

# Problem Sets

Fall 1998

## 7.03 Problem Set 1

Due before 5 PM on Thursday, September 24

Hand in answers in recitation section or in the box outside the class

**1. (a)** You and your lab partner have isolated twenty His<sup>-</sup> yeast mutants. Mutants 1 – 10 are of mating type  $\alpha$  and mutants 11 – 20 are of mating type **a**. Your partner has performed complementation tests by mating each  $\alpha$  strain to each **a** strain. The table below shows the results where a "+" at the intersection of the two parental strains indicates that the diploid can grow without histidine added to the medium, and a "-" indicates that the diploid can't grow without histidine. Unfortunately, after all of the data are collected and the plates are thrown away, your partner spills coffee on the table wiping out much of the data. From the data that remains see if you can reconstruct the full table.

		strains of mating type $\alpha$										
		wild- type	1	2	3	4	5	6	7	8	9	10
strains of mating type <b>a</b>	11	+	-	-				-	+			
	12		-	-		+	-			-		
	13			-					-		-	
	14			+	-			+				+
	15		-				-				+	
	16			+		-				-		
	17		+			+			+		+	+
	18				-		+		+	-	-	
	19				+	-			+			+
	20	+					-					+

**(b)** Determine how many different genes are represented by your collection of His<sup>-</sup> mutants. For your answer you should indicate which mutations fall into the same complementation group. Any remaining ambiguities in the assignment of mutations to complementation groups should be explained.

**2. (a)** In a cross between a male mouse from a true-breeding black strain and a female from a true-breeding tan strain all of the  $F_1$  progeny are gray.

Based on the information that you have at this stage, is it possible that a single gene determines the differences in coat color among the two parental strains and progeny? If so, what coat colors should appear in the  $F_2$  generation and at what frequencies?

**(b)** In fact, when  $F_1$  mice are crossed among themselves, the following  $F_2$  progeny are produced: 30 gray mice, 10 black mice, 8 tan mice, and 2 dark brown mice.

Propose a genetic model to account for the existence of the dark brown  $F_2$  mice. For your answer give the genotypes of the parental mice, the  $F_1$  mice and each class of  $F_2$  mice.

**(c)** Use the Chi-square test to show that the observed frequencies fit with the expected frequencies based on your model. For your answer, give the observed and expected phenotypic ratios, the degrees of freedom, your calculated value for  $\chi^2$ , and a rough estimate of the p value.


**(d)** Returning to the original true-breeding parental strains, a male from the tan strain is crossed to a female from the black strain. As expected all of the female  $F_1$  mice are gray, but to your surprise all of the male  $F_1$  mice are black.


Propose a genetic model to account for this new data. Give the genotypes of the male and female parental mice and the male and female  $F_1$  mice. Finally, predict the types of mice that will appear in the  $F_2$  generation. Specify the coat colors, sex, and expected frequency of each class.


**(e)** Given the genetic model that you have developed in part <sup>d</sup>(d), return to the crosses described in parts (a) and (b) and determine what the  $F_3$  generation will look like. Again, specify the coat colors, sex, and expected frequency of each class.

Use model developed in part (d) to derive expected phenotypes & frequencies for  $F_3$  from cross in 2a

**3.** Each of the following pedigrees carries two different recessive traits indicated as follows:

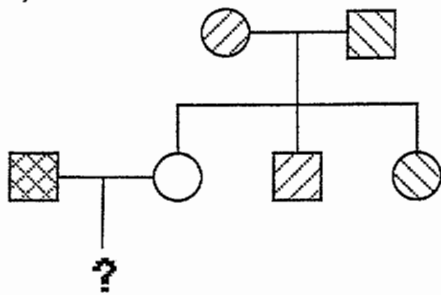
 = individual with trait 1

 = individual with trait 2

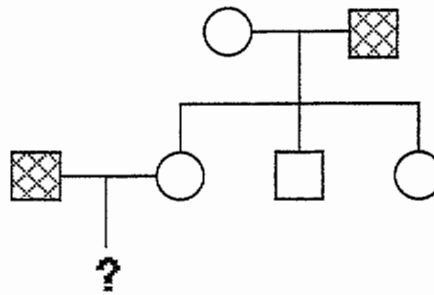
 = individual with both trait 1 and trait 2

The loci for the two traits are linked on the same autosome and lie 20 cM apart. For each pedigree, calculate the probabilities that the individual indicated by ? will have only trait 1, only trait 2, or both traits.

