Problem Sets

Fall 1994

7.03 Problem Set 1

due in class Friday, September 16

All three problems will be graded. Some parts of these problems are trickier than they might first appear.

1. Yeast cells have in their membranes a number of different proteins that transport amino acids into the cell. (Incidentally, it is these transporters that allow cells with auxotrophic mutations in amino acid biosynthetic pathways to take up the needed amino acid and therefore to grow on rich medium). The transport system that is responsible for uptake of the amino acid arginine will also take up the toxic compound canavanine. It is easy to isolate mutants that are defective in the arginine transporter by selecting for mutants that are resistant to canavanine. Nine canavanine resistant mutants (can^r) are isolated. Four can^r mutants are isolated in a haploid yeast strain of mating type α (strains 1 - 4) and five can^r mutants are isolated in a haploid strain of mating type **a** (strains 5 - 9). All possible pairwise crosses as well as crosses to wild type are performed as indicated in the table below. When the resulting diploid can not grow on medium with canavanine a (–) is indicated at the intersection of the two parental strains. When the resulting diploid can grow with canavanine a (+) is indicated.

		strains of mating type α				
		1	2	3	4	wild type
strains of mating type **a**	5	–	–	–	–	–
	6	–	–	–	+	–
	7	+	+	+	+	+
	8	+	–	+	–	–
	9	–	–	–	–	–
	wild type	–	–	–	–	–

(a) Which mutants are recessive and which are dominant for canavanine resistance?

(b) How many different genes or, more precisely, complementation groups appear to be required for uptake of canavanine. Indicate which mutations are in the same complementation group. Also indicate any ambiguities in the assignment of complementation groups.

(c) Propose a simple experiment to resolve any ambiguities in the assignment of strains 5 and 9 to complementation groups.

2. **(a)** A new mouse mutant is discovered that has a tail shorter than normal. When a short tailed mouse is crossed to wild type about half of the F1 are normal and half have short tails. Is the the short tailed mutation dominant or recessive? When two short tailed F1 mice are crossed what fraction of the F2 would you expect to have short tails?

(b) After college you get a job as a genetics expert in a biotechnology company. Your first job is to make a true-breeding line with the short tailed mutation. A year and many crosses later you have still failed to produce a line that only gives short tailed progeny. Afraid that you are going to lose your job you recall something you learned many years ago doing the first 7.03 problem set. What hypothesis can you come up with to explain your failure to isolate true-breeding short tailed mice, that is, mice that are homozygous for the short tailed mutation (assume complete penetrance of the short tailed mutation). According to this hypothesis, when two short tailed mice are crossed what fraction of the progeny would be expected to have short tails?

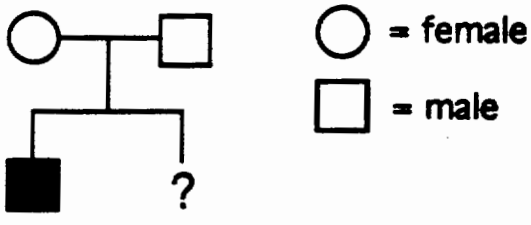
(c) For a cross between two short tailed mice approximately how many progeny would you expect to have to score in a typical experiment, to distinguish the hypothesis in part **b** from the hypothesis in part **a**? (For this problem an exact analytical solution would be very hard to obtain so just assume an average outcome given the expectation for part **b** and then use the Chi-square table to find the approximate number of progeny that you would need to score to show a significant deviation from the outcome expected from part **a** at the $p < 0.05$ significance level).

(d) Another geneticist in your company has independently isolated another short tailed mouse. After performing some crosses she finds that this mouse has similar properties to the mutant you are working with. That is, the new short tailed strain

produces a mixture of normal and short tailed mice when crossed to wild type and attempts to generate a true breeding short tailed strain have failed. Propose an experiment and interpretation of the data that you would use to determine whether the two mutations are alleles of the same gene.

(e) If you think about it, the test you proposed for part **d** is actually a complementation test. Given that complementation tests are only meaningful when performed with recessive mutations, explain why the test you are proposing is valid.

3. Consider the following simple pedigree where a couple have a son with a very rare trait. The couple is going to have another child (sex not yet known) and they ask your advice about the chances that their next child will be affected with the trait.



(a) Assume that the trait is autosomal recessive with complete penetrance. What is the probability that the next child will have the trait? If the next child doesn't exhibit the trait, what is the probability that they are a carrier?

(b) Assume that the trait is X-linked recessive with complete penetrance. What is the probability that the next child will have the trait? If the next child is a girl what is the probability that she will be a carrier?

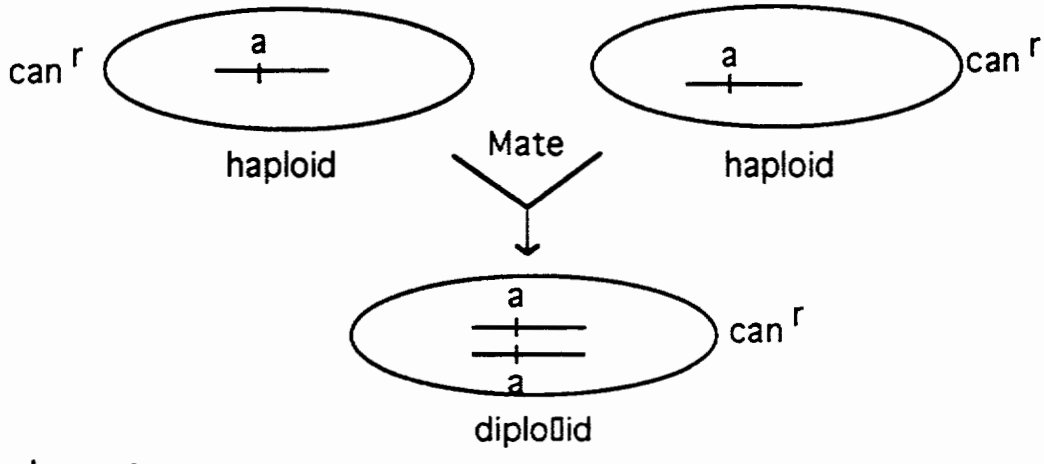
(c) Assume that the trait is autosomal dominant with incomplete penetrance. If the probability of expression of the trait in a heterozygote is 20%, what is the probability that the next child will exhibit the trait? If the next child does not exhibit the trait, what is the probability that they carry the trait?

7.03 Problem Set 1 Answers

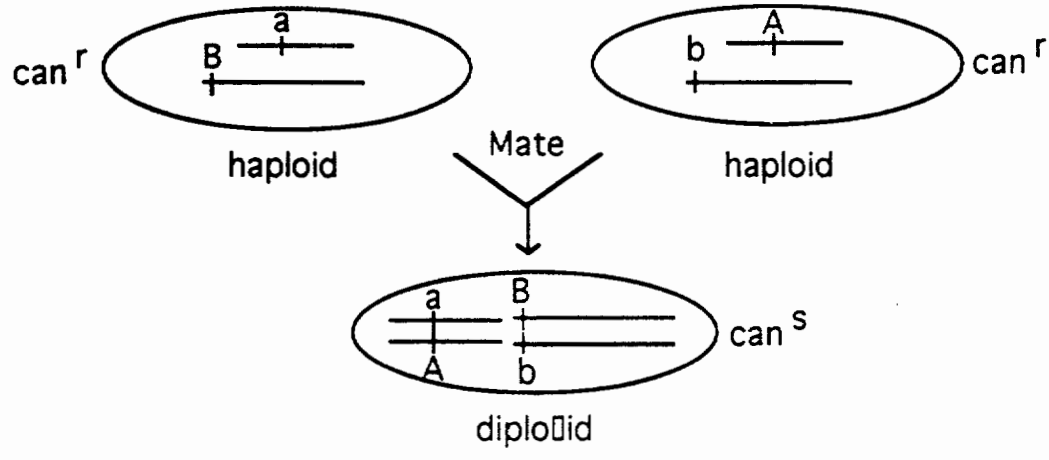
1. (a) All mutants are recessive except mutant 7. Mutant 7 is dominant to wild type because resulting diploids are canavanine resistant (can^r) when mated to a wild type strain. All others are recessive with respect to wild type because the resulting spores are canavanine sensitive (can^s) when mated to a wild type strain. The crucial information is the result of the cross to wild type. Remember that the terms dominant and recessive refer to phenotypes of alleles, and that they are relative terms (i. e. A mutant phenotype may be recessive to wild type, but dominant to different mutant phenotype, as long as the phenotypes can be distinguished).

(b) Illustrated below is a schematic representation of the two possible results of a complementation test:

Noncomplementors:



Complementors:



The lower case letters represent recessive canavanine resistant mutations while the upper case letters represent the wild type canavanine sensitive alleles of the genes. In the case of the noncomplementors, the resulting diploid will contain two recessive canavanine resistant alleles in the same gene (genotype a/a). Therefore, canavanine will not be imported into the cells, and the cells will live. In the case of the complementors, the diploid will contain a functional wild type allele of each of two genes (genotype A/a B/b). In this case canavanine will still be imported into the cells, and the cells will die.

The dominant mutant 7 cannot be assigned to a complementation group, because complementation testing cannot be performed on dominant mutations. The complementation groups are as follows:

4,6
1,8,3
2
5*
9*

*The ambiguity results from the fact that mutants 5 and 9 were never mated to one another and it's not know whether they are in the same complementation group. Therefore, there are either 4 or 5 complementation groups.

Note: You can infer that mutants 1 and 3 are in the same complementation group, because both are in the same complementation group as 8.

Assigning complementation groups using the canavanine resistant phenotype appears to be different than using auxotrophic markers such as his^- because the canavanine noncomplementors live, while the histidine noncomplementors die, on their respective selective media. However, the principle behind the two tests is the same.

(c) The problem is that strains 5 and 9 cannot be crossed to one another because they are of the same mating type, and the appropriate complementation tests cannot be done. One experiment that could resolve this ambiguity would involve mating the **a** strain 5 to an α wild type strain to form a diploid. Sporulation of this diploid will yield four types of spores:

mating type α can^r
mating type α can^s
mating type **a** can^r
mating type **a** can^s

The canavanine resistant spore that has the α mating type (which will have the same mutation as in strain 5) can then be crossed to the original strain 9 (mating type **a**). The mutants can then be assigned to the same or different complementation groups depending on whether the diploids are sensitive or resistant to canavanine, respectively. (This procedure can also be done with strain 9, which then can be mated back to the original strain 5).

2. (**a**) The short-tailed mutation (S) is dominant because in a cross with WT mice, we see short-tailed mice in the F1 generation. We can also determine that the original short-tailed mouse was heterozygous (S/+) because only 1/2 off the F1 have short tails. When two short-tailed F1 mice are crossed (S/+ x S/+), we predict a 3:1 ratio of affected to unaffected progeny, meaning that 3/4 of the F2 will have short tails. This is because mice of both genotypes S/+ and S/S will have short tails as the mutation is dominant.

(**b**) A true breeding strain would contain only mice homozygous for the short-tailed mutation (S/S), Failure to generate a true breeding strain probably results from a failure to produce homozygous mutant mice. That is, the only short-tailed mice in the stock are heterozygous for the mutation. This situation would result if the short-tailed mutation was homozygous lethal, with homozygous affected mice (S/S) dying before birth. According to this hypothesis, we expect a 2:1 ratio of heterozygous affected mice (S/+) to unaffected mice (+/+) from a cross of heterozygotes In other words, 2/3 of the progeny of a cross of short-tailed mice would have short tails.

(**c**) In order to distinguish the two hypotheses, we need to perform a chi-square test. We make the expected case homozygous lethality and the observed case homozygous viability.

	Expected (homozygous lethal)	Observed (homozygous viable)
Short tails	2/3	3/4
Long tails	1/3	1/4

In order to determine the number of progeny we need to score, we can use trial and error to narrow in on an approximate number. We want our answer to be significant at the p=0.05 level with one degree of freedom, so from the table our answer should be approximately 3.8. By solving the Chi square test using different numerical values for the observed and expected values, we can determine that we need to score slightly more than 100 progeny in order to determine which of the hypotheses is more likely.

We can also arrive at an exact numerical solution. If the total number of progeny that we need to score is N, then the number of individuals in a given class is the frequency of the class times N. We solve for N by substituting N times the frequency for each phenotypic class in the Chi square formula. We want our results to be significant at the p=0.05 level with one degree of freedom. From the table of Chi square values, this is 3.8. By this method, we arrive at an answer of 123 progeny.

(d) In order to determine whether these two short tailed mutations are allelic, we perform a complementation test by assaying homozygous lethality. This is possible because both of these mutations appear to have two distinct phenotypes, a dominant short tailed phenotype and a recessive embryonic lethality phenotype. If the two mutations are in the same gene, they will fail to complement each other and the homozygous mutant flies will be embryonic lethal. In this case, we expect a ratio of 2/3 short-tails to 1/3 long tails. If the two mutations are in two different genes, they will complement each other and none of the flies will die during embryogenesis. In this case, we expect a ratio of 3/4 short tails to 1/4 long tails.

(e) The test proposed in (d) is valid because it is a complementation test using a recessive phenotype, which in this case is embryonic lethality. While the short-tailed mutation is thought of as dominant with respect to tail length, it is recessive with respect to lethality because only mice with two copies of the mutation (S/S) die.

3(a) Let's call the recessive allele responsible for the trait 'a' and its wild type counterpart '+'.
The genotype of the son with the trait must be a/a. His parents must both be carriers of this trait and are of genotype a/+. So the probability of the next child being affected would be :
P(a/a)= 1/2 x 1/2 =1/4

If the next child doesnot exhibit the trait, his genotype could be +/+, +/a, but not a/a, so the probability of his genotype being +/a would be:

$$P(a/+ \mid +/+ \text{ or } +/a) = \frac{1/2}{1/2 + 1/4} = 2/3$$

(b) All males are hemizygous for the X chromosome. The son must be of genotype a/O. His father is not affected therefore is of genotype +/O. His mother must be the carrier from whom he inherited a. She is of the genotype +/a.

