## LECTURE 5: LINKAGE AND GENETIC MAPPING

Reading for this and previous lecture: Ch. 5, p 123-141
Problems for this and previous lecture: Ch. 5, solved problems I, II; \#2-5, 7-9, 12, 14, 15, 20, 21, 23, 24 and 27

## Recombination results when crossing over during meiosis separates linked genes

## Reciprocal exchanges between homologous chromosomes are the physical basis of

 recombination. Using chromosomes that had cytologically visible abnormalities, Creighton and McClintock working with maize, and Stern, working with Drosophila, showed that recombination depends upon the physical exchange of equal parts between maternal and paternal chromosomes during meiosis. Both groups followed chromosomes that were physically marked with cytologically visible abnormalities, so that a maternally-derived homolog could be easily distinguished from the paternal one. The marked chromosomes also carried mutations in order to monitor which progeny were the result of recombination.Your book covers Stern's experiments; we'll review Creighton and McClintock's study. They used two different forms of Chromosome 9 in their experiment. One form was normal, and the other had two cytological abnormalities, a heterochromatic knob at one end and a piece of chromosome 8 fused at the other end (a translocation, which you'll hear more about later in the course). In addition to these physical differences, these two chromosomes were genetically marked to detect recombination events. One marker gene controlled kernel color (C, colored; c, colorless) and the other controlled kernal starch metabolism (Wx, starchy; wx, waxy). They set up the following cross ( $*=$ knob, $\sim=$ translocation):

```
* C Wx ~ / c wx x c wx/ c wx
```

All the parental types resulting from this cross retained the parental type chromosomal cytology. The chromosomes of the recombinant progeny ( C wx and c Wx ) were then examined. Creighton and McClintock showed that each recombinant carried only one of the cytological abnormalities ( $* \mathrm{C} w x$ and $\mathrm{c} \mathrm{Wx} \sim$ ), indicating that physical exchange between the cytologically marked and normal chromosome 9 homologs had occurred in the previous generation. Creighton and McClintock's paper ended with the statement "The foregoing evidence points to the fact that cytological crossing-over occurs and is accompanied by the expected types of genetic cross-over.

## Recombination frequencies for pairs of genes reflect the distances between them along a chromosome.

Since genes are arranged linearly on chromosomes, Morgan proposed that different gene pairs exhibited different linkage rates because the closer together two genes are the less likely they will be separated by a recombination event. That is the probability of a crossover occurring between two genes increases with the distance separating them. Sturtevant, an undergraduate in Morgan's lab, suggested that recombination frequency could be used to gauge the physical distance between two genes: $1 \% \mathrm{RF}=1 \mathrm{cM}=1$ map unit. Recombination frequency $=\#$ recombinants/total progeny x 100.

## Experimental recombination frequencies between two genes are never greater than $\mathbf{5 0 \%}$.

 Recombinants among the F2 progeny are never in the majority. Genes on different chromosomes yield $50 \%$ recombination frequency because of independent assortment. Genes that lie far apart on the same chromosome also show $50 \%$. The only way to tell for sure whether the two genes are on the same chromosome is to show definite linkage with other genes that lie in between them. How do we do that? By mapping.Next time, we'll talk about how to set up crosses to determine the positions of genes relative to each other along a chromosome and to gauge the distance between them.

## Mapping: locating genes along a chromosome

## Two-point crosses: comparisons help establish relative gene positions

Let's look at Sturtevant's undergraduate thesis data. Consider three X-linked genes y, m and w. By looking at two-point crosses (crosses tracing two genes at a time), he used the recombinant data to order the genes.

|  | Recomb. Frequency |
| :--- | :---: |
| y-w | 1.1 |
| w-m | 32.8 |
| m-y | 34.3 |

MAP
y --------- w----------------------------------------------------------- m $\quad 32.8$ m.u.
1.1 m.u.

More data (work through the numbers to add $v$ and $r$ to the map and fill in the distances):

| y-v | 33 |  |  |
| :--- | :--- | :--- | ---: |
| $y-r$ | 42.9 (farthest apart) | v-m | 4.0 |
| w-v | 32.1 | v-r | 24.1 |
| w-r | 42.1 | m-r | 17.8 |

MAP


There are limitations to two-point crosses. With crosses involving only two genes at a time, it may be difficult to determine gene order if some of the gene pairs lie close together. See the smaller map above. In addition, the actual distances often do not match up. The distance between $y-r$ in a two point cross is 42.9 m.u., but the distance calculated by adding up all the intervening distances ( $\mathrm{y}-\mathrm{w}+\mathrm{w}-\mathrm{v}+\mathrm{v}-\mathrm{m}+\mathrm{m}-\mathrm{r}$ ) is $55 \mathrm{~m} . \mathrm{u}$.

