Name:
Recitation Section:

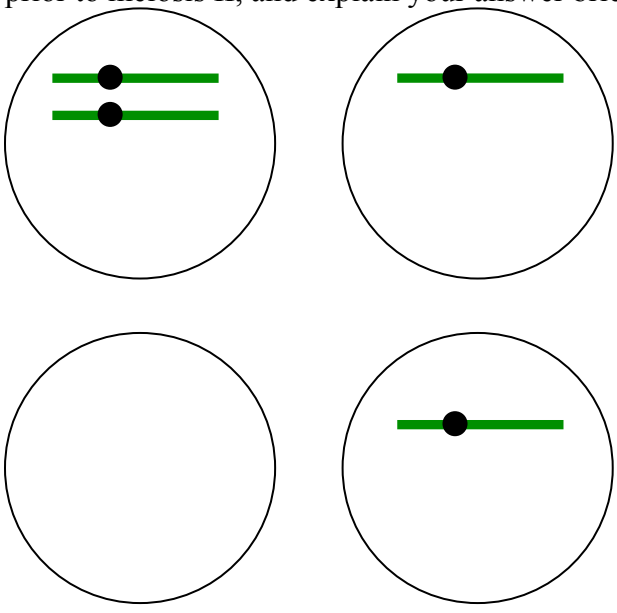
7.03 Problem Set 2 Answer Key
Due Monday, March 2, 2015 by 3 PM

1. You have identified some yeast mutant strains that, when performing meiosis to transition from a diploid state to a haploid state (**sporulation**), end up with daughter spores with abnormal chromosome number (also known as **ploidy**). [5 PTS TOTAL]

a) Through whole-genome sequencing, you discover that one of your mutant strains has a mutation in the gene *SMC1*, which encodes a subunit of the cohesin complex. Briefly explain how this mutant strain might result in abnormal ploidy in daughter cells.
[1 PT]

Cohesin keeps sister chromatids attached during meiosis, which ensures that when they are aligned on the metaphase plate, the spindle fibers from the opposite poles of the cell are attached to one of each. When the homologs or sister chromatids are separated, the spindle fibers must be attached to one of each or daughter cells will get uneven numbers of each chromosome.

b) As it turns out, you are able to copy the mutation of *SMC1* in your yeast strains in mice, and observe the germ cells that result. Let also assume that we are able to selectively induce this mutation during specific stages of meiosis. Considering the case where **only one** of the homologous chromosomes experiences non-disjunction, draw the resulting daughter cells after this mutation is activated just prior to meiosis II, and explain your answer briefly:



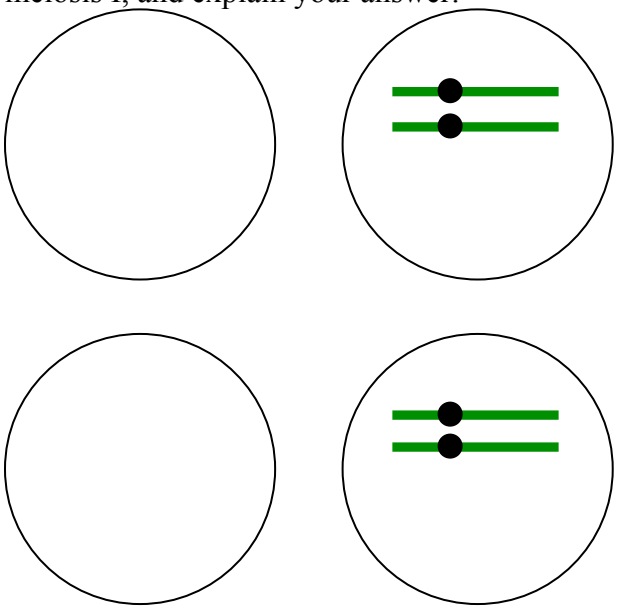
If non-disjunction occurs in meiosis II for one set of homologous chromosomes, then one daughter cell will get two sister chromatids, and one will get zero. The

Name:
Recitation Section:

other two cells will be unaffected (they will get one sister chromatid each).

[1 PT] – Drawing
[1 PT] – Explanation

c) Now draw the resulting daughter cells after this mutation is activated **only during** meiosis I, and explain your answer:



During which stage of meiosis does the product of this gene act? Explain briefly.
If non-disjunction occurs in meiosis I, the two homologous chromosomes will end up in only one of the daughter cells after this stage. Then, in meiosis II, the "daughters" of this daughter will end up with two sister chromatids (double the normal number), while the others will end up with none.