The next child would exhibit the trait if it is a boy of genotype a/O or a girl of genotype a/a.
P(affected) = P(a/O)+ P(a/a) =1/2 x 1/2 + 0 = 1/4

If the next child is a girl, she could be of genotypes +/a or +/+.

$$P(+/a \mid +/a \text{ or } +/+) = \frac{1/2 \times 1/2}{1/2 \times 1/2 + 1/2 \times 1/2} = 1/2$$

(c) The affected son is of genotype A/+ (capital A is used in this case to stand for the dominant allele.) One of his parents must be of the same genotype. The other could be of genotype +/+ or, less likely, A/+.

Let's assume only one parent carries A. Then
P(affected)= P(A/+) x 20% = 1/2 x 20% = 10%

$$P(\text{carrier} \mid \text{not affected}) = \frac{P(\text{carrier})}{P(\text{carrier}) + P(\text{non-carrier})}$$

$$= \frac{P(A/+) \times 80\%}{P(A/+) \times 80\% + P(+/+)} = \frac{1/2 \times 80\%}{1/2 + 1/2 \times 80\%} = 4/9$$

7.03 Problem Set 2

due in class Friday, September 30

1. You have obtained two mutant strains of the yeast *S. cerevisiae*. One strain is canavanine resistant (can^r) meaning that it can grow on the toxic compound canavanine, whereas wild-type strains are sensitive to canavanine (can^s). The other strain is a histidine auxotroph (his^-). In order to analyze the mutants, you cross each mutant strain to a can^s His^+ wild-type strain, sporulate the diploid, and analyze the resulting tetrads. You get the following results:

can^r x can^s always gives 2 can^s : 2 can^r spores
his^- x His^+ always gives 2 His^+ : 2 his^- spores

(a) What do these segregation patterns imply about the genes involved?

Next you wish to determine linkage between the *can* gene and the *his* gene. You mate a can^r His^+ strain to a can^s his^- strain and sporulate the diploid. Because you are in a hurry to map these genes you decide to not take the trouble to dissect tetrads. Instead you mix a large number of tetrads, break them apart, and analyze their phenotypes one spore at a time. Doing the analysis this way you do not know which spores came from the same tetrad. You get the following spores:

can^r His^+ x can^s his^- → 192 can^r his^+
170 can^s his^-
32 can^r his^-
29 can^s his^+

(b) Use these results to determine the map distance in centimorgans (cM) between the *can* gene and the *his* gene. For this problem map distance is calculated in the classical way as recombinant spores divided by total spores, quantity multiplied by 100 .

Your colleague maps the same genes by performing tetrad analysis. From the can^r His^+ x can^s his^- mating and subsequent sporulation, you find the following three classes of tetrads containing spores with the listed phenotypes:

Class I	**Class II**	**Class III**
can^r His^+	can^r His^+	can^r his^-
can^s His^+	can^r His^+	can^r his^-
can^r his^-	can^s his^-	can^s His^+
can^s his^-	can^s his^-	can^s His^+

There are 22 class I tetrads, 81 class II tetrads, and 3 class III tetrads.

(c) i) Designate each of the three classes of tetrads as PD, NPD, or T.

ii) Using this data calculate the map distance in cM between the *can* and *his* genes using the mapping formula for tetrad analysis. Which map distance is more accurate? Also explain why the map distance calculated from your data is smaller or larger than that calculated from your colleague's data.

You have recently isolated a mutant that is auxotrophic for arginine (arg^-). This mutation also segregates 2:2 when crossed to an Arg^+ wild-type strain as in part **a**. You want to map the *arg* gene relative to the *can* gene. Considering the results from the previous mapping experiment, you decide to perform tetrad analysis. From the can^r Arg^+ x can^s arg^- cross and subsequent sporulation, you find the following three classes of tetrads containing spores with the listed phenotypes:

Class I	**Class II**	**Class III**
can^r Arg^+	can^s Arg^+	can^r Arg^+
can^s Arg^+	can^s Arg^+	can^r Arg^+
can^s arg^-	dead spore	can^s arg^-
dead spore	dead spore	can^s arg^-

There are 28 class I tetrads, 4 class II tetrads, and 74 class III tetrads.

(d) Certain spores in the can^r arg^+ x can^s arg^- cross are not viable. What are the genotype of the dead spores? Give a physiological explanation why spores of this genotype will not grow. Think about what can^r means as described in problem #1 of problem set 1.

(e) i) Designate each of the three classes of tetrads as PD, NPD, or T.

ii) Using this data calculate the map distance in cM between the *can* and *arg* genes.

Next, you are given a mutant strain that is auxotrophic for tryptophan (trp^-). You know that this mutation is tightly linked to the centromere of its chromosome and that it segregates 2:2 when crossed to a Trp^+ (wild-type) strain. You want to map the *trp* gene with respect to the *can* gene. From the can^r Trp^+ x can^s trp^- cross and subsequent sporulation, you find the following three classes of tetrads containing spores with the listed phenotypes:

Class I	**Class II**	**Class III**
can^r Trp^+	can^r Trp^+	can^r trp^-
can^s Trp^+	can^r Trp^+	can^r trp^-
can^r trp^-	can^s trp^-	can^s Trp^+
can^s trp^-	can^s trp^-	can^s Trp^+

There are 12 class I tetrads, 42 class II tetrads, and 39 class III tetrads.

(f) i) Designate each of the three classes of tetrads as PD, NPD, or T.

ii) Are the *can* and *trp* genes on the same or different chromosomes? Calculate the map distance in cM between the *can* gene and its centromere.

2. In your Drosophila genetics lab class you are working with several 3rd chromosome eye mutations. One of these mutations, neon, (ne), is a recessive mutation that causes the flies to have bright pink eyes, (as opposed to the wildtype eye color which is red) . Another mutation, Lumpy, (Lp), is a dominant mutation that causes the ommatidia of the flies eyes to clump making the eyes look lumpy. You are asked to map the position of neon relative to lumpy and a third chromosome body mutation, ebony, (eb), that causes the flies to have a darker body color. You cross a female heterozygous for ebony, Lumpy and neon to a male that is homozygous for ebony and neon but wildtype for Lumpy. You get the following results:

+	eb	ne	1569
Lp	+	ne	13
Lp	eb	ne	3
Lp	+	+	1579
Lp	eb	+	405
+	eb	+	12
+	+	+	4
+	+	ne	415

(a) i) Draw a map of neon, (ne), relative to ebony, (eb), and lumpy, (lp).

ii) What were the genotypes of the true breeding parents of the female heterozygote used in this cross?

iii) Considering the fact that testcrosses are usually done by crossing to a homozygote why is it acceptable for you to cross the female heterozygote to a male that is wildtype for Lumpy?

Next, you are given another newly identified 3rd chromosome eye mutation, piglet, (a recessive mutation that causes the flies to have light pink eyes). You are asked to map the piglet mutation relative to Lumpy and ebony. You cross a female heterozygous for Lumpy, piglet and ebony to a male that is homozygous for Lumpy and ebony. You score the progeny and get the following results:

+	+	pg	1
+	eb	+	349
+	+	+	1393
Lp	+	pg	351
Lp	+	+	12
Lp	eb	+	2
Lp	eb	pg	1378
+	eb	pg	10

(b) i) What are the phenotypes of the true breeding parents of the female heterozygote?

ii) Draw a map of piglet relative to Lumpy and neon and ebony.

You are next asked to determine which of the eye mutations, if any, lie within the same gene. You decide to cross a female fly that is homozygous for piglet to a male that is homozygous for neon. You score the F1 progeny and find that all of the flies have pink eyes.

(c) i) Are piglet and neon in the same gene?

ii) Lumpy is dominant so you can't do a complementation test with Lumpy but given what you know about the mapping and complementation data what can you say about the gene in which Lumpy lies?

Your next assignment in Drosophila project lab is to determine the relative map positions of 3 newly-identified genes: plaid(pl), stripe(st), and polka dot(pd). Your TA has conducted a preliminary analysis of these loci in which she has determined that they are all on the X chromosome. You begin your analysis by crossing pl/pl, st/st, +/+ females and +, +, pd males. You then perform a testcross by crossing pl, st, pd males with the triply heterozygous F1 females. The following F2 progeny are scored:

+ + +	169
pl + pd	45
pl + +	42
pl st pd	180
+ + pd	174
+ st pd	37
pl st +	187
+ st +	53

(d) i) Draw a linkage map of these loci and indicate any ambiguities in the gene order.

ii) Do your results fit the assertion of the TA that all three mutations are on the same chromosome? In particular, discuss why the recombination frequencies between the 3 loci are not additive.

3. *C. elegans* is a small nematode worm. There are two sexes, males and self-fertilizing hermaphrodites which contain both eggs and sperm. It is possible to either cross males to hermaphrodites or allow the hermaphrodites to self-fertilize. Males (XO) crossed to hermaphrodites (XX) produce 1/2 male: 1/2 hermaphrodite progeny. When selfed, hermaphrodites produce hermaphrodites. One of the great advantages of *C. elegans* as a genetic organism is that true breeding mutant lines can be made easily by several rounds of selfing of hermaphrodites.

Selfing a hermaphrodite gives the same result as crossing two animals of the same genotype together. Given the mutation (a) recessive to the wild-type allele (A):

A/a x A/a

↓

1/4 A/A
1/2 A/a
1/4 a/a

A/a

↓ (selfed)

1/4 A/A
1/2 A/a
1/4 a/a

Mapping experiments involving hermaphrodites often take advantage of the fact that heterozygous mutations become evident in the next generation when the animals are selfed.

You are interested in the worm nervous system, so you do a screen to isolate behavioral mutants in hermaphrodites. Three of the mutants breed true and are given the names *spz* (*spaz*), *unc* (*unc*oordinated) and *slo* (*slow*). Worms with these phenotypes are designated Spz, Unc, and Slo. You cross wild-type males to each of the mutants and find that all F1s are wild-type.

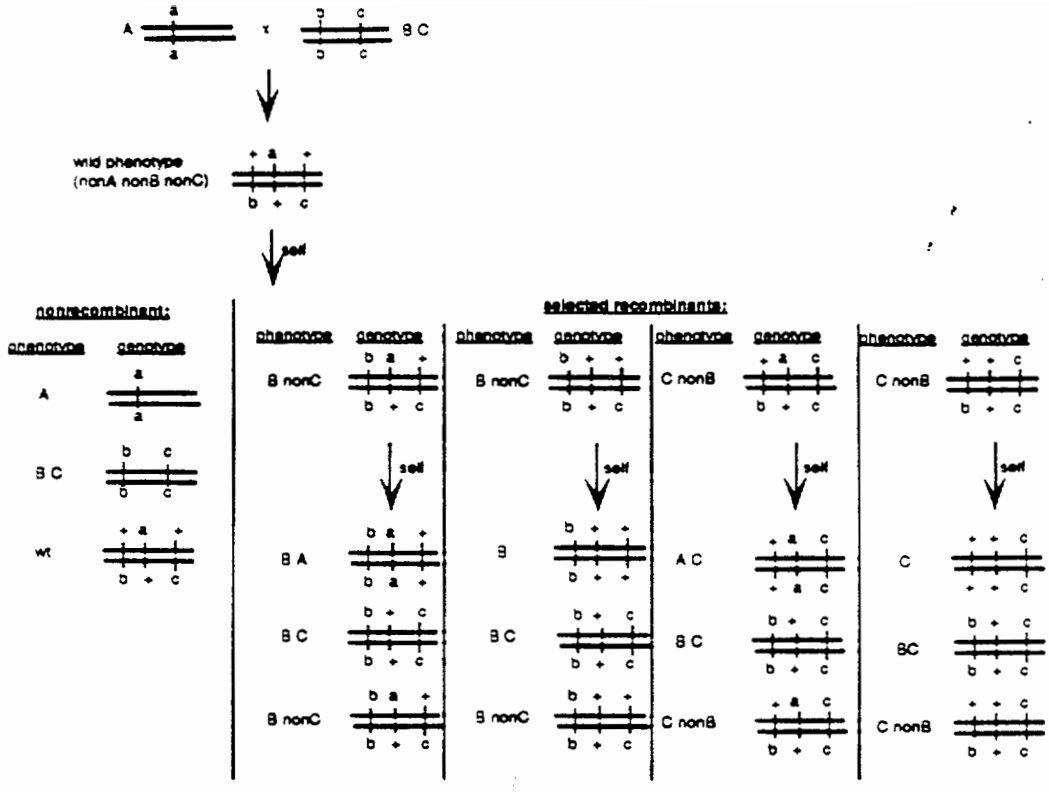
(a) *dpy* (*dumpy*) is a well-characterized mutation on chromosome III (an autosome) which is useful for mapping. You cross *dpy* males to *unc* hermaphrodites. All F1s are wild-type. Hermaphrodite F1s are selected and selfed. The following progeny are seen:

Phenotype	number
wild-type	1023
Dpy nonUnc	482
Unc nonDpy	501
Dpy Unc	1

Is the *unc* mutation linked to *dpy*?

(b) Next you cross wild-type males to *dpy unc* hermaphrodites. F1 hermaphrodites are selected and selfed. You find that 1% of F2s are Dpy nonUnc. What is the map distance between *dpy* and *unc*? (Be careful!)

(c) It is possible to directly select recombinants between two mutations and use the data to determine the position of a third mutation with respect to the first two. For example, to map a mutation a with respect to linked b and c, you would select phenotypically B nonC and C nonB F2 animals as shown below. If 50% of B nonC animals produce A progeny, and 50% of C nonB animals produce A progeny, a lies halfway between b and c.



You find that the *slo* mutation is also linked to *dpy*. When a *dpy unc* hermaphrodite is crossed to a *slo* male, all the cross progeny are wild-type. The F1 are selfed and you select 30 Unc nonDpy F2 hermaphrodites. When selfed, 6/30 of these animals segregate F3 Slo progeny and 24/30 never segregate Slo progeny. You also select 30 Dpy nonUnc F2 hermaphrodites. 20/30 of these animals segregate F3 Slo progeny and 10/30 do not. Where does *slo* map with respect to *unc* and *dpy*? Draw a map, including map distances. For your calculation, use the distance determined in part b.

d) The Spz phenotype is very similar to the Unc phenotype. You cross *spz* males to *unc dpy* hermaphrodites. The F1 cross progeny are all Unc.