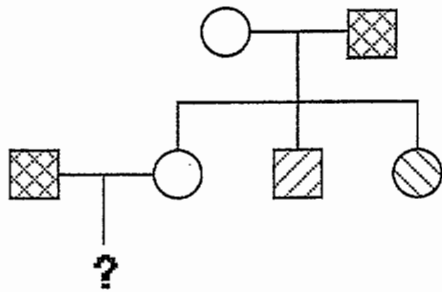
(a)



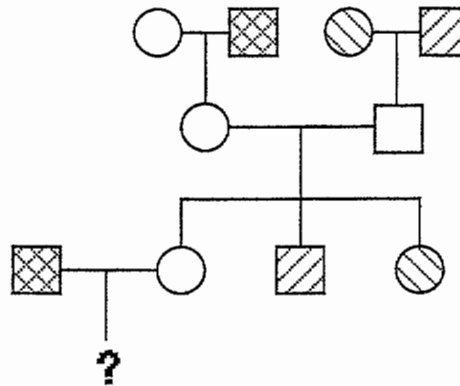
(b)



(c)



(d)



1a.

		strains of mating type $\alpha$										
		wild-type	1	2	3	4	5	6	7	8	9	10
strains of mating type a	11	+	-	-	+	+	-	-	+	-	+	+
	12	+	-	-	+	+	-	-	+	-	+	+
	13	-	-	-	-	-	-	-	-	-	-	-
	14	+	+	+	-	+	+	+	+	-	-	+
	15	+	-	-	+	+	-	+	+	-	+	+
	16	+	+	+	+	-	+	+	+	-	+	+
	17	+	+	+	+	+	+	+	+	-	+	+
	18	+	+	+	-	+	+	+	+	-	-	+
	19	+	+	+	+	-	+	+	+	-	+	+
	20	+	-	-	+	+	-	-	+	-	+	+

1b. There are at least five different complementation groups (and hence at least 5 genes represented). The first group contains mutants 1,2,5,6,11,12,15, and 20. The second group contains 3, 9, 14, and 18. The third group contains 4, 16, and 19. The fourth group contains only 17. A fifth group contains 7. 10 may belong to this last group, or perhaps its own group; it cannot be determined for certain from this data. Lastly mutants 13 and 8 are dominant mutations and cannot be assigned to any complementation groups based on the data presented in this table.

2a. If we were to assume that one gene determined the coat color, then this would be an example of co-dominance. In this case the  $F_2$  would have phenotypes at the following ratio: 1Black:2Gray:1Tan.

2b. Since this ratio is approximately a 9:3:3:1 ratio, a reasonable hypothesis would be a two gene model. In this case the black parental mouse would have a genotype of  $B/B$   $t/t$  and the tan parental mouse would have a genotype of  $b/b$   $T/T$ . All of the  $F_1$ 's would have the genotype  $B/b$   $T/t$ . The black  $F_2$ 's will have two recessive tan traits ( $t/t$ ) and at least one dominant black trait ( $B/?$ ). The Tan  $F_2$ 's will have two recessive black traits ( $b/b$ ) and at least one dominant tan trait ( $T/?$ ). If the  $F_2$  is dark brown it will have two pairs of recessive traits ( $b/b$   $t/t$ ). All other genotypes ( $B/?$   $T/?$ ) will have a gray phenotype.