[1 PT] – Drawing
[1 PT] – Explanation

Name:
Recitation Section:

2. You have isolated three novel Drosophila mutants in the lab and want to map them to the Drosophila genome. **cr** makes flies fly in circles, **lg** causes them to grow an extra set of legs and **ye** mutants have yellow eyes. After some preliminary work you determine that these mutant traits are each caused by a single gene and are recessive. You also suspect they might be linked to each other, and to a known recessive fly marker – **eb** (ebony body). [5 PTS TOTAL]

a) Design a set of crosses that will allow you to score offspring with both parental and recombinant genotypes for all four genes. Assume you have access to any true-breeding fly you desire. Explain briefly why you chose these crosses.

$\frac{cr\ lg\ y\ eb}{cr\ lg\ y\ eb}$ x $\frac{+\ +\ +\ +}{+\ +\ +\ +}$ [0.25 PT]

F1: $\frac{cr\ lg\ y\ eb}{+\ +\ +\ +}$ female x $\frac{cr\ lg\ y\ eb}{cr\ lg\ y\ eb}$ male [0.25 PT]

Recessive phenotypes can only be seen on a recessive background, and recombination only occurs in female flies. [0.5 PT]

b) You carry out your crosses and find that you obtain offspring with the following phenotypes (+ denotes wild type). Calculate the genetic distances between the four genes of interest and draw a map of their positions relative to each other. Please show your work.

	Phenotype	**Number of Offspring**
1	ye lg cr eb	70
2	+ + + eb	10
3	ye lg cr +	13
4	ye + cr eb	80
5	+ lg + +	75
6	+ + cr eb	2
7	ye lg + +	3
8	ye + + +	2
9	ye lg + eb	22
10	+ lg + eb	13
11	ye + cr +	11
12	+ + + +	74
13	+ lg cr +	20
14	+ + cr +	24
15	ye + + eb	17
16	+ lg cr eb	4
	Total	440

Name:
Recitation Section:

	Reciprocal Pairs	Independently assorted	Number of Offspring
Parental			
1	ye cr eb	lg	70
12	+ + +	+	74
4	ye cr eb	+	80
5	+ + +	lg	75
Single Cross Over			
9	ye + eb	lg	22
14	+ cr +	+	24
15	ye + eb	+	17
13	+ cr +	lg	20
Single Cross Over			
3	ye cr +	lg	13
2	+ + eb	+	10
11	ye cr +	+	11
10	+ + eb	lg	13
Double Cross Over			
7	ye + +	lg	3
6	+ cr eb	+	2
8	ye + +	+	2
16	+ cr eb	lg	4
	Total		440

Distances can be calculated by looking at the cross, two factors at a time. In this way we can calculate the distance for each pair of genes.

Distance eb to ye:
[(10+13+2+3+4+2+13+11)/440]*100 = 13.2 ± 1.73 cM

Distance eb to cr:
[(10+13+17+20+22+24+13+11)/440]*100 = 29.5 ± 2.59 cM

Distance eb to lg:
[(10+13+17+20+80+75+2+3)/440]*100 = 50 ± 3.36 cM

Distance cr to ye:
[(17+20+2+3+4+2+22+24)/440]*100 = 21.4 ± 2.2 cM

Since the lg is 50cM away from eb, we consider it to be unlinked. If we calculated out the distance of lg from ye or cr, we would also get values around 50cM.

Name:
Recitation Section:

If you look carefully at all reciprocal pairs of genotypes that have similar numbers, you will find that recombinants = parentals for lg relative to any other gene. This is a hallmark of an unlinked gene.

eb 13.2 ye 21.4 cr

_____

lg

____

[1PT] – Distances
[0.5 PT] – Map with eb-ye-cr
[0.5 PT] – Map with lg unlinked

c) Explain why some distances between genes do not add up and any other ambiguities you may come across.

The eb-cr distance is less than the sum of the eb -ye and ye-cr distances because a two factor cross cannot take account of double crossover events. Particularly, classes 6,7,8,16 do in fact result from double-crossover events, but in a two-factor cross, they appear to simply be parentals. They are therefore not included in the numerator of our calculation for distance between eb-cr, and the numerator is thus smaller, and the final distance is smaller.
Ambiguity – lg is either unlinked, or so far away that it appears unlinked.

[1 PT] – No points deducted if ambiguity not mentioned

d) A few months later, you discover another dominant mutant phenotype, **Pi,** that causes your flies to have pink eyes. After attempts to place this gene onto the map you created in (b) you find that out of 500 offspring you only get one offspring that is recombinant for ye and Pi.
What is the genetic distance between these two mutations? If the **ye** gene is 0.09 Mbps and given a recombination rate of 3.3 cM/Mbp in Drosophila, do you think **Pi** & **ye** are alleles of the same gene?