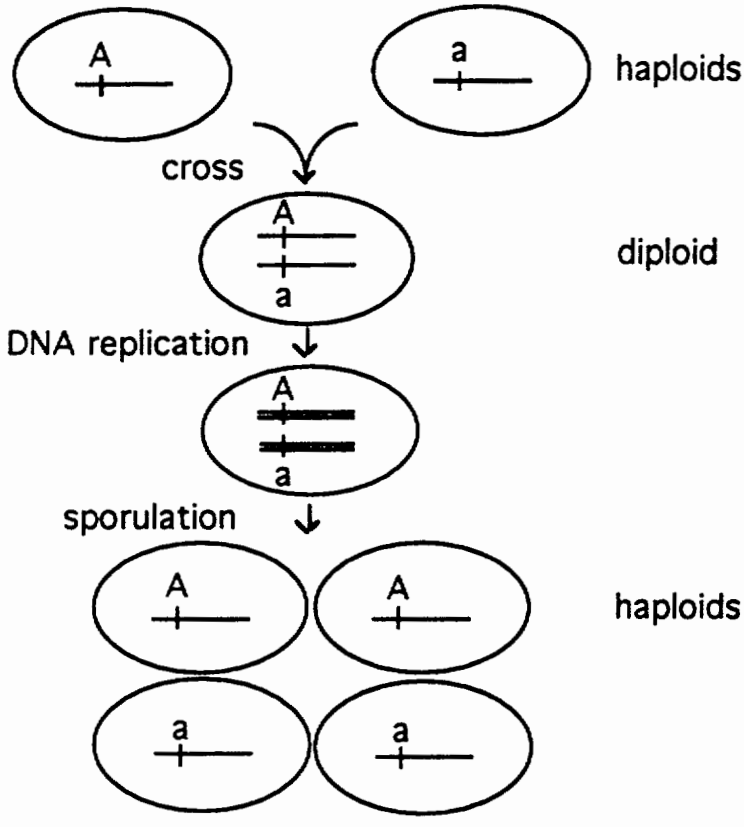
i) Does *unc* complement *spz*?

ii) You perform a second cross of *spz* males to *dpy slo* hermaphrodites and self the F1s. You select Slo nonDpy F2 hermaphrodites. When selfed, 0/30 segregate F3 Spz progeny. What proportion of Dpy nonSlo F2s are expected to segregate Spz progeny?

iii) Propose a model for the relationship between the *spz* and *unc* mutations which explains the results above.

7.03 Problem Set 2 Answers

1. (a) As shown below, when you cross a wild type strain with a mutant strain, 2:2 segregation indicates that there are two alleles of a single gene segregating.



2 A: 2 a

(b) Using the formula given: map distance equals one hundred times the number of recombinant spores divided by the total number of spores, one can calculate the distance simply by recognizing which are the recombinant spores. The recombinants in this cross are the can^r his^- and the can^s his^+ spores, and therefore the calculation is as follows:

$$\frac{32+29}{192+170+32+29} \times 100 = 14.4 \text{ cM}$$

This method for determining map distance is analogous to that in *Drosophila* and other higher diploid organisms. The difference lies in the fact that there is a free-living haploid phase in the yeast life cycle, and

therefore, the phenotypes of the haploids can be determined. Whereas, in *Drosophila*, the haploid gametes are not free-living, and the phenotypes can only be determined in diploids.

(c) i) The class designations are as follows:

> Class I tetrads are the tetratypes. They are recognized by the fact that they contain all four possible combinations of the markers, both parental and recombinant.
>
> Class II tetrads are the parental ditypes. They are recognized by the fact that they contain only spores of the parental genotypes, two spores of each genotype.
>
> Class III tetrads are the nonparental ditypes. The are recognized by the fact that they contain only spores of the recombinant genotypes, two spores of each genotype.

ii) We know these genes are linked because the number of nonparental ditypes is much less than the number of parental ditypes. Using the following formula for determining map distance between two linked genes:

$$\left(1/2\right)\left(\frac{T+6NPD}{PD+T+NPD}\right)\left(100\right) = \text{map distance}$$

the map distance between the *can* and *his* genes is calculated as follows:

$$\left(1/2\right)\left(\frac{22+6(3)}{81+22+3}\right)\left(100\right) = 18.9 \text{ cM}$$

The map distance calculated using your colleague's data is more accurate because the map distance calculated using random spore analysis does not take into account double recombinants. Some of these double recombinants will be counted as single recombinants, and others are not detected because they have the parental genotype. Because you observe less recombination events by random spore analysis than you are able to observe by tetrad analysis, the map distance is smaller for random spore analysis.

(d) The genotype of all the dead spores is can^r arg^-. One explanation for the lack of viability is as follows: Since canavanine is imported into the cell via the arginine transporter (see problem#1 in problem set 1), mutants for canavanine uptake may not be able import arginine from the media. If

a can^r strain that has a defect in the arg transporter is also auxotrophic for arginine (meaning it cannot synthesize its own arginine) it has no source of the amino acid arginine and therefore cannot grow, even on media supplemented with arginine.

(e) i) First, we know the mutant phenotype is governed by a single locus, because the mutation segregates 2:2 when crossed to wild type (see part a). Knowing the genotype of the dead spores is can^r arg^-, the class designations are as follows:

Class I tetrads are the tetratypes
Class II tetrads are the nonparental ditypes
Class III tetrads are the parental ditypes

ii) We know the two loci are linked (for reasons explained in part c), and therefore the map distance between the *can* and the *arg* genes can be calculated as follows:

$$\left(1/2\right)\left(\frac{28+6(4)}{74+28+4}\right)\left(100\right) = 24.5 \text{ cM}$$

(f) i) Again, we know the mutant phenotype is governed by a single locus, because the mutation segregates 2:2 when crossed to wild type. The class designations are as follows:

Class I tetrads are the tetratypes
Class II tetrads are the parental ditypes
Class III tetrads are the nonparental ditypes

ii) First, we are given that the *trp* gene is tightly linked to its centromere. Since the number of parental ditypes approximately equals that of the nonparental ditypes, and the number of parental ditypes is greater than one-fourth the number of tetratypes, we know the *can* and *trp* genes are linked to different centromeres. Therefore we also know the genes are on different chromosomes. Given that the *trp* gene is tightly linked to its centromere, the distance between the *can* gene and its centromere is calculated by the following formula:

$$\left(1/2\right)\left(\frac{T}{PD+T+NPD}\right)\left(100\right) = \text{map distance from gene to centromere}$$

The map distance between the *can* gene and its centromere is calculated below:

$$\left(1/2\right)\left(\frac{12}{42+39+12}\right)\left(100\right) = 6.5 \text{ cM}$$

2. (a) The map of *neon* relative to *Lumpy* and *ebony* is the following:

eb ——— *Lp* ——— *ne*
20.7 **0.8**

First, determine the order of the mutations by comparing the genotypes of the nonrecombinant progeny (the largest class of progeny) with that of the double recombinant progeny (the smallest class of progeny). One can determine which mutation is in the middle by determining which mutation has to be "flipped" in the parental chromosomes to give us the double recombinant chromosomes.

Nonrecombinant		**Double recombinant**
+ eb ne	vs	*Lp eb ne*
Lp + +		*+ + +*

Since *Lp* is the mutation that is "flipping" we know that it is the mutation in the middle and thus the order is:

eb Lp ne or *ne Lp eb*

Without information concerning the place of *eb* in relation to the other markers on the chromosome it is impossible to distinguish between the two orders written above.

To determine the distance between *eb* and *Lp* and *Lp* and *ne* simply count the number of progeny that result from a recombination event between the two genes respectively and divide this number by the total number of progeny and multiply by 100.

Progeny resulting from a crossover between *Lp* &*eb*	
Lp eb ne	3
+ + +	4
Lp eb +	405
+ + ne	415

Progeny resulting from a crossover between *Lp* & *ne*	
Lp + ne	13
+ eb +	12
+ + +	4
Lp eb ne	3

827 32

$\frac{827}{4000} \times 100 = 20.7$ $\frac{32}{4000} \times 100 = 0.8$

To determine the distance between *eb &ne* , add up the total number of progeny resulting from a single crossover between *eb &ne* and add to this twice the number of double recombinant progeny and divide the total by the total number of progeny and multiply by 100.

Progeny resulting from a crossover between *eb &ne*

Lp eb +	405
+ + *ne*	415
Lp + *ne*	13
+ *eb* +	12
	845

$\frac{(845+2x7)}{4000} \times 100 = 21.5$ mu

The distance between *eb&ne* should be the sum of the distance between *Lp & eb* and *Lp&ne* since map distances are additive.

(ii) The genotype of the parents of the female heterozygote can be determined by looking at the phenotypes of the nonrecombnant progeny. In this case the phenotype of the nonrecombinant progeny are + *eb ne* and *Lp* + + and thus the genotype of the female heterozygote is $\frac{+\ eb\ ne}{Lp\ +\ +}$

Therefore, the genotypes of the parents must be +/+, *eb/eb, ne/ne* and *Lp/Lp*, +/+, +/+.

Parents $\frac{+\ eb\ ne}{+\ eb\ ne} \times \frac{Lp\ ++}{Lp\ ++}$

Female heterozygote $\frac{+eb\ ne}{Lp\ +\ +}$

(iii) In *Drosophila,* recombinant progeny are scored by their phenotype. Thus, when trying to map recessive mutations, one usually crosses to a male that is homozygous recessive for the trait in order to be able to

visualize and score the mutation and thus be able to score recombinant progeny properly.
(b) (i) Using the same logic we used in part a (ii), the genotypes of the parents are +/+, +/+, +/+ and *Lp/Lp*, *eb/eb*, *pg/pg* and thus the phenotypes are wildtype and Lumpy, ebony, piglet.
(ii) Using the same logic we used in part a (i) the map is:

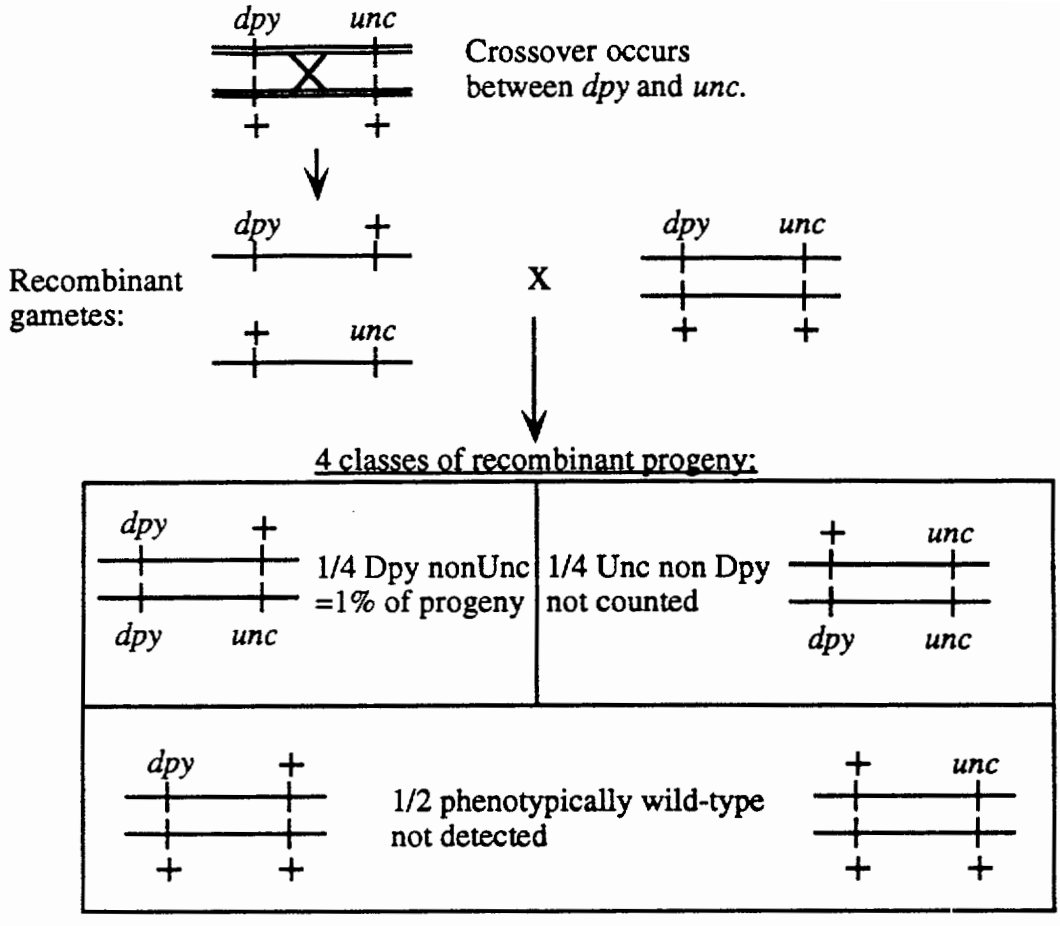
eb		pg	Lp	ne
	20.1		**0.7**	**0.8**

(c) (i) The wildtype phenotype for *Drosophila* eye color is red eyes. If *ne* and *pg* complemented,(ie were in different genes), the flies would have red eyes. Since all of the flies have pink eyes we can assume that the mutations fail to complement and are thus in the same gene.
(ii) The complementation data indicate that *ne* and *pg* are mutations in the same gene. Since *Lumpy* maps between *piglet* and *neon* and effects the phenotype of the eye, it is likely that the *Lumpy* mutation lies within the same gene as the *ne* and *Lp* mutations.

(d) i) In order to solve a three-point mapping problem, it is necessary to know which alleles are on each homolog in the F1 female. In this case, we are told the genotypes of each of her parents, so we know that her genotype is pl st + / + + pd. Next, we want to determine the order of the three genes, which we can do by identifying the double recombinant classes and inferring the original gene order from them. In this problem, there are two possible gene orders because four of the classes of F2 progeny are approximately equivalent. Therefore, the two possible orders are pl st pd and st pl pd. In order to calculate the genetic distance between each of the loci, we simply count up the number of recombinant gametes within each interval.

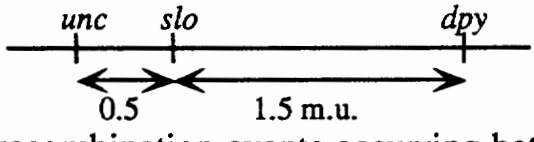
pl-st interval: 45+42+37+53/887 = 20 cM
st-pd interval: 180+169+37+42/887 = 48 cM
pl-pd interval: 180+169+53+45/887 = 50 cM

From these data, we conclude that pl and st are linked to each other, and neither one is linked to pd. Therefore, either of the following maps is accurate.