2c.

$$X^2 = \Sigma(O-E)^2/E$$

Phenotypic class	Observed	Expected	(O-E) <sup>2</sup> /E
Gray	30	9/16*50 = 28.125	0.125
Black	10	3/16*50 = 9.375	0.041
Tan	8	3/16*50 = 9.375	0.202
Dark Brown	2	1/16*50 = 3.125	0.405

$$X^2 = 0.773$$

4 phenotypic classes – 1 = 3 degrees of freedom. Referring to page 138 in your textbook, a  $X^2$  value of 0.773 with 3 degrees of freedom corresponds to a p value between 0.5 and 0.9. This p supports the likelihood that this trait is controlled by two genes that independently segregate.

2d.

This data fits a model in which there are two genes influencing coat color, with the “tan” gene located on the X-chromosome and the “black” gene located on an autosome. In this case the parental female would be B/B  $X^T/X^t$  and the parental male would be b/b  $X^T/Y$ . The  $F_1$  males would have the genotype B/b  $X^t/Y$  (black), and  $F_1$  females would have the genotype B/b  $X^T/X^t$  (gray).

In the  $F_2$  generation there would be the same four phenotypic classes as in problem 2b, but at different ratios. The ratios would be 3 gray: 3 black: 1 tan: 1 dark brown, with an equal number of males and females in each class.

2e. The calculation for the phenotypic ratios for  $F_3$  is simplified if you consider the frequency of the two genes separately. Considering the trait for black coat color first, the alleles segregate identically in both males and females in the  $F_3$  at a ratio of 1BB:2Bb:1bb. The trait for tan coat color will appear 3 $X^T$ Y:1 $X^t$ Y in males and 3 $X^T$  $X^T$ :4 $X^T$  $X^t$ :1 $X^t$  $X^t$  for females.

To determine the frequencies of the complete genotypes for  $F_3$ , simply multiply the probabilities for each trait together. The results of this multiplication would be:

Male			Female		
Genotype	Probability	Phenotype	Genotype	Probability	Phenotype
BB $X^T$ Y	3/16	Gray	BB $X^T$ $X^T$	3/32	Gray
BB $X^t$ Y	1/16	Black	BB $X^T$ $X^t$	4/32	Gray
Bb $X^T$ Y	6/16	Gray	BB $X^t$ $X^t$	1/32	Black
Bb $X^t$ Y	2/16	Black	Bb $X^T$ $X^T$	6/32	Gray
bb $X^T$ Y	3/16	Tan	Bb $X^T$ $X^t$	8/32	Gray
bb $X^t$ Y	1/16	D. Brown	Bb $X^t$ $X^t$	2/32	Black
			bb $X^T$ $X^T$	3/32	Tan
			bb $X^T$ $X^t$	4/32	Tan
			bb $X^t$ $X^t$	1/32	D. Brown

The  $F_3$  phenotypic ratio for males will be 9Gray:3Black:3Tan:1DB. For females the  $F_3$  phenotypic ratio will be 21Gray:7Tan:3Black:1DB.

3a. 40% will express trait 1  
 40% will express trait 2  
 10% will express neither  
 10% will express both

3b. 40% will express neither  
 40% will express both  
 10% will express trait 1  
 10% will express trait 2

3c. Same as 3b.

3d. The genotypes of the  $F_1$  generation can be determined to be  $12/++$  for the female and  $+2/1+$  for the male. This means that there are several possible genotypes for the mother of the unknown child. There can basically be 5 genotypes that will allow the mother of the unknown to be unaffected by either trait, each occurring at the following rates:

Genotype	Probability	Phenotype
$++/++$	$4/54$	Normal
$++/+2$	$17/54$	Normal
$++/1+$	$17/54$	Normal
$++/12$	$8/54$	Normal
$+2/1+$	$8/54$	Normal

Now each of these possible genotypes have to be crossed with the doubly homozygous male. The easiest way to do this is consider each case separately and then weight the ratios with the probability of each genotype for the mother

The final answer will be

$25/54$  WT                       $2/27$  Both                       $25/108$  Trait 1                       $25/108$  Trait 2

(It doesn't add up to 100 because percentages are approximate)

## Problem Set 1 Detailed Solutions

I think the answers to Problem 1 are relatively complete. If you need more help understanding this problem, please come to Office Hours and we'll discuss it...

