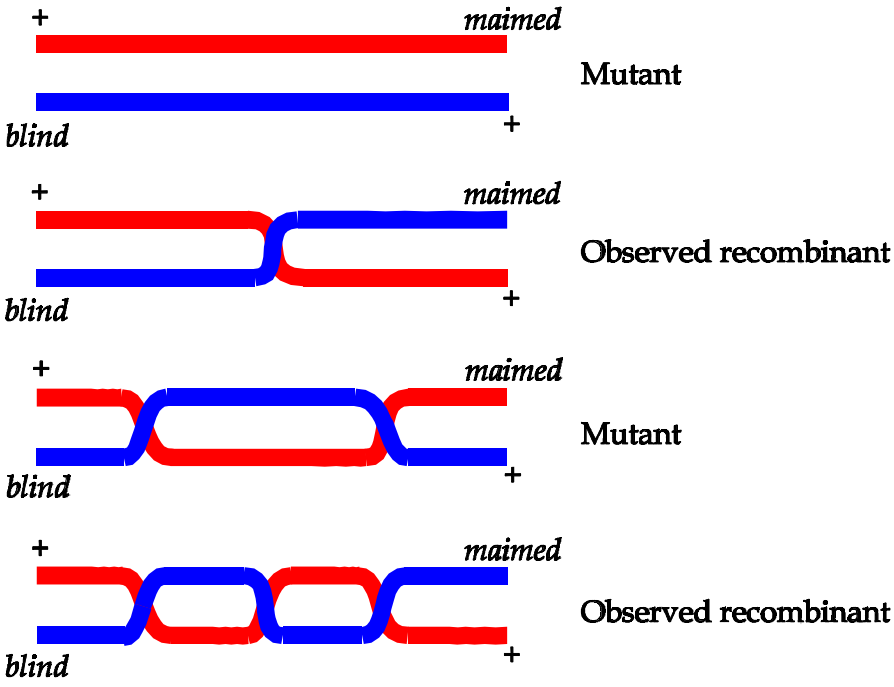

Fig. 1: Phenotypes resulting from odd/even number of Crossovers. Synapse between two homologous chromosomes and the outcome of different numbers of crossover events. Plus indicates wild-type allele.

the number of events is even, even when recombination has occurred. With large distances, you're doing some serious undercounting of recombination events and thus arriving at a distance that is a significant underestimate of the true distance.

We can correct for multiple recombinations by solving the following equation:

Observed Recombination Frequency = Rf = P(1 event) + P(3 events) + P(5 events) + ...

If recombination is random, then each probability is given by the Poisson equation (which turns up all over biology:

$$P(\mathbf{n} \text{ events}) = \mathbf{m}^{\mathbf{n}} \cdot e^{-\mathbf{m}} \; / \; \mathbf{n}!$$

where **m** is the most likely number of events (recombinations). Since map distance (**d**) is defined as the average number of recombinations, then **m** = **d**. The probabilities of each term can thus be calculated, and summing over all odd numbers of events, we get:

$$\text{(Eq 1)} \qquad \mathbf{Rf} = (1 - e^{-2\mathbf{d}}) / 2$$

(the derivation of this relationship is left as an exercise for the reader) (and if you get it, please let me know!).

Let's see if this equation makes any sense. What's the relationship between the recombination frequency and the true distance? What's **Rf** when **d** is very large? (Hint: What is $e^{-2\mathbf{d}}$ if **d** is a huge number) What about when **d** is very small? (Hint: e^{-x} is approximately (1-x) when x is much less than 1). These relationships are shown in Fig. 2 to the right. The red line is the actual relationship between **Rf** and **d**. The dotted line shows the relationship if you take **Rf** to be equal to **d** (the usual assumption in mapping). Clearly, this assumption is good only for small map distances.

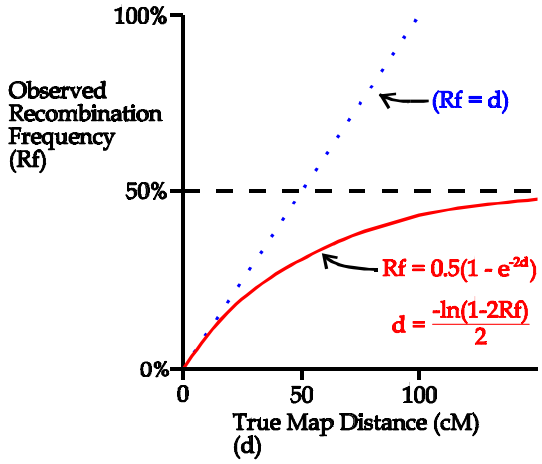

Fig. 2: Relationship between Recombination Frequency and map distance. Solid line represents curve based on Equation 1. Dotted line represents relationship if **Rf** were a true measure of map distance

Usually, you have the recombination frequency, and you want to calculate the true distance, so the equation can be rearranged:

$$\text{(Eq 2)} \qquad \mathbf{d} = -[\ln(1-2\mathbf{Rf})] / 2$$

If we're all together on this, then you should be able to see what happens and why when you try to calculate the true map distance given an recombination frequency of 50%. You can use Equation 2 to make your map distances calculated from recombination frequency more accurate.