The cross above shows that we only detect 1/4 of recombination events in our experiment.
Another important factor to remember is that both male and female germlines are heterozygous *dpy unc*/+ +. Here we have measured the recombination rate in both germlines together. Assume that recombination rate in the two are equal (this has been experimentally shown).
1% x 4 x 1/2 => 2 m.u.

c)



Fraction of recombination events occurring between *unc* and *slo* = (6 + 10)/(30 + 30) = 4/15
map distance between *unc* and *slo* = 4/15 x 2 m.u. = 0.5 m.u.

Fraction of recombination events occurring between *slo* and *dpy* = (24 + 20)/(30 + 30) = 11/15
map distance between *slo* and *dpy* = 11/15 x 2 m.u. = 1.5 m.u.

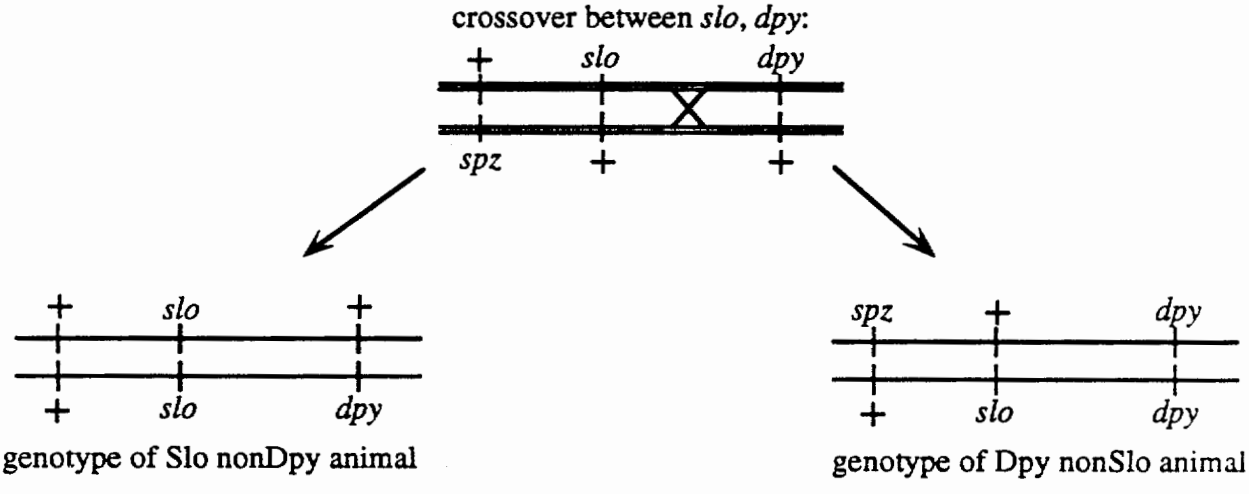
d) i) The cross is: male *spz/spz* x *unc dpy/unc dpy* hermaphrodite
NOTE: Here, *dpy* is a useful marker to distinguish self progeny from cross progeny.

Case 1: If spz and unc complement each other, they are mutations in different genes. The F1 genotype can be written *spz/+ , unc dpy/+ +* and the F1 phenotype should be nonUnc.

Case 2: If spz and unc are mutations in the same gene, they will fail to complement each other. The F1 genotype is written *spz +/unc dpy*, and the F1 phenotype is Unc.

The results indicate that *unc* does not complement *spz*.

ii) 30/30 of Dpy nonSlo F2s should segregate Spz progeny.
The mapping results are consistent with *spz* mapping outside the interval between *dpy* and *slo*:



genotype of Slo nonDpy animal

genotype of Dpy nonSlo animal

The mapping data are also consistent with the results of the complementation test.

iii) In part i) we found that *unc* and *spz* fail to complement each other. In part ii) we found that *spz* and *unc* map to the same place with respect to known markers. Together, these data strongly support the hypothesis that *unc* and *spz* are mutations in the same gene.

7.03 Problem Set 3

due in class October 19,1994

1. The lysogenic phage Ω produces turbid plaques when grown on a lawn of the bacterium *Bacillus subtilis*. You have isolated five mutants of phage Ω, all of which produce clear plaques when grown on *B. subtilis*. All five mutant strains (**a** through **e**) were determined to be recessive with respect to wild-type. In order to determine which mutations are located in the same gene, you decide to perform a complementation test. This is done by infecting the host bacterium with pairwise combinations of the mutant phage and determining whether the spots of growth are clear or turbid.

	a	b	c	d	e
a	clear	turbid	clear	clear	turbid
b		clear	turbid	turbid	clear
c			clear	clear	turbid
d				clear	turbid
e					clear

a. Assign the mutations to complementation groups.

Next you wish to determine where the genes are located in the phage genome. You decide to map them with respect to a known essential gene in which you have an amber mutation, am1. When the am1 mutant phage is plated on a lawn of the wild-type bacteria (restrictive host) no plaques are formed. But, in a *B. subtilis* amber suppressing strain (permissive host), the am1 phage will produce normal plaques. In order to map the position of the clear genes you pick two mutants (**a** and **b**) to map with respect to am1. You coinfect the amber suppressing strain with the am1 mutant phage and each of the two clear mutants. The resulting phage are titered on both the permissive and restrictive host strains. The total number of plaques formed on the permissive host is ~$1X10^4$ for each of the crosses. The number of turbid plaques formed on the restrictive host is given in the table below:

am1 x **a**	223
am1 x **b**	382

b. Calculate the distance in map units between the am1 mutation and two of the clear mutations.

You now wish to map all the mutations in the clear genes with respect to one another. This is done by infecting the wild-type bacteria with pairwise combinations of the mutants and titering the resulting phage on the wild-type bacteria. The total number of plaques formed for each cross is ~$1X10^4$, and the number of turbid plaques formed for each cross is given in the table below:

	a	b	c	d	e
a	-	146	52	53	228
b		-	204	201	77
c			-	5	260
d				-	258
e					-

c. Draw a map of the mutations, including am1, and show approximate gene boundaries. Remember to include map distances between the mutations.

Since mutations **c** and **d** are so close to one another, their order with respect to the other mutations is ambiguous. To assign the order of the two mutations, you perform a three factor cross between the **c** and **d** clear mutations, and am1. Since you are dealing with three genes and only two phenotypes, you will not be able to assign all of the recombinant classes. However, you will be able to determine gene order by setting up two crosses one of which will produce wild-type plaques by a double crossover and the other by a single crossover. In the first cross, an am1 mutant phage which also has the **c** mutation is crossed to a phage with only the **d** mutation, In the second cross, an am1 mutant phage which also has the **d** mutation is crossed to a phage with only the **c** mutation. You coinfect on the permissive host (amber suppressing strain) and titer on both the permissive and restrictive hosts. In order to achieve an increased sensitivity for this experiment you increase the amount of plaques that are formed on the permissive host to $5X10^4$. The number of turbid plaques that are formed on the restrictive host for each of the two crosses is given below:

cross 1 1 turbid plaque

cross 2 26 turbid plaques

d. Draw a new map of the mutations including the order of the **c** and **d** mutations (do not include map distances).

A friend in your lab sequenced the clear mutations and determined that mutation **c** is a missense mutation and mutation **d** is a frameshift mutation in which one base pair has been added (+1 reading frame). You have isolated a new mutation (**f**), which your friend sequenced as well and determined that it is also a frameshift mutation but one in which one base pair has been deleted (-1 reading frame). She also informs you that mutation **c** is a change in the same base pair that is deleted in mutant **f**. To be thorough, you decide to map mutation **f** with respect to the other clear mutations. You do this by infecting the wild-type *B. subtilis* with pairwise combinations of mutant **f** and the other mutants, titering the resulting phage on the wild-type host, and counting the number of turbid plaques that are formed. The map distances determined by these crosses are approximately the same for mutant **f** as for mutant **c**, except the distance between **f** and **d** is two times larger than the distance between **c** and **d**. You repeat the experiment several times and always come up with the same result.

e. Give a logical explanation for this discrepancy in the map distance.

Phage Ω is known to have a recombination frequency of 2 m.u./1000 bp.

f. Given the map distance between mutations **c** and **d** (calculated in part c), approximately how many base pairs are there between these mutations?

3. a. You want to order and determine the the relative distance between five auxotrophic mutations, f-, g-, h- i-and j-. You infect strain 1, (f+, g+ h+, i+, j+) with transducing phage P1. The lysate is used to infect strain 2, (f-, g-, h-, i-, j-). The cotransduction analysis is carried out by selecting for one marker and then screening 100 selected clones for the presence of other markers. You get the following data:

Selected marker	Screened marker	# Out of 100 selected clones
g+	i+	94
g+	h+	75
g+	j+	8
g+	f+	70
h+	i+	78
i+	f+	75
i+	j+	12
j+	h+	39
f+	h+	95

(i) Order the markers and indicate their relative distance from each other. Indicate any ambiguities.

b. To clear up some of the ambiguities you repeat the experiment above and select 1000 j+ colonies for further study. The data are shown below.

Screened markers	Number of colonies
f-h-	600
f+h+	345
f+h-	50
f-h+	5

(i) Does this data clear up any of the ambiguities? If so, explain how. Draw out the crossovers that gave rise to each of the four classes.

Note: Questions "2 + "4 were purposely omitted.

7.03 Problem Set 3 Answers

1. a. When a combination of two mutant phage produces a turbid plaque, the two mutations are in different complementation groups, and when a combination of two mutants produces a clear plaque, the mutations are in the same complementation group. There are only two complementation groups for the five clear mutations. Mutations **a**, **c**, and **d** are in one complementation group and mutations **b** and **e** are in the other.

b. The distance these two mutations are from one another is proportional to the number of recombinants that are produced when the phage are crossed. The recombinants from the cross between the am1 mutant phage and the **a** clear mutant are the am1 **a** mutant phage and the wild-type phage. The growth characteristics of all the phage from this cross are as follows (this also applies to the am1 x **b** cross):

Phage genotype	Restrictive host (wild-type)	Permissive host (amber suppressor)
am1 +	no plaques	turbid plaques
+ **a**	clear plaques	clear plaques
am1 **a**	no plaques	clear plaques
+ +	turbid plaques	turbid plaques

Therefore to calculate the map distance one would count the number of turbid plaques formed on the restrictive host and divide this by the total number of plaques formed on the permissive host and multiply the quantity by one hundred.. This number is then multiplied by two, because the wild-type phage, which produce turbid plaques on the restrictive host, are only one-half of the recombinant progeny. The calculation for the am1 x **a** cross is shown below:

$$\text{am1 x a} \quad \frac{223}{1 \times 10^4} \times 2 \times 100 = 4.5 \text{ m.u.}$$

c. Again the distance these two mutations are from one another is proportional to the number of recombinants that are produced when the phage are crossed. The recombinants from the mutant **a** and mutant **b** cross are the **a b** double mutant phage and the wild-type phage. The growth characteristics of all the phage that result from this cross are shown below:

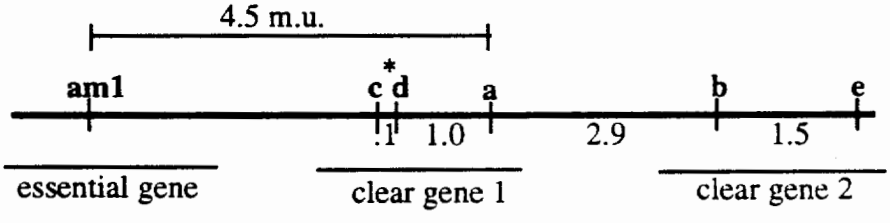
Phage genotype		Wild-type host
b	+	clear plaques
+	**a**	clear plaques
b	**a**	clear plaques
+	+	turbid plaques

In this case one would calculate the map distance by counting the number of turbid plaques formed and dividing this by the number of total plaques formed. (Note there is no restrictive host being used in this cross.) This number is then multiplied by one hundred. Assuming that the double mutants will show the clear phenotype, we again multiply this number by two because we are only counting half of the recombinants. A sample calculation for the **a** x **b** cross is shown below:

$$\mathbf{a \times b} \quad \frac{146}{1 \times 10^4} \times 2 \times 100 = 2.9 \text{ m.u.}$$

The map, distances between mutations, and approximate gene boundaries are shown below:

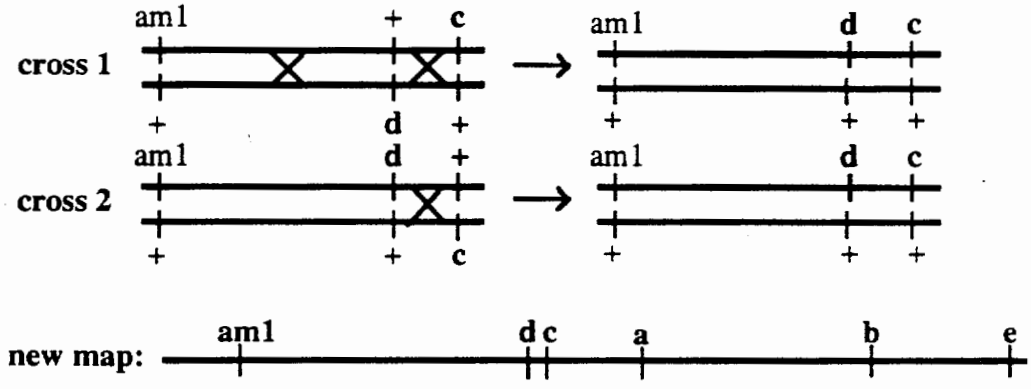
	a	b	c	d	e
a	-	2.9	1.0	1.0	4.6
b		-	4.1	4.0	1.5
c			-	0.1	5.2
d				-	5.2
e					-



* Note that the order of genes d and c, and the gene boundaries are ambiguous

d. First, we know that only the wild-type phage (+ + +) will produce turbid plaques on the restrictive host and that the only way to obtain these phage is by recombination. Since the cross between the am1 **c** double

mutant phage and the **d** mutant phage (cross 1) produces far less wild-type recombinants than in cross 2, we infer that these wild-type recombinants were produced by a double crossover, as shown in the figure below. Therefore, the wild-type phage from the cross between the am1 **d** double mutant phage and the **c** mutant phage (cross 2) were produced by a single crossover between **c** and **d**. Therefore we know that the order has to be am1-d-c.



e. Since mutations **c** and **f** map to the same exact location, we know that the distance from mutation **d** to each of the mutations **c** and **f** has to be the same. In a mapping experiment such as this, only half of the recombinants between two mutations will yield wild-type phage, while the other half of the recombinants will be double mutants. We expect that the double mutants will exhibit the same phenotype as the single mutants (clear plaques) and assume they cannot be observed. Therefore, we count the turbid plaques, which represent the half of the recombinant phage which are wild-type, and multiply this number by two to get the real number of recombinants (as described above). However, in this case, the double mutant consists of two frameshift mutations which compensate for each other (one adds a base pair, and one deletes a base pair). Therefore, the double mutant has the wild-type phenotype and produces turbid plaques. Since we already counted all of the recombinants by counting the turbid plaques, there is no need to multiply the distance by two, and hence the discrepancy in the map distances.