#2

- a. (2 pts) If one gene determines coat color, then the co-dominance (or incomplete dominance, depending on whether you consider gray a mixture of tan and black or a separate color) explains the appearance of three phenotypes.
- b. (2 pts) The answer key seems clear about this one...you need at least two genes to explain 4 phenotypes. (only three phenotypes are possible in a one gene model)
- c. (2 pts) In the Chi-square test, in order to calculate the expected values, you must take the observed total number of progeny and divide into the appropriate frequencies for each class...
  - e.g. Since there are 50 total progeny, the expected values are 28.125, 9.375, 9.375, 3.125. It's OK that these are fractions since this is a statistical test.
- d. (2 pts) The fact that the ratio of phenotypes between males and females is different alerts you to a sex-linked trait. You must, however, still account for the 4 phenotypes in part b...thus two genes must be involved. To be consistent with the data already presented in (b), only one gene can be sex-linked. The frequencies can be determined by dealing with each gene independently since neither is linked.

(Black female)                      (tan male)  
 $B/B X^t/X^t$     x     $b/b X^T/Y$

F1     $B/b X^t/X^T$     x     $B/b X^t/Y$   
       (gray females)    (black males)

For the autosomal gene in the F2:

BB 1/4  
 Bb 2/4  
 bb 1/4

For the sex-linked gene in the F2:

females:	males:
$X^t/X^T$ 1/2	$X^T/Y$ 1/2
$X^t/X^t$ 1/2	$X^t/Y$ 1/2

Females:

$B/B X^t/X^T$  (1/4)(1/2) = 1/8 gray  
 $B/B X^t/X^t$  (1/4)(1/2) = 1/8 black  
 $B/b X^t/X^T$  (1/2)(1/2) = 2/8 gray  
 $B/b X^t/X^t$  (1/2)(1/2) = 2/8 black  
 $B/b X^T/X^T$  (1/4)(1/2) = 1/8 tan  
 $B/b X^t/X^t$  (1/4)(1/2) = 1/8 dark brown

Males:

$B/B X^T/Y$  (1/4)(1/2) = 1/8 gray  
 $B/B X^t/Y$  (1/4)(1/2) = 1/8 black  
 $B/b X^T/Y$  (1/2)(1/2) = 2/8 gray  
 $B/b X^t/Y$  (1/2)(1/2) = 2/8 black  
 $b/b X^T/Y$  (1/4)(1/2) = 1/8 tan  
 $b/b X^t/Y$  (1/4)(1/2) = 1/8 dk brn



e. For this section of the problem the cross is as follows:

(Tan female)  $b/b X^T/X^T$  x (black male)  $B/B X^t/Y$  \*note that this is the reciprocal cross from d.

F1  $B/b X^t/X^T$  x  $B/b X^T/Y$

F2	$B/B X^t/X^T$	1/8	$B/B X^t/Y$	1/8
	$B/B X^T/X^T$	1/8	$B/B X^T/Y$	1/8
	$B/b X^t/X^T$	2/8	$B/b X^t/Y$	2/8
	$B/b X^T/X^T$	2/8	$B/b X^T/Y$	2/8
	$b/b X^t/X^T$	1/8	$b/b X^t/Y$	1/8
	$b/b X^T/X^T$	1/8	$b/b X^T/Y$	1/8

F3 You can deal with each locus independently:

For the "B" locus: there is an equal number of each allele in the population. Since the F2 will be mating at random, you can follow the probability of an F3 inheriting a certain allele, instead of following each of the possible crosses.