Distance between Pi & ye:
1/500*100 = 0.2 ± 0.2

ye gene length in cM = 3.3cM/Mbp *0.09 Mbp = 0.297 cM

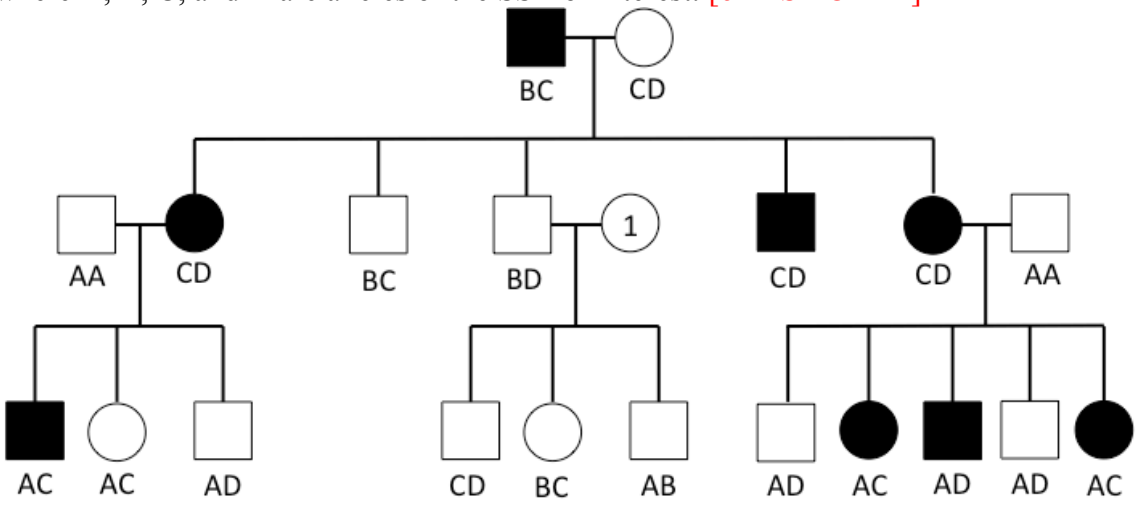
It is possible that they are in the same gene because 0.2 < 0.297 cM.

[0.25 PT] - Pi-ye distance

Name:
Recitation Section:

[0.25 PT] – ye length
[0.5 PT] – conclusion

3. You are a geneticist who has been asked to help pinpoint the inheritance of a newly discovered disease (maybe they will name it after you, if you can figure out its inheritance). Your work has led you to believe that this disease is linked to alleles of a simple sequence repeat (SSR). Based on this, your collaborators have painstakingly collected samples from families and sequenced them, resulting in the following pedigree, where A, B, C, and D are alleles of the SSR of interest: [6 PTS TOTAL]



a) What is the most likely mode of inheritance of this disease? (autosomal or sex-linked? dominant or recessive?) Briefly explain your answer.

Autosomal dominant, both males and females are affected, and about 50% of offspring are affected.

[1 PT]

b) What can you infer about the genotype of the individual marked 1?

1's genotype must be AC.
The right son has an A, which he cannot have gotten from his father, so he got it from his mother. Knowing this, the mother's other allele must be C, because she must pass it on to her other children.

[1 PT]

Name:
Recitation Section:

c) Is it likely that this SSR is completely linked to the disease? Explain briefly. What would the LOD score be for this scenario?
The simplest assumption is that C is the SSR allele linked to the disease, because all of the affected individuals in the second generation received a C from their affected father. (In addition, three of the affected four in the second generation also have that C.) In this case, there are two recombinants: the only daughter of the 2nd-generation parents on the left and the second son (third child) of the 2nd-generation parents on the right. Thus the SSR and the disease are separable and cannot be linked. The LOD score is then equal to negative infinity/undefined.

[1 PT]

d) Estimate a recombination frequency between the trait and this SSR. Calculate the LOD score based on only the individuals in the second generation for this estimated frequency.
In total, there are 13 informative meioses in this pedigree (5 in the second generation, 8 in the third). The middle family's offspring are all uninformative because they all inherited a wild-type allele from both parents. We can observe twp recombinants out of these 13 (the only daughter of the 2nd-generation parents on the left and the second son (third child) of the 2nd-generation parents on the right), and thus we can estimate the recombination frequency to be 2/13 = 0.15.