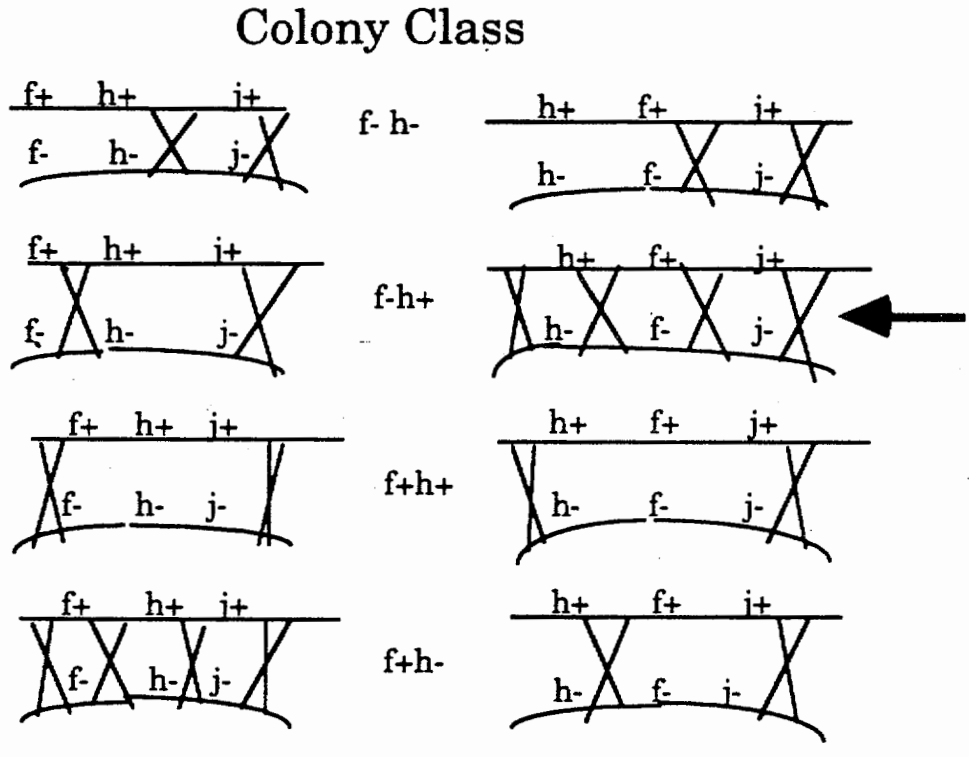
f. The map distance between the **d** and **c** mutations is 0.1 m.u., so the number of base pairs between these two mutations is calculated as shown below(the number of bp between mutations **d** and **f** is the same since mutations **c** and **f** map to the exact same location):

$$\frac{1000 \text{ bp}}{2 \text{ m.u.}} \times 0.1 \text{ m.u.} = 50 \text{ bp}$$

3. (a) In order to determine the order of the markers on the chromosome, we need to know some cotransduction frequencies. The cotransduction frequency is simply defined as the number of clones that contain two particular donor markers divided by the total number of clones selected. This is a measure of how often the two markers get packaged into the same phage head which in turn is a measure of the distance between the markers. in this problem, we are given the number of clones out of 100 that show cotransduction for the two markers. Therefore, the cotransduction frequency is simply equal to the (number of clones/total number of clones or 100) x 100. The contransduction frequency in this problem is the same as the number of clones.

Using the knowledge that cotransduction frequencies are inversely related to distance, we observe that i+ and g+ are close as are f+ and h+. Similarly, i+ and g+ are about the same distance from f+ and h+. The error inherent in these measurements does not allow us to order i+ and g+ relative to f+ and h+ nor f+ and h+ relative to either i+ or g+. We also see that j+ is far from f+, but even farther from i+ (probably more than one headful since we see no cotransductants). This information places j+ on the other side of

(b) These data can be used to resolve the ambiguity in the order of f+ and h+ relative to the j+ marker. This experiment is similar to a three factor cross in the sense that we will use the rare class of crossover recombinants to determine the marker order. To do this, draw out the crossovers that give rise to the four classes of colonies for the two possible orders, f+, h+, j+ and h+ ,f+, j+. These are shown below. The rarest class of recombinants results from the most crossovers. In this case, the class is j+,f-,h+, which results from a quadruple crossover event. Only if the order is h f j will a quadruple crossover give rise to a j+, f-, h+ colony.



---

4. a. i) The a gene must be contained within clone #1 since only clone #1 is sufficient to complement the mutation in strain A. Similarly, the b gene must be contained within clone #2. Gene a and b are right next to each other. Since both clone X and clone 1 and 2 when cotransformed complement the c mutation, whereas clone 1 or 2 alone cannot, it is clear that both gene a and gene b wild-type activities are required to complement the c mutation. The C strain must be mutant for both gene a and gene b.

7.03 Problem Set 4

Due in class Monday, October 24

1. You are studying the hypothetical nut (nitrogen utilization) operon in the bacterium *Bacillus subtilis*. This bacterium is able to use glutamine or ammonia as a nitrogen source. It is known that one of the genes in the operon, the nutC gene, encodes an enzyme which converts glutamine to glutamate and ammonia (NH_3), that the cell can then use as a nitrogen source. It is also known that the presence or absence of ammonia does not affect the levels of nutC activity. When the bacteria are grown on media with glutamine and ammonia, the levels of nutC activity are high, and when the bacteria are grown on ammonia alone the levels of nutC are low. You have isolated mutants in two complementation groups, nutB and nutA, which affect glutamine's control of nutC activity. You construct the following strains to analyze the regulation of the nut operon:

		Levels of nutC activity	
	Strains	**- glutamine**	**+ glutamine**
1)	$nutA^+$ $nutC^+$	15	200
2)	$nutA^+$ $nutC^-$	0	0
3)	nutA-1 $nutC^+$	0	0
4)	$nutA^+$ $nutC^+$/ F' $nutA^+$ $nutC^+$	30	400
5)	nutA-1 $nutC^+$/ F' $nutA^+$ $nutC^+$	15	200
6)	nutA-1 $nutC^+$/ F' $nutA^+$ $nutC^-$	0	0
7)	nutA-1 $nutC^-$/ F' $nutA^+$ nutC+	15	200

For questions a-c indicate which strains (by number) you used in determining your answer.

a. i) Is the nutA-1 mutation dominant or recessive?

ii) Does the nutA-1 mutation act in cis or trans?

Strains	Levels of nutC activity - glutamine	+ glutamine
8) nutB$^+$ nutC$^+$	15	200
9) nutB$^+$ nutC$^-$	0	0
10) nutB-1 nutC$^+$	0	0
11) nutB$^+$ nutC$^+$/ F' nutB$^+$ nutC$^+$	30	400
12) nutB-1 nutC$^+$/ F' nutB$^+$ nutC$^+$	30	400
13) nutB-1 nutC$^+$/ F' nutB$^+$ nutC$^-$	15	200
14) nutB-1 nutC$^-$/ F' nutB$^+$ nutC$^+$	15	200
15) nutB-2 nutC$^+$	200	200
16) nutB-2 nutC$^+$/ F' nutB$^+$ nutC$^+$	400	400
17) nutB-2 nutC$^+$/ F' nutB$^+$ nutC$^-$	200	200
18) nutB-2 nutC$^-$/ F' nutB$^+$ nutC$^+$	200	200

b. i) Is the nutB-1 mutation dominant or recessive?

ii) Does the nutB-1 mutation act in cis or trans?

c. i) Is the nutB-2 mutation dominant or recessive?

ii) Does the nutB-2 mutation act in cis or trans?

d. What are the expected levels of nutC activity (with and without glutamine) in the following strains:

i) nutB-2 nutA-1 nutC$^+$
ii) nutB-1 nutA$^+$ nutC$^+$/ F' nutB-2 nutA$^+$ nutC$^+$
iii) nutB-2 nutA-1 nutC$^+$/ F' nutB$^+$ nutA$^+$ nutC$^+$

e. Present a molecular model for this regulation and draw a diagram showing the role of glutamine, DNA sites, and trans acting elements which are needed for the regulation.

2. You are interested in smactose metabolism in *Escherichia fakii.* Wild-type cells are capable of breaking down the smactose into sictose and moctose in a reaction catalysed by the smactase enzyme. When *E. fakii* cells are grown on media without smactose, smactase activity is not induced. You have isolated strains mutant in three unlinked genes involved in smactase regulation and activity.

genotype	- smactose	+ smactose
smcA+ smcB+ smcC+	0	100
smcA- smcB+ smcC+	0	0
smcA+ smcB- smcC+	100	100
smcA+ smcB+ smcC-	0	0
smcA+ smcB- smcC-	100	100
smcA- smcB- smcC+	0	0
smcA- smcB+ smcC+/ F' smcA+	0	100
smcA+ smcB- smcC+/ F' smcB+	0	100
smcA+ smcB+ smcC-/ F' smcC+	0	100

You mate a strain carrying the amber-suppressing tRNA mutation on F' (F'supE) with the F$^-$ smcA- smcB+ smcC+ strain. You isolate the merodiploid smcA- smcB+ smcC+/ F'supE and find that smactase activity is restored to wild-type level in this strain.

i)What is the nature (recessive or dominant) of the smcA mutation?
What is the molecular nature of the smcA mutation?

ii) What is the nature (recessive or dominant) of the smcB mutation? How does smcB function in smactase regulation?

iii) What is the nature (recessive or dominant) of the smcC mutation? How does smcC function in smactase regulation?

iv) smcA- and smcC- strains have similar phenotypes in the tests above, but they behave differently in double mutant combinations with smcB-. Based on this data, is smcA or smcC more likely to be the structural gene for the smactase enzyme?

v) Propose a genetic pathway for smactase regulation that is consistent with all the data above and explains the role of smcA, smcB, and smcC in smactase production and regulation.

Another mutant strain (smcD-) is isolated. smcD is found to be closely linked to smcB by P1 transduction experiments. When mutant strains are assayed for smactase activity, you find:

genotype	- smactose	+ smactose
smcA+ smcB+ smcC+ smcD+	0	100
smcA+ smcB+ smcC+ smcD-	0	0
smcA- smcB+ smcC+ smcD-	0	0
smcA+ smcB- smcC+ smcD-	100	100
smcA+ smcB+ smcC- smcD-	0	0
smcA+ smcB+ smcC+ smcD-/ F' smcB+ smcD+	0	0
smcA+ smcB- smcC+ smcD-/ F' smcB+ smcD+	0	100
smcA+ smcB- smcC+ smcD+/ F' smcB+ smcD-	0	0

vi) What is the nature (recessive or dominant) of the smcD mutation? Is the smcD mutation cis- or trans- acting with respect to smcB? Propose an explanation for the behavior of the smcD mutation consistent with the model proposed in part v.

7.03 PROBLEM SET 4
ANSWERS

1. **a. i)** The nutA-1 mutation is recessive to wild-type ($nutA^+$). This can be determined from strain 5, where the merodiploid has a copy of both nutA-1 and $nutA^+$, but it has the inducible phenotype of the wild-type strain 1 instead of the uninducible phenotype of the strain with the nutA-1 mutation (strain 3).

ii) The nutA-1 mutation acts in cis. This can be determined from strains 6 and 7. In strain 6, the nutA-1 uninducible phenotype is expressed when nutA-1 is in cis to $nutC^+$. In strain 7, $nutA^+$ is in cis to $nutC^+$, and there are wild-type levels of nutC activity. Therefore, nutA is the operator site for nutC.

Note: The levels of activity are doubled in strain 4 because there are two wild-type copies of the nutC gene which includes the region of DNA encoding the protein (nutC) and the operator (nutA). This explains why the levels of the merodiploid strains 5 and 7 are identical to the wild-type haploid (strain 1) instead of being doubled. In both strains 5 and 7 there is only one functional copy of the nutC gene.

b. i) The nutB-1 mutation is recessive to wild-type ($nutB^+$). This can be determined from strain 12, where the merodiploid has a copy of both nutB-1 and $nutB^+$, but it has the inducible phenotype of the wild-type strain 8 instead of the uninducible phenotype of the strain with the nutB-1 mutation (strain 10).

ii) When nutB-1 is either in cis or in trans to $nutC^+$ (merodiploid strains 13 and 14), wild-type levels of nutC activity are produced. Therefore, we determine that the nutB-1 mutation acts in trans.

c. i) The nutB-2 mutation is dominant to wild-type ($nutB^+$). This can be determined from strain 16, where the merodiploid has a copy of both nutB-2 and $nutB^+$, but it has the constitutive phenotype of the strain with the nutB-2 mutation (strain 15) instead of the inducible phenotype of the wild-type strain 8.

ii) Since the nutB-1 mutation acts in trans, we conclude the nutB-2 mutation also acts in trans, but this should be determined from the data given for strains 17 and 18. Since nutB-2 is a dominant mutation, we must

determine if the nutB-2 phenotype is expressed when the nutB-2 mutation is in cis or trans to nutC$^+$. When nutB-2 is either in cis or trans to nutC$^+$ the levels of nutC activity are constitutive (nutB-2 phenotype), which allows us to conclude that nutB-2 acts in trans.

d.