For instance out of 8 possible alleles:

$B/B$  There is a 1/4 probability that a parent of this genotype will donate an allele to the progeny, and then if this parent does donate an allele, a 100% probability that the allele will be B. (1/4)

$B/b$  There is a 1/2 probability that a parent of this genotype will donate an allele to the progeny, and then if this parent does donate an allele, a 1/2 probability that the allele donated will be B. (1/4)

There is a 1/2 probability that a parent of this genotype will donate an allele to the progeny, and then if this parent does donate an allele, a 1/2 probability that the allele donated will be b. (1/4)

$b/b$  There is a 1/4 probability that a parent of this genotype will donate an allele to the progeny, and then if this parent does donate an allele, a 100% probability that the allele will be b. (1/4)

Thus, both alleles can be donated with equal frequency  $(1/4 + 1/4) = 1/2$

So the ratios for this locus in the F3 are:

So the ratios for this locus in the F3 are:

$$\begin{array}{l} B/B (1/2)(1/2) = 1/4 \\ B/b (1/2)(1/2) + \\ b/B (1/2)(1/2) \quad \quad \quad \} = 2/4 \\ b/b (1/2)(1/2) = 1/4 \end{array}$$

At the other locus, the same principle can be applied:

$$F_2: \begin{array}{ll} X^t X^T 1/2 & X^T Y 1/2 \\ X^T X^T 1/2 & X^t Y 1/2 \end{array}$$

$$F_3: \begin{array}{l} \text{♀: } X^t X^t \\ \quad \uparrow \quad \quad \swarrow \\ \text{from mother} \quad \text{from father} \\ (1/2)(1/2) \times (1/2) = 1/8 \end{array}$$

$$\begin{array}{l} \text{♂: } X^t Y \\ (1/2)(1/2) = 1/4 \end{array}$$

$$\begin{array}{l} X^t X^T \\ \quad \uparrow \quad \quad \swarrow \\ \text{from mother} \quad \text{from father} \\ (1/2)(1/2) \times (1/2) = 1/8 \end{array}$$

$$\begin{array}{l} X^T Y \\ (1/2)(1/2) + 1/2 = 3/4 \end{array}$$

$$\begin{array}{l} X^T X^t \\ \quad \uparrow \quad \quad \swarrow \\ \text{from mother} \quad \text{from father} \\ [(1/2)(1/2) + (1/2)] \times (1/2) = 3/8 \end{array}$$

if the heterozygote is the mother

if the homozygote is the mother

$$\begin{array}{l} X^T X^T \\ \quad \uparrow \quad \quad \swarrow \\ \text{from mother} \quad \text{from father} \\ [(1/2)(1/2) + (1/2)] \times (1/2) = 3/8 \end{array}$$

Now... we can combine the probabilities from the two loci...

$$\begin{array}{l}
 \rightarrow x^t x^t \quad (1/4)(1/8) = 1/32 \\
 BB \rightarrow x^t x^T \quad (1/4)(1/8) = 1/32 \\
 \downarrow x^T x^t \quad (1/4)(3/8) = 3/32 \\
 \downarrow x^T x^T \quad (1/4)(3/8) = 3/32
 \end{array}
 \left. \vphantom{\begin{array}{l} \rightarrow x^t x^t \\ \rightarrow x^t x^T \\ \downarrow x^T x^t \\ \downarrow x^T x^T \end{array}} \right\} 4/32$$

$$\begin{array}{l}
 \rightarrow x^t x^t \quad (1/2)(1/8) = 1/16 = 2/32 \\
 Bb \rightarrow x^t x^T \quad (1/2)(1/8) = 1/16 = 2/32 \\
 \downarrow x^T x^t \quad (1/2)(3/8) = 3/16 = 6/32 \\
 \downarrow x^T x^T \quad (1/2)(3/8) = 3/16 = 6/32
 \end{array}$$