The second generation's 5 informative meioses are all unphased, and thus we can not determine which allele was linked with the disease in the father. So, we have to give each possibility (B or C) equal weight. Remember that LOD = $\log_{10}$(P(this genotype | linkage)/P(this genotype | no linkage)).

$$LOD = log_{10}[\frac{\left(\frac{1}{2}\right)(0.15^5 + 0.85^5)}{0.5^5}] = 0.85$$

[0.5 PT] – Correctly estimate θ
[0.75 PT] – Calculate LOD score with phase penalty

Note: if we're solving this the way Prof. Hemann did in class on 2/25, then for the family members in the second generation:

P(genotypes | linkage @ θ=0.15) = ½(P(genotypes | B is linked with disease) + P(genotypes | C is linked with disease))

P(genotype | linkage @ θ=0.15) = ½((1/2*0.15*1/2)^5 + (1/2*0.85*1/2)^5)
(the extra ½ is because the unaffected parent is a heterozygote for the SSR, unlike the unphased example from lecture 10)

Name:
Recitation Section:

P(genotype | no linkage) = (½ * ¼)^5

e) Calculate the LOD score for the individuals in the third generation with the same estimate from part (d).

For the third generation, the 8 informative meioses are all phased, in that we can clearly see that their grandfather's C allele is most likely linked to the disease. There are two recombinants, as identified earlier, so the LOD score is calculated thus:

$$LOD = log_{10}[\frac{(0.15^2 * 0.85^6)}{0.5^8}] = 0.34$$

[0.75 PT] – LOD score

Solving this the other way:
P(genotypes | linkage @ θ=0.15) = P(genotypes | C is linked with disease)
P(genotypes | linkage @ θ=0.15) = P(recombinants)*P(non-recombinants)
P(genotypes | linkage @ θ=0.15) = (1/2*0.15)^2*(1/2*0.85)^6

P(genotypes | unlinked) = (1/2*1/2)^8

f) Calculate the LOD score for this entire pedigree. Based on this calculation, can you infer that this SSR is linked to the disease from this pedigree alone?
Remember that LOD scores, because of the mathematics of logarithms, can be added together. So, the total LOD score for this pedigree is 0.85+0.34=1.19. We need a LOD score greater than 3 to infer linkage, so we can not infer that this SSR is linked to the disease. Guess you won't be getting a disease named after you after all.

[0.5 PT] – Total LOD score
[0.5 PT] – Conclusion

Name:
Recitation Section:

4. You have isolated four different yeast mutants by mutagenesis that can not grow on medium lacking leucine. You cross each of these haploid mutants to wild type yeast and observe the following diploids. [4 PTS TOTAL]

Haploid genotype	**Diploid genotype**	**Diploid phenotype**
leu1	leu1/LEU1 or leu1/+	WT
leu2	leu2/LEU2	leu-
leu3	leu3/LEU3	WT
leu4	leu4/LEU4	WT

[0.5 PT] – Fill in table

a) What can you deduce about your four mutants based on these crosses?

leu1, leu3 and leu4 cause recessive phenotypes, while leu2 causes a dominant phenotype.

[0.5 PT] – Mutant conclusion

You then obtain *MAT*a and *MAT*α strains of each of these mutants and cross them to each other to obtain diploids. You sporulate the resulting diploids and obtain 120 tetrads each. You observe the following phenotypes:

Diploid	**Tetrads**	**Tetrad #**	**Tetrad Type**
leu1/leu2	4 leu-	120	PD
leu1/leu3	4 leu- 2 leu- , 2 leu+ 3 leu- , 1 leu+	18 22 80	PD NPD TT
leu1/leu4	3 leu- , 1 leu+ 4leu-	14 106	TT PD

Name:
Recitation Section:

b) Identify the tetrad types observed in the table above. What can you deduce about the relationship between these four mutant alleles based on the tetrads observed?

leu1 and leu2 are either alleles of the same gene or are very tightly linked.

Leu1 (leu2) and leu3 are unlinked as you observe a 1:4:1 ratio of PD:TT:NPD.

Leu1 and leu4 are linked tightly enough that no double crossovers can occur.

[0.5 PT] – Fill in the table
[1 PT] – Three explanations for linkage

c) Calculate the genetic distance between these genes and draw a map.

leu1 to leu3 = 100* ((T+6NPD)/2*tetrads) = 100*((80 + 6*22)/2*120) = 88.3 cM
This is very large and can be considered unlinked (For yeast tetrads a distance of 80cM or greater can be considered unlinked).

leu1 to leu4 = 100* (T/2*tetrads) = 100* (14/240) = 5.83 cM

leu1/2 5.83cM leu4

leu3

[0.5 PT] - leu1 to leu 3 distance
[0.5 PT] – leu1 to leu4 distance
[0.5 PT] – Map with leu3 unlinked