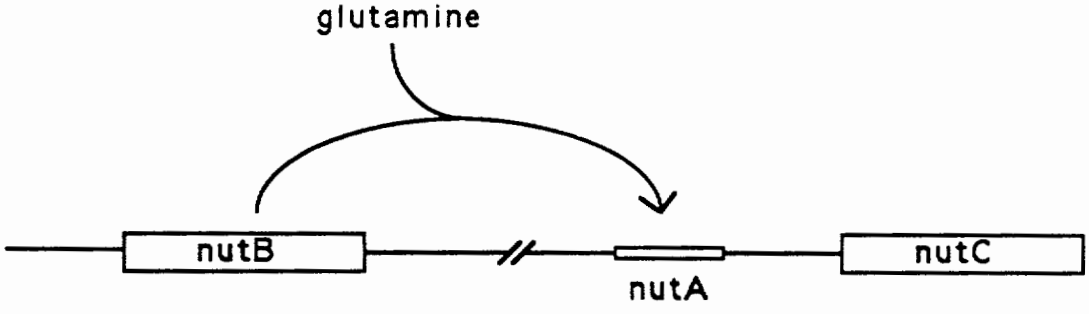
	Levels of nutC activity	
Strains	**- glutamine**	**+ glutamine**
i) nutB-2 nutA-1 nutC$^+$	0	0
ii) nutB-1 nutA$^+$ nutC$^+$/ F' nutB-2 nutA$^+$ nutC$^+$	400	400
iii) nutB-2 nutA-1 nutC$^+$/ F' nutB$^+$ nutA$^+$ nutC$^+$	200	200

Given the model in part e., the explanation for the levels of activity in the first strain is as follows: Even though the nutB-2 mutation is dominant and constitutive, the nutA-1 mutation is in the operator, and therefore the nutB-2 protein product will not be able to bind the nutA-1 operator and will not activate transcription in the presence or absence of glutamine.

The levels in the second strain are 400 in the presence or absence of glutamine because there is the dominant nutB-2 mutation is on the F' plasmid, and therefore will always activate transcription, even though the other copy of nutB is the nonfunctional nutB-1. The levels are 400 because there are two functional nutC genes with functional copies of the operator, nutA.

The levels of the third strain are 200 in the presence and absence of glutamine because the dominant constitutive nutB-2 mutation will always activate transcription of the nutC gene. The levels are 200 instead of 400 because only one copy of the nutC gene has a functional copy of the nutA operator where the activator (nutB) will bind.

e. The model, as diagrammed below, is as follows: The nutB gene encodes a protein that acts as an activator of the nutC gene. NutA, which is the operator site for the nutC gene, is where nutB binds to activate transcription of nutC. The nutB protein product will only activate nutC transcription when glutamine is present. Therefore, a recessive loss of function mutation in nutA (nutA-1) prevents the activator from binding even in the presence of glutamine, and no transcription of nutC can occur. The nutB-1 is a recessive loss of function mutation. This mutation is also responsible for a lack of nutC activity, because there is no functional nutB in a haploid with the mutation to activate transcription. NutB-2 is a dominant mutation that results in constitutive nutC activity, because the protein binds the nutA operator site even in the absence of glutamine.



2. **i)** The strain with the genotype smcA- smcB+ smcC+/F' smcA+ shows wild-type regulation; therefore, the smcA mutation is recessive. The merodiploid smcA- smcB+ smcC+/F' supE also shows wild-type regulation, indicating that the supE amber suppressor tRNA is able to suppress the smcA mutation. Therefore, the mutation must be a change within the smcA coding sequence to an amber stop codon which truncates the smcA protein.

ii) The strain with the genotype smcA+ smcB- smcC+/F' smcB+ shows wild-type regulation; therefore, the smcB mutation is recessive. The smcA+ smcB- smcC+ strain shows constitutive activity in the presence and absence of smactose. The smcB gene encodes a negative regulator of smactase activity.

iii) The strain with the genotype smcA+ smcB+ smcC-/F' smcC+ shows wild-type regulation; therefore, the smcC mutation is recessive. The smcA+ smcB+ smcC- strain is uninducible for smactase activity in either the presence or absence of smactose. Therefore, the smcC gene may encode a positive regulator of smactase activity. There is also the possibility that smcC encodes the smactase enzyme - see iv).

iv) A mutation in the structural gene for the smactase enzyme should cause a lack of activity (uninducible phenotype) alone and in all double mutant combinations. The smcA+ smcB+ smcC- strain is uninducible, but the smcA+ smcB- smcC- strain has constitutive smactase activity. Therefore, smcC can not encode the structural gene for the smactase enzyme. Both the smcA- smcB+ smcC+ strain and the smcA- smcB- smcC+ strain are uninducible for smactase activity. Therefore, smcA is more likely to encode the smactase enzyme.

These genetic data do not rule out the possibility that smcA is a positive regulator of smactase activity and that none of our genes encodes the smactase enzyme.

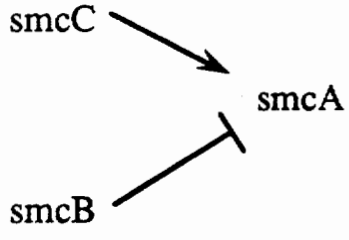
v) smcA probably encodes the smactase enzyme.
smcB is a (net) negative regulator of smactase.
smcC is a (net) positive regulator of smactase.

There are two simple models we can draw based on this information.

Model 1: linear pathway

In model 1, smcC positively regulates smcA by negatively regulating smcB, the repressor of smcA.

Model 2: smcC and smcB act directly on smcA

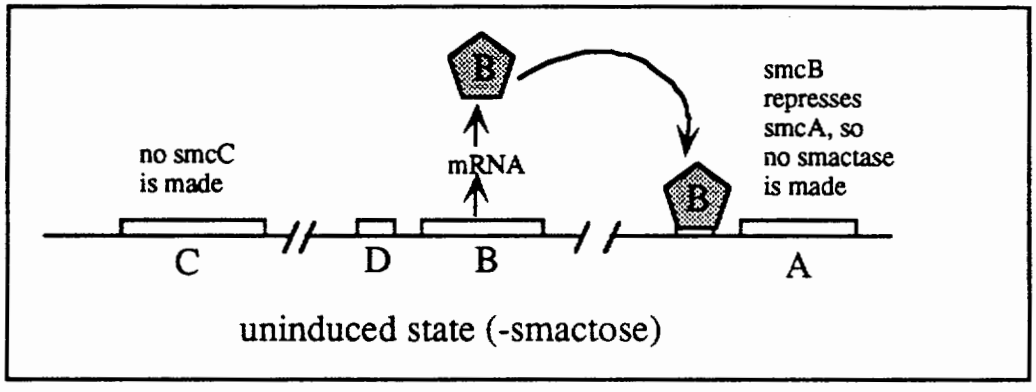


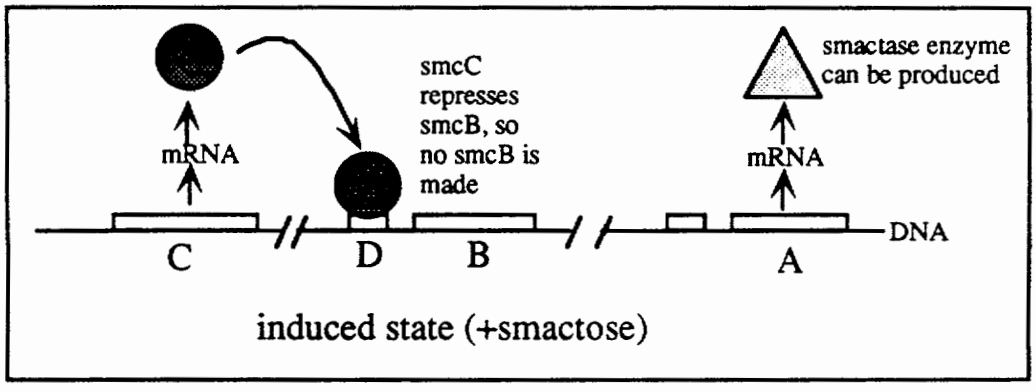
Note that genetic pathways describe regulatory circuits; they are not molecular or biochemical explanations!!

vi) Both the smcA+ smcB+ smcC+ smcD- strain and the smcA+ smcB+ smcC+ smcD-/ F' smcB+ smcD+ strain are uninducible for smactase activity, so the smcD mutation is dominant.

The smcA+ smcB- smcC+ smcD-/ F' smcB+ smcD+ strain shows wild-type inducible regulation of smactase activity, but the smcA+ smcB- smcC+ smcD+/ F' smcB+ smcD- strain is uninducible. The smcD- mutant phenotype is expressed when smcD- is in cis- to smcB+. Therefore, the smcD mutation is cis- acting with respect to smcB.

We know that (1) smcB negatively regulates smcA; and (2) smcD- is a cis-acting mutation with respect to smcB that results in the opposite phenotype to that associated with smcB- (smcB- is constitutive, smcD- is uninducible.) Therefore, smcD is likely to be a repressor binding site in the smcB operator. This fits well with model 1 above since smcC could be positively regulating smcA by inhibiting smcB. The following diagram may make all of this a little clearer:





Note that smcD can not be a "super-repressor" mutant of smcB since the smcA+ smcB- smcC+ smcD-/ F' smcB+ smcD+ shows wild-type regulation.

7.03 Problem Set 5

Due in class Monday, November 7

Note: Question #1 was purposely omitted.

2. You are interested in the genetics of onion color. The wild-type color of your favorite onion strain is red. White and yellow mutant onion strains have been isolated.

a) You obtain two true-breeding strains of mutant onions which are yellow. You cross these two strains and find all the F1 onions are red. You cross the F1 to each other and find the following phenotypes in the progeny.

yellow #1 x yellow #2

↓

F1: all red

F1 x F1

↓

F2: 563 red : 437 yellow

i) How can crossing two yellow strains produce red F1 progeny?
ii) Propose a model to explain the phenotypes of the F2 progeny.

b) The biochemical pathway required for making the red pigment is known. Strains mutant for any of the steps in this pathway produce yellow onions. It is

possible to bypass the requirement for a particular biochemical intermediate by supplying the missing compound in the soil. In addition to yellow strains #1 and #2, you also have yellow strains #3 and #4. You supplement the soil with compounds A-E and test whether the plants grown on it can produce red (wild-type) onions.

	media supplemented with:				
strain	A	B	C	D	E
1	+	+	-	+	-
2	-	+	-	-	-
3	+	+	-	-	-
4	+	+	-	+	+

Draw a diagram of the biochemical pathway leading to the synthesis of the red pigment. Include both the intermediates and the genes in your model.

c) You also obtain two true-breeding mutant strains that are white.

I) You cross the white#1 strain with a true-breeding yellow strain (you get the same results whether you use yellow strain #1, #2, #3, or #4) and find that all the F1s are white. You then cross the F1s with each other and find the following:

yellow x white#1

↓

F1: all white

F1 x F1

↓

F2: 750 white : 188 red : 62 yellow

How many genes are involved? What are the genotypes of the following strains: white#1, F2 white, F2 red, and F2 yellow? Explain how the genotypes give rise to the color phenotypes.

II) You cross the white#2 strain with a true-breeding red wild-type strain and find that all the F1s are red. You then cross the F1s with each other and find the following:

red x white#2

↓

F1: all red

F1 x F1

↓

F2: 564 red : 187 yellow : 249 white

How many genes are involved? What are the genotypes of the following strains: white#2, F2 red, F2 yellow, and F2 white? Explain how yellow F2 progeny arise and account for the ratios of the progeny classes.

3. For a UROP project you are investigating arginine biosynthesis in yeast. You are given three mutations in arg1: arg1-1, arg1-2, and arg1-3. Your professor tells you that arg1-3 is a deletion of the coding region. You also have in hand two alleles of arg2: arg2-1 and arg2-2.

The arg mutants are unable to grow on minimal media lacking arginine. You mutagenize a haploid arg1-1 strain with EMS and isolate four colonies capable of growing in the absence of arginine.

To investigate the nature of these four "revertants" you cross to a wild type haploid, sporulate the diploid, dissect the tetrads, and score them for ability to grow in the absence of arginine (arg^+).

"Revertant"	Tetrad Types	Number
A X arg^+	4 arg^+: 0 arg^-	30
	3 arg^+: 1 arg^-	115
	2 arg^+: 2 arg^-	25
B X arg^+	4 arg^+: 0 arg^-	142
	3 arg^+: 1 arg^-	0
	2 arg^+: 2 arg^-	0
C X arg^+	4 arg^+: 0 arg^-	16
	3 arg^+: 1 arg^-	90
	2 arg^+: 2 arg^-	18
D X arg^+	4 arg^+: 0 arg^-	26
	3 arg^+: 1 arg^-	101
	2 arg^+: 2 arg^-	23

(a) What does the tetrad analysis tell you about the nature of the "revertants"? Explain how the three tetrad types arise. Give two possible explanations for "revertant B".

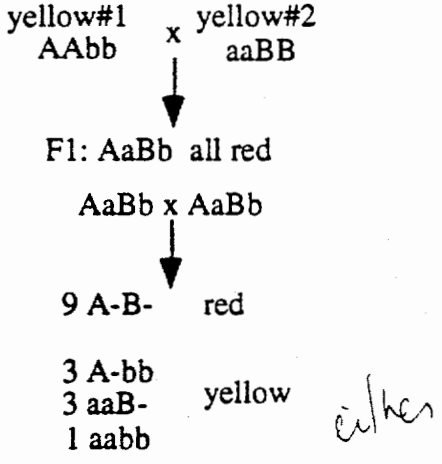
You wish to examine the "revertants" in the presence of your other arg1 alleles as well as in arg2 mutants and so do more matings with the original "revertants" and tetrad dissections. **The arg2 gene is tightly linked to arg1.**

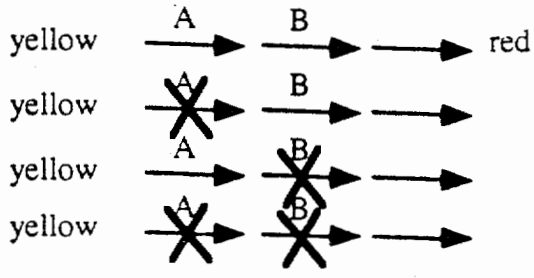
Revertant A X arg1-3	All tetrads are 2 arg^+: 2 arg^-	
(The same results are obtained in a cross to arg1-2)		
Revertant A X arg2-1	0 arg^+: 4 arg^-	21
	1 arg^+: 3 arg^-	90
	2 arg^+: 2 arg^-	23
(Similar ratios are obtained in a cross to arg2-2)		

7.03 Problem Set 5 Answers

2. a) i) Since all the F1 progeny are red, which is the wild-type color of the onion strain, we can conclude that the mutations in the yellow#1and yellow#2 strains are recessive and they complement each other. Therefore they must lie in different genes.

ii) The two phenotypes of the F2 progeny show a 9red:7yellow ratio, indicating that there are two unlinked genes involved. The red F2 must be of genotype A-B-; while the yellow F2 must consist of three genotypes: A-bb, aaB-, and aabb. One model that fits the data would be that both A and B are required for producing the red pigment. Onions mutant for either A or B or both A and B would be yellow.