$$\begin{array}{l}
 \rightarrow x^t x^t \quad (1/4)(1/8) = 1/32 \\
 bb \rightarrow x^t x^T \quad (1/4)(1/8) = 1/32 \\
 \downarrow x^T x^t \quad (1/4)(3/8) = 3/32 \\
 \downarrow x^T x^T \quad (1/4)(3/8) = 3/32
 \end{array}
 \left. \vphantom{\begin{array}{l} \rightarrow x^t x^t \\ \rightarrow x^t x^T \\ \downarrow x^T x^t \\ \downarrow x^T x^T \end{array}} \right\} 4/32$$

$$\begin{array}{l}
 \rightarrow x^t Y \quad (1/4)(1/4) = 1/16 \\
 BB \rightarrow x^T Y \quad (1/4)(3/4) = 3/16
 \end{array}$$

$$\begin{array}{l}
 \rightarrow x^t Y \quad (1/2)(1/4) = 1/8 = 2/16 \\
 Bb \rightarrow x^T Y \quad (1/2)(3/4) = 3/8 = 6/16
 \end{array}$$

$$\begin{array}{l}
 \rightarrow x^t Y \quad (1/4)(1/4) = 1/16 \\
 bb \rightarrow x^T Y \quad (1/4)(3/4) = 3/16
 \end{array}$$

3.a. (2 pts) Given that the parents genotypes are 12/12 and 1+ / +2,

For trait 1: the only possible genotype is 12/1+. This genotype will result if:

- a. there is no recombination in the gamete of the mother (.8) AND
- b. she donates the 1+ gamete (.5)

Thus, the probability that the child will express Trait 1 is  $(.8)(.5) = .4$

For trait 2: the only possible genotype is 12/+2. This genotype will result if:

- d. there is no recombination in the gamete of the mother (.8) AND
- e. she donates the +2 gamete (.5)

Thus, the probability that the child will express Trait 2 is  $(.8)(.5) = .4$

For both: the only possible genotype is 12/12. This genotype will result if:

- a. there IS recombination in the mother (.2) AND
- b. she donates the 12 gamete (.5)

Thus, the probability that the child will express both traits is  $(.2)(.5) = .1$

For neither: the only possible genotype is 12/++. This genotype will result if:

- d. there IS recombination in the mother (.2) AND
- e. she donates the ++ gamete (.5)

Thus, the probability that the child will express neither trait is  $(.2)(.5) = .1$

b. (2 pts) Given that the parents genotypes are 12/12 and ++/12,

For trait 1: the only possible genotype is 12/1+. This genotype will result if:

- a. there IS recombination in the mother (.2) AND
- b. she donates the 1+ gamete (.5)

Thus, the probability that the child will express Trait 1 is  $(.2)(.5) = .1$

For trait 2: the only possible genotype is 12/+2

- a. there IS recombination in the mother (.2) AND
- b. she donates the +2 gamete (.5)

Thus, the probability that the child will express Trait 2 is  $(.2)(.5) = .1$

For both: the only possible genotype is 12/12. This genotype will result if:

- a. there is no recombination in the gamete of the mother (.8) AND
- b. she donates the 12 gamete (.5)

Thus, the probability that the child will express both traits is  $(.8)(.5) = .4$

For neither: the only possible genotype is 12/++. This genotype will result if:

- a. there is no recombination in the gamete of the mother (.8) AND
- b. she donates the ++gamete (.5)

Thus, the probability that the child will express Trait 2 is  $(.8)(.5) = .4$

c. (2 pts) Given that the parents' genotypes are 12/12 and 12/++... the answer is the same as for b.

- d. (4 pts) Given that the mother's genotype can not be definitively determined, this part must be done in two steps.