## Three-point crosses: a faster, more accurate way to map genes

Example:
P: $\quad \operatorname{vg} b \mathrm{pr} / \mathrm{vg} \mathrm{b}$ pr female $\quad \mathrm{x} \quad \mathrm{vg}+\mathrm{b}+\mathrm{pr}+/ \mathrm{vg}+\mathrm{b}+\mathrm{pr}+$ male
F1: $\quad \mathrm{vg} \mathrm{b}$ pr/vg+b+pr+

Test cross: $\quad \mathrm{vg} \mathrm{b} p \mathrm{f} / \mathrm{vg}+\mathrm{b}+\mathrm{pr}+\mathrm{F} 1$ female $\mathrm{x} \quad \mathrm{vg} \mathrm{b} p \mathrm{f} / \mathrm{vg} \mathrm{b}$ pr male

## Test Cross Progeny:

| 1779 | vg b pr | Parental type |
| :---: | :--- | :---: |
| 1654 | $\mathrm{vg}+\mathrm{b}+\mathrm{pr}+$ | $"$ |
|  |  |  |
| 252 | $\mathrm{vg}+\mathrm{b} \mathrm{pr}$ | recombinant for vg relative to b and pr |
| 241 | $\mathrm{vg} \mathrm{b}+\mathrm{pr}+$ | $"$ |
|  |  |  |
| 131 | $\mathrm{vg}+\mathrm{b} \mathrm{pr}+$ | recombinant for b relative to vg and pr |
| 118 | $\mathrm{vg} \mathrm{b}+\mathrm{pr}$ | $"$ |
|  |  |  |
| $\frac{13}{4197}$ | $\mathrm{vg} \mathrm{b} \mathrm{pr}+$ | $\mathrm{vg}+\mathrm{b}+\mathrm{pr}$ |

distance between vg and b is: $(252+241+131+118) / 4197 \times 100=17.7 \mathrm{~m} . \mathrm{u} .(* * *$ this is not correct - see below!)
distance between vg and pr is $(252+241+13+9) / 4197 \times 100=12.3 \mathrm{~m} . \mathrm{u}$.
distance between b and pr is $(131+118+13+9) / 4197 \times 100=6.4 \mathrm{~m} . \mathrm{u}$.

The recombination frequencies show that $v g$ and $b$ are separated by the largest distance and therefore are the outside genes. The two smallest classes represent the double crossover events, in which pr alleles are recombined relative to vg and b . The arrangement of alleles in double recombinants indicates the relative order of the three genes, because in gametes formed by double crossovers, the gene whose alleles have recombined relative to the parental configuration must be the one in the middle.

## Map: <br> vg-------------------pr--------------b

But, again, the distance separating the outside genes vg and $b$ ( 17.7 map units) is smaller than the sum of the two intervening distances $(12.3+6.4=18.7)$. Why? Because the distance between vg and $b$ that we calculated does not account for all the recombination events (double crossovers).
Our three-point cross data lets us adjust the recombination frequency. How? We add the recombination frequency of the double crossovers twice, since each individual in the double crossover groups is a result of two exchanges between vg and $b$. DOUBLE CROSSOVERS are needed to generate gametes in which the middle gene is recombined relative to the two flanking
genes. Thus when we calculate the genetic distance between the two outside markers, vg and $b$, we must add in the double crossovers twice.

## *** CORRECTION TO ABOVE TO ACCOUNT FOR DOUBLE CROSSOVERS:

distance between $v g$ and $b$ is $(252+241+131+118+13+13+9+9) / 4197 \times 100=18.7 \mathrm{~m} . \mathrm{u}$.

## A few additional points about mapping:

- Mapping reveals the relative order of genes, not the actual physical distance.
- The most accurate maps are made by summing the genetic distances of genes lying close together (small intervals). One needs to connect genes that lie far apart through the genes that lie between them.
- INTERFERENCE: A measure of the independence of crossovers from each other. (That is, does a crossover in one region affect the likelihood of a crossover in an adjacent region?)

Calculating interference: First of all, what is the probability of double crossovers occuring? Consider our example of vg , pr , and b linkage. We can calculate the probability of a double crossover using the Law of the Product rule. As long as a crossover in one region does not affect the probability of a crossover in another region, the probability of a double crossover is simply the product of their separate probabilities.

Probability of single crossover between vg and $\mathrm{pr}=0.123$ (corresponds to 12.3 map units) Probability of single crossover between pr and $\mathrm{b}=0.064$ (corresponds to 6.4 map units) Probability of both is $0.123 \times 0.063=0.0079$ or $0.79 \%$
But the number actually observed was different: $([13+9] / 4197) \times 100=0.52 \%$, meaning that the two crossovers are not occurring independently, but instead a crossover in one region reduces the likelihood of a crossover in an adjacent region.

Coefficient of coincidence $=$ frequency observed $/$ frequency expected $=0.52 / 0.79=0.66$ (thus, in our example, only $66 \%$ of the expected DCOs took place).
Interference $=1-$ coefficient of coincidence $=1-0.66=0.34$.
If interference is zero, this means that the frequency of double crossovers is as expected and that a crossover in one region occurs independently of a crossover in an adjacent region. If interference is 1 , this means that interference is complete and that no double crossovers are observed because a crossover in one region completely blocks a crossover in an adjacent region.