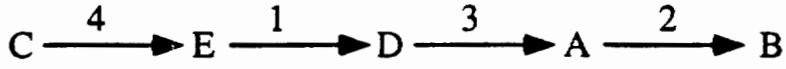


b) In a "complementation by feeding" experiment, an intermediate produced in later steps of a biochemical pathway would be capable of complementing

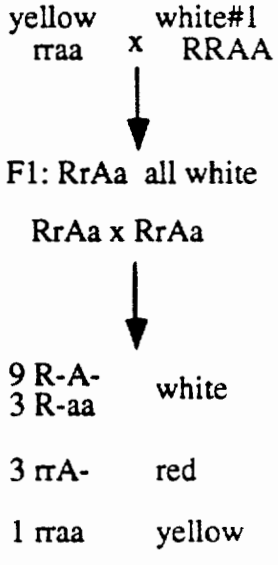
more mutations. Thus we can figure out that the order of intermediates produced along the pathway is C-E-D-A-B.

On the other hand, mutations that disrupt earlier steps of a biochemical pathway would be complemented by a greater number of intermediates. By this criterion, we can figure out that the order of gene action is 4-1-3-2.

Combining the order of intermediates with that of gene action, we can conclude that the biochemical pathway for producing the red pigment must be

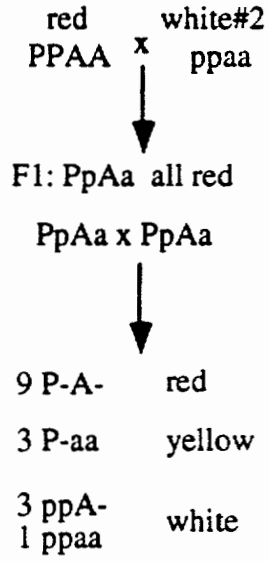


c) i) The 12:3:1 ratio observed in the F2 progeny suggests that there are two unlinked genes involved. One possible model is that a regulatory gene R is involved in determining the onion color. The dominant allele R prevents color expression in onions, and the recessive allele r permits color expression. R is epistatic to the genes involved in the pigment biochemical pathway. By this model, since red onions are observed in the F2 progeny, we know that white#1 must carry the wild-type "red" allele, and the wild-type "red" allele must be masked by dominant allele R, which makes the onion white. . So white#1 is RRAA (A is the wild-type allele of any of the genes 1-4 in the above pathway.) The yellow parental strain is rraa. All the F1 are RrAa. In the F2 progeny, onions carrying the R allele (R-A- + R-aa) would be white, rrA- would be red, and rraa would be yellow.



ii) The 9:3:4 ratio observed in the F2 progeny suggests that there are two unlinked genes involved. One can make sense of the data by introducing

another regulatory gene P. The recessive allele p prevents color expression and the dominant allele P permits color expression. In the homozygous condition, pp is epistatic to the genes in the pigment biochemical pathway. Since yellow onions appear in the F2 progeny, we know that white #2 strain must carry the yellow mutation.



Gene P could also encode an enzyme responsible for making the yellow pigment.

3. a) The classes of tetrad types observed for revertants A, C, and D indicate that the arg^+ colonies isolated from the arg1-1 strain contain extragenic suppressors in addition to the original arg1-1 mutation. We come to this conclusion because these revertant strains are able to segregate arg^- spores when crossed to an arg^+ strain, indicating that the arg1-1 mutation is still present. The arg^- phenotype was masked in the original revertant strains due to the presence of another mutation capable of suppressing the arg^- phenotype. During meiosis, the suppressor will become separated from the arg1-1 mutation at some frequency, and thus the arg^- phenotype revealed. You can see that the suppressor is unlinked from the tetrad ratios. The parental ditype tetrads have 4 arg^+spores, the nonparental ditypes have 2 arg^+ spores, and the tetratypes are 3 arg^+. Because PD:T:NPD is 1:4:1 the suppressor gene is unlinked from the arg1 gene.

The key observation concerning revertant B is that none of the tetrads produced by crossing it to an arg^+ strain contain any arg^- spores, ie. all of

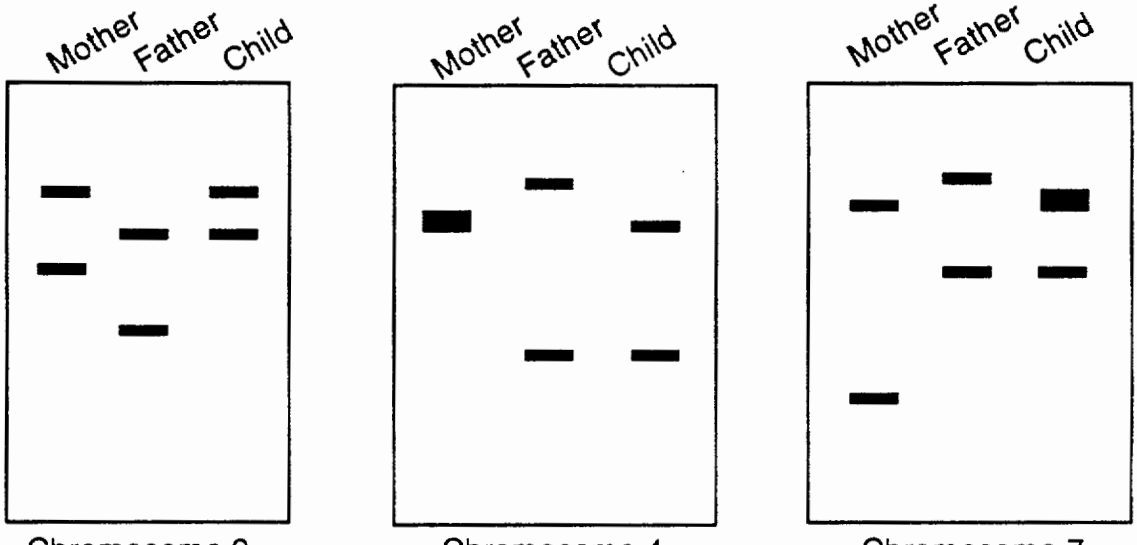
the tetrads are parental ditypes. This suggests that revertant B may be a genotypic revertant, where the arg1-1 mutation itself has been reverted. Another possibility is that the revertant strain contains a very tightly linked suppressor that is recombinationally inseparable from the arg1-1 mutation. This would most likely be produced by another mutation within the arg1 gene, an intragenic suppressor, but could also be caused by an extragenic suppressor that is located very close to the arg1 gene.

7.03 Problem Set 6

Due Friday, November 18

1. You are a new doctor at a prestigious hospital in Boston. A very sick newborn is brought to you with a large number of malformations, including cerebral, cardiovascular, and muscular defects. You set out to analyze the genetic basis of the disease.

Polymorphic markers linked to the centromeres have been identified on different chromosomes. The polymorphisms are assayed by Southern blot. Genomic DNA from each family member is cut with restriction enzymes and run on gels. The DNA in the gels is transferred to nitrocellulose filters. The blots are then probed with labeled DNA capable of identifying the different polymorphisms. The following patterns are observed with markers for chromosomes 2, 4, and 7:



Chromosome 2 marker | Chromosome 4 marker | Chromosome 7 marker

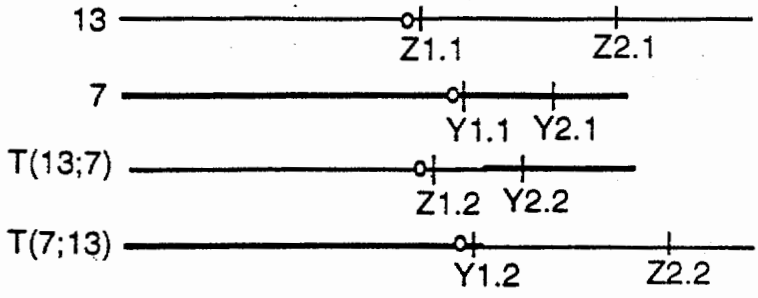
An equal amount of DNA was loaded in each lane. The band intensities, therefore, correspond to the copy number of the markers.

a. What type of chromosomal abnormality is observed in the child? Identify the parent in which the event must have occurred. Draw the aberrant meiosis and identify whether the change occurred in meiosis I or meiosis II.

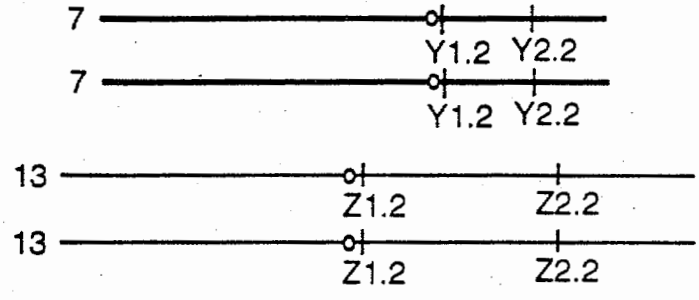
You then check medical databases from other hospitals for other cases of the disease. You find a case where a child from a father carrying the reciprocal translocation T(7;13) is affected, although both parents are normal. The mother's karyotype is normal.

Y1, Y2, Z1, and Z2 are polymorphic markers associated with chromosomes 7 and 13, respectively. Different variants of the markers are associated with the normal and translocated chromosomes.

Father's genotype



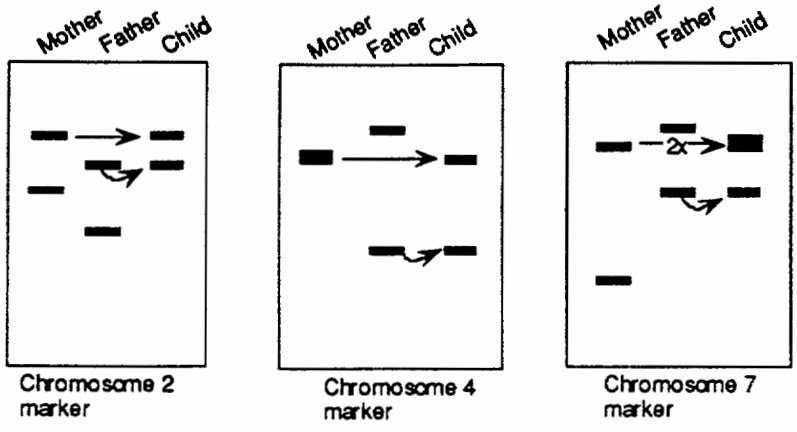
Mother's genotype



Note: Questions #2 + #3 were purposely omitted.

Problem Set 6 Answers

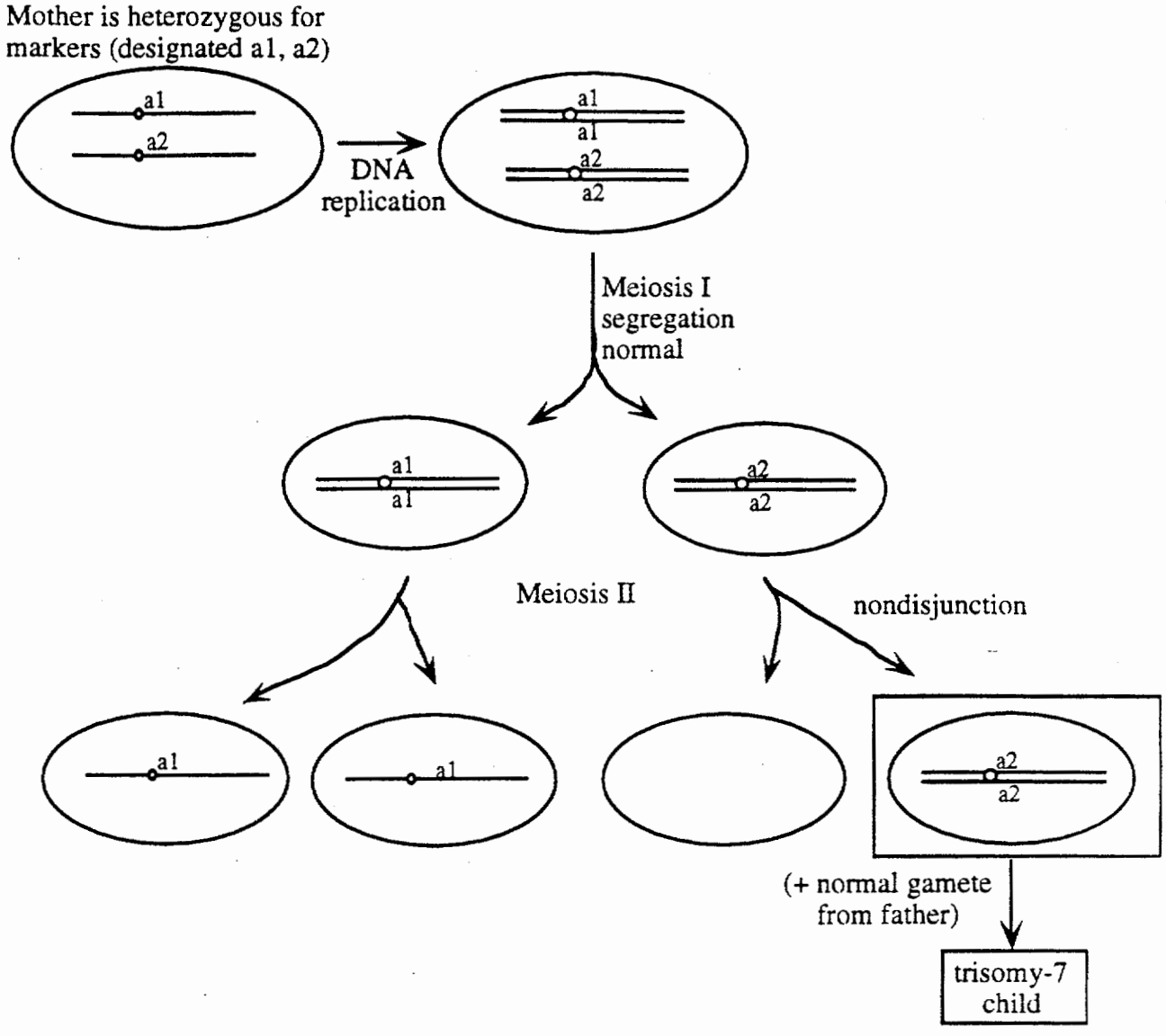
1. a. In the chromosome 2 and 4 blots, the child is heterozygous for the polymorphisms, having inherited different alleles of the polymorphism from each parent. An unusual pattern is observed for the chromosome 7 marker, however. The child has inherited one allele from the father as well as <u>two</u> copies of one of the mother's alleles.



Since the polymorphism is closely linked to the centromere, we infer that the child is trisomic for chromosome 7. Since the child has two copies of the mother's allele, the nondisjunction event must have occurred during meiosis in the mother. Since the child is carrying two copies of the same allele from the mother, nondisjunction must have occurred during meiosis II.

If nondisjunction had occurred during meiosis I, then the child would carry a copy of each of the mother's alleles (both a1 and a2), and three bands with different sizes would be visible in the child's lane of the Southern blot.

The following diagrams the aberrant meiosis:

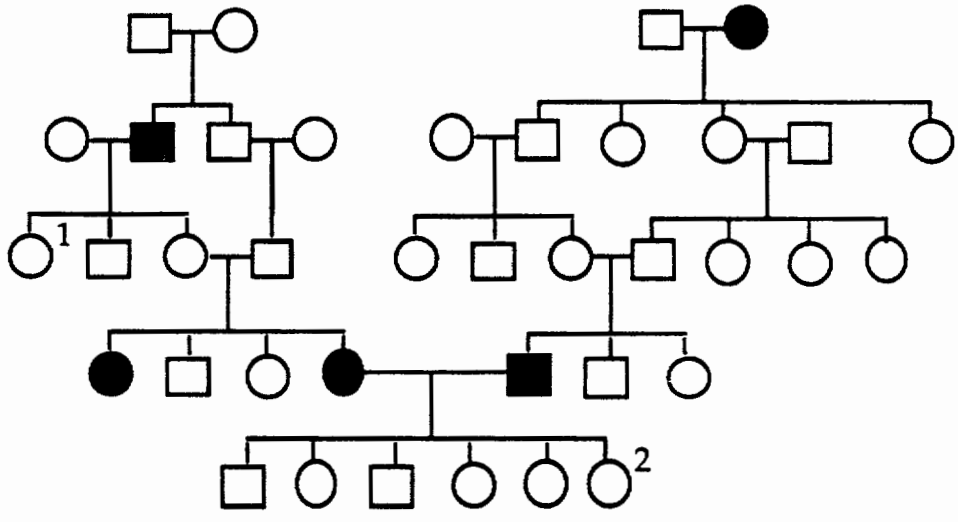

Mother is heterozygous for markers (designated a1, a2)
a1
a2
DNA replication
a1
a1
a2
a2
Meiosis I segregation normal
a1
a1
a2
a2
Meiosis II
nondisjunction
a1
a1
a2
a2
(+ normal gamete from father)
trisomy-7 child

Problem set 7

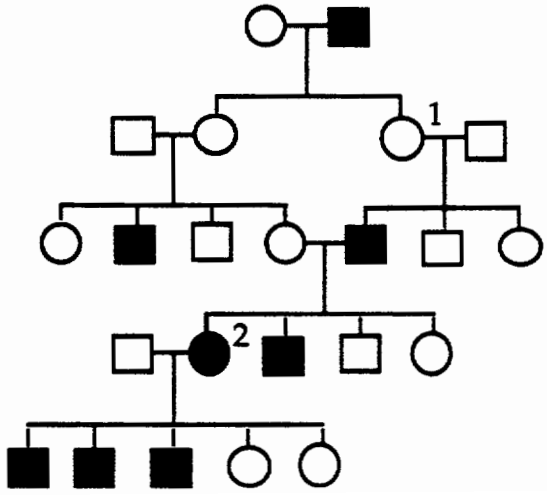
due in class December 9

1. a) In the following pedigrees, individuals affected with certain traits are indicated by the solid symbols. Squares represent males, and circles represent females. For each pedigree, deduce the most likely mode of inheritance of the trait and the probable genotype(s) of the marked individuals. Assume complete penetrance of the trait in all pedigrees.

i) Two families report to you with apparently the same disease. Fortunately they have intermarried, so you are able to investigate transmission within each family as well as in the children from this intermarriage.



ii)



b) The large pedigree on the following page exhibits segregation of an autosomal dominant form of the disease retinitis pigmentosa. This form of the disease has been shown to be linked to a DNA marker which has 4 alleles (1, 2, 3, and 4). Affected individuals are all heterozygous for the disease allele (D) and the disease is completely penetrant.

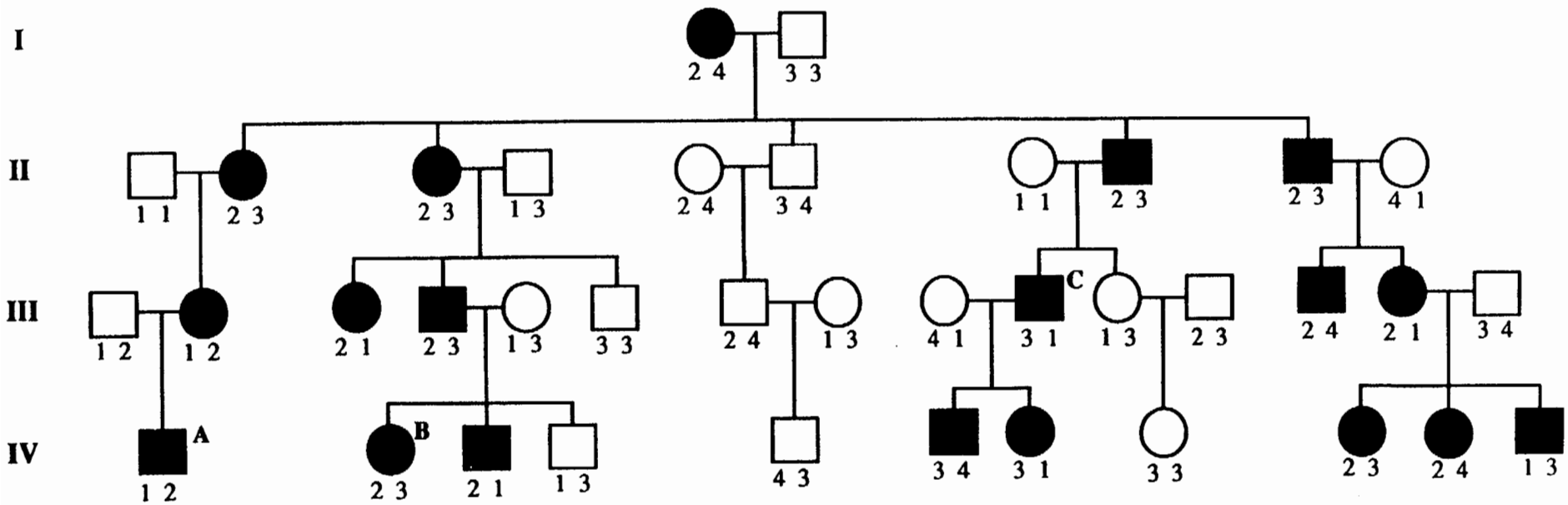
i) Assign the phase in the individuals marked A, B, and C. By phase we mean which alleles of the disease locus and which alleles of the marker locus are on the same homologs. In any of these 3 individuals is the phase ambiguous?

ii) Which allele of the marker appears to be segregating with the disease?

iii) Are there any unaffected individuals who carry the allele of the marker which segregates with the disease (from part ii)? How can you explain this?

iv) Indicate individuals in which a recombination event between the disease locus and the marker occurred in their gametes (i.e. not the individuals who inherited the recombinant chromosome). Individuals can be denoted by their generation (I, II, III, or IV) and their horizontal position counting from left to right. For example, individual B is denoted as IV-2 and individual C is denoted as III-10.

Note: Questions #2 + #3 were purposely omitted.


I
II
III
IV
2 4
3 3
1 1
2 3
2 3
1 3
2 4
3 4
1 1
2 3
2 3
4 1
1 2
1 2
2 1
2 3
1 3
3 3
2 4
1 3
4 1
3 1
C
1 3
2 3
2 4
2 1
3 4
1 2
A
2 3
B
2 1
1 3
4 3
3 4
3 1
3 3
2 3
2 4
1 3

Problem Set 7 Solutions

1. a) i) In each of the two separate families (not including the last generation), the trait appears to segregate in a simple autosomal recessive manner. That is, both males and females are affected at about the same frequency, the trait can skip generations, and the trait appears in consanguineous matings (parents are related). However, when affected individuals from each family have children with one another, none of the children are affected. If this trait were governed by a recessive allele of a single autosomal locus, then you would expect all the children from a mating between two affected individuals to be affected. Since this is not the case, it appears that the two mutations complement one another. In other words, there must be two genes (a and b) that affect this trait. Therefore, the genotype of individual 1 is +/a ; +/+, and the genotype of individual 2: +/a ; +/b

ii) In this pedigree, males are affected much more frequently than females, and it appears as if the trait is passed from mothers to sons. Therefore, one would conclude that this pedigree exhibits an X-linked recessive pattern of inheritance. In the fourth generation, there is both an affected male and female because both the mother and the father have a mutant allele. The affected son inherits his one affected X chromosome from his mother, while the affected daughter inherits one affected X chromosome from each parent. Since individual 1 has an affected son, is unaffected herself, and has an affected father, her genotype must be $X^{+}X^{d}$ (where d is the mutant allele). Since individual 2 is affected, her genotype must be $X^{d}X^{d}$.

mutant allele). Because this person is affected, his mother must be m/m and therefore he must have inherited an m allele from her. Unfortunately we know nothing about his father and he could have inherited either an m or a + allele from his father. (Note: An affected mother who has no affected children must be a heterozygote.) Finally, the genotype of individual 2 must be m/m because she has all affected children.

b) i) The phase of individual A is ambiguous because both his parents are heterozygous for the same markers. We know that his mothers must have the mutant allele on the same chromosome as marker allele 2 (see part 2), but we are uncertain whether he inherited the 2 allele from his mother or his father (the same applies to the 1 allele). Therefore, he could have inherited the mutant allele on a non-recombinant chromosome (with allele 2) or on a recombinant chromosome (with allele 1).

The phase of individual B and C are as follows (where D denotes the mutant allele, and the marker alleles are represented by 1,2, and 3):

B:

+	3
D	2

C:

+	1
D	3

Since individual B has the disease allele, and the 2 allele of the marker we know that he must have inherited both of these from his father, therefore D and 2 must lie on the same chromosome (he could not have inherited these from his mother, because she carries neither).

The same logic applies in the case of individual C.

ii) Allele 2 appears to be segregating with the disease, because allele 2 is inherited with the disease allele in almost all affected individuals.

iii) Several unaffected individuals carry the marker allele 2 (for example, individuals II-5 and II-12). This is possible because the marker is not the <u>cause</u> of the disease. Allele 2 just happens to be linked to the disease allele most often in this particular

pedigree, and another family may show linkage between a different allele of the marker and the disease.

iv) A recombination event occurred in individuals II-8 and III-14. Individual IV-11 inherited the recombinant chromosome (where the disease is linked to marker allele 3 instead of allele 2) from III-14. Individual III-10 inherited recombinant chromosome (where the disease is linked to marker allele 1 instead of allele 2) from II-8 and subsequently passed this chromosome on to IV-6 and IV-7.

Population Genetics Sample Problem

a) Consider a rare recessive disorder. For population I which is in Hardy-Weinberg equilibrium, the frequency of carriers is 10^{-3}. What is the expected frequency of affected individuals?

b) What is the probability that a child of a marriage between second-cousins from population I will be affected?

c) If the fitness of affected individuals (homozygotes) is 0.5 and the carrier (heterozygote) has the same fitness as homozygous normal individuals, what is the mutation frequency?

d) If the mutation frequency is very low but there is a selective advantage to being a carrier of the trait, what is the estimated heterozygote advantage (h)?

e) Consider a population II for which the frequency of affected individuals is 10^{-4} and the frequency of carriers is 2×10^{-2}. Is population II in Hardy-Weinberg equilibrium?

f) As a result of migration, an island is peopled by equal numbers of individuals from populations I and II. On this island, what will the frequency of the disorder be if there is random mating between individuals of the two populations? What would the frequency be if there were no intermating between the two populations?

Answers

a) $10^{-3} = f(A/a) \approx 2q$, $q = 5 \times 10^{-4}$

$f(a/a) = q^2 = \boxed{2.5 \times 10^{-7}}$

b) F for second cousins is 1/64

The frequency of affected offspring will be: $F \cdot q = \boxed{7.8 \times 10^{-6}}$

c) With no selection on heterozygote. $\hat{q} = \sqrt{\mu/s}$

$5 \times 10^{-4} = \sqrt{\mu/s}$ $\quad \boxed{\mu = 1.25 \times 10^{-7}}$

d) For balanced polymorphism $sq = h$

$h = (0.5)\ 5 \times 10^{-4} = \boxed{2.5 \times 10^{-4}}$

e) $q = 10^{-2}$, $q^2 = 10^{-4}$ $\quad \boxed{\text{Yes}}$

f) $q_{\text{pop I}} = 5 \times 10^{-4}$

$q_{\text{pop II}} = 10^{-2}$

i) For random mating q for the combined population is 5.5×10^{-3}

$\boxed{q^2 = 3 \times 10^{-5}}$

ii) With no mixing the average q^2 for the two subpopulations is

$\text{ave } q^2 = \dfrac{2.5 \times 10^{-7} + 10^{-4}}{2} \approx \boxed{5 \times 10^{-5}}$