First, identify the possible genotypes of the mother and determine their relative frequencies. There are five possible genotypes for the mother:

$++/++$   $+2/1+$   $+2/++$   $1+/++$   $+/12$

To determine the relative frequencies of each genotype, you can apply a version of a Punnett's square, with her father's possible gametes and the probability of each gamete being donated to the child along one side of the square, and the mother's on the other side. The probability for each gamete is determined as in parts a-c

	her father's gametes			
	1+ (.4)	+2 (.4)	++ (.1)	12 (.1)
++ (.4)	1+/++ (.4)(.4)	+2/++ (.4)(.4)	++/++ (.1)(.4)	++/12 (.1)(.4)
12 (.4)			++/12 (.1)(.4)	
1+ (.1)		+2/1+ (.4)(.1)	++/1+ (.1)(.1)	
+2 (.1)	1+/+2 (.4)(.1)		++/+2 (.1)(.1)	

I have only filled in the relevant squares, the genotypes that are possible for the mother, given that we know she is unaffected by either trait. We know the mother cannot be any of the other genotypes; therefore, to determine the RELATIVE frequencies of the five possible genotypes we must normalize the values in the table. Earlier in the class we defined probability as the number of outcomes that satisfy our conditions divided by the total number of outcomes. In this case, the number of outcomes is the sum of the probabilities that one genotype occurs divided by the sum of the probabilities in the filled in squares (which becomes our value for the number of possible outcomes).

For  $++/++$   $0.04/0.54$

For  $++/1+$   $(0.16 + 0.01)/0.54$

For  $++/+2$   $(0.16 + 0.01)/0.54$

For  $++/12$   $(0.04 + .04)/0.54$

For  $+2/1+$   $(0.04 + 0.04)/0.54$

**For Trait 1:**

Three of the genotypes listed above can yield gametes, which, when combined with the father's 12 gamete will result in a child with only trait 1. ( $++/1+$ ,  $++/12$ ,  $+2/1+$ ).

The probability that the child will have only trait 1 can be calculated as follows:

$$(1/2)(.17/.54)^{++/1+} + (1/2)(.08/.54)^{++/12}(.2) + (1/2)(.08/.54)^{+2/1+}(.8) = .23$$

↑ the prob. that the mother will donate the 1+ gamete  
 ↑ the prob. that she is this genotype  
 ↑ the prob. that the mother will donate the 1+ gamete  
 ↑ the prob. that she is this genotype  
 ↑ the prob. that there is no recombination.

For Trait 2:

Three of the genotypes listed above can yield gametes, which, when combined with the father's 12 gamete will result in a child with only trait 2. ( $++/+2$ ,  $++/12$ ,  $+2/1+$ )

This value is the same as for Trait 1, calculated in an identical manner. = .23

For both traits:

Two of the genotypes listed above can yield gametes, which, when combined with the father's 12 gamete will result in a child with both traits. ( $++/12$ ,  $+2/1+$ )

The probability that the child will have both traits can be calculated as follows:

$$(1/2)(.08/.54)^{++/12}(.8) + (1/2)(.08/.54)^{+2/1+}(.2) = .07$$

For neither trait:

All of the genotypes can result in a child expressing neither of the two traits.

This probability can be calculated as follows:

$$.04/.54 + (1/2)(.17/.54) + (1/2)(.17/.54) + (1/2)(.08/.54)(.2) + (1/2)(.08/.54)(.8) = .46$$

$$\frac{++}{++}$$

$$\frac{++}{1+}$$

$$\frac{++}{+2}$$

$$\frac{1+}{+2}$$

$$\frac{++}{12}$$

## 7.03 Problem Set 2

Due before 5 PM on Thursday, October 1

Hand in answers in recitation section or in the box outside the class

- 1.** Two different true-breeding *Drosophila* lines are crossed and F1 females from this cross are then crossed to males from a line that is homozygous for four different recessive traits. A total of 1000 progeny from these crosses are then evaluated for each of the four traits. For simplicity, the recessive traits will be designated **a**, **b**, **c**, and **d** while the corresponding dominant traits will be designated with a "+". The phenotypes and number of each of the sixteen possible phenotypic classes are given below:

<u>Phenotype</u>	<u>Number</u>	<u>Phenotype</u>	<u>Number</u>
<b>a b c d</b>	2	<b>+ b c d</b>	200
<b>+ + + +</b>	2	<b>a + + +</b>	208
<b>a b c +</b>	11	<b>a b + +</b>	9
<b>+ + + d</b>	10	<b>+ + c d</b>	12
<b>a b + d</b>	1	<b>a + + d</b>	31
<b>+ + c +</b>	3	<b>+ b c +</b>	36
<b>a + c d</b>	34	<b>+ b + d</b>	195
<b>+ b + +</b>	41	<b>a + c +</b>	205

**(a)** What can you conclude about the phenotypes of the two true-breeding parental lines. It will not be possible to specify their phenotypes exactly so please note the nature of the ambiguities that remain. (Hint: before you begin your analysis, it might be helpful to use two-factor cross data to determine which genes are linked.)

**(b)** Draw a genetic map showing the relative positions and genetic distances between the genes for each of the four traits. Linked genes should be grouped together on the same chromosomal segment and any unlinked genes should be placed on a different chromosomal segment.

**2. (a)** There are on average a total of 50 chiasma per meiosis in human males. How many chiasma per meiosis should there be in females, given that the total length of the genetic map measured in females is twice that of males.

**(b)** Consider two human genes that are 10 cM apart. Imagine there that there are features on the chromosomes at the positions of each of the genes that are visible in a microscope observing meiosis in the formation of sperm. If a large number of male meiotic cells were examined, what fraction of these would have a chiasma between the two chromosomal features that are 10 cM apart? Assume that every chiasma leads to a crossover between two of the homologous chromatids and ignore the possibility of double crossovers and crossovers between sister chromatids.

**(c)** Consider two autosomal markers that are found to be 8 cM apart as determined by crosses in the general human population (recombination within males and females are determined with equal frequency). If only crosses where recombination in males were considered what would the measured distance between the markers be, given (as above) that the total length of the genetic map measured in females is twice that of males.



## Problem 1

Part (a) (5 points)

Two true-breeding parents had offspring that included the F1 females that were test crossed with true-breeding recessive males to give the progeny data. By finding which markers are linked and are traveling on the same chromosome in the F1 females, you can get the phenotypes of the true-breeding parents.

Begin by evaluating the combinations of 2-factor cross data:

(i) analysis for markers a & b

$$\begin{aligned} a b &= 2 + 11 + 1 + 9 = 23 \text{ (rec)} \\ ++ &= 2 + 10 + 3 + 12 = 27 \text{ (rec)} \\ a + &= 34 + 208 + 31 + 205 = 478 \text{ (par)} \\ + b &= 41 + 200 + 36 + 195 = 472 \text{ (par)} \end{aligned}$$

a & b are linked  $\Rightarrow$  distance =  $100 \times 50/1000 = 5$  cM

parental  $\Rightarrow$   $\begin{array}{c} a + \\ \text{-----} \\ + b \end{array}$

(ii) analysis for markers a & c

$$\begin{aligned} a c &= 2 + 11 + 34 + 205 = 252 \\ ++ &= 2 + 10 + 41 + 195 = 248 \\ a + &= 1 + 208 + 9 + 31 = 249 \\ + c &= 3 + 200 + 12 + 36 = 251 \end{aligned}$$

all classes are equally represented, so a & c are unlinked.

(iii) analysis for markers a & d

$$\begin{aligned} a d &= 2 + 1 + 34 + 31 = 68 \text{ (rec)} \\ ++ &= 2 + 3 + 41 + 36 = 82 \text{ (rec)} \\ a + &= 11 + 208 + 9 + 205 = 433 \text{ (par)} \\ + d &= 10 + 200 + 12 + 195 = 417 \text{ (par)} \end{aligned}$$

a & d are linked  $\Rightarrow$  distance =  $100 \times 150/1000 = 15$  cM

parental  $\Rightarrow$   $\begin{array}{c} a + \\ \text{-----} \\ + d \end{array}$

(iv) 2-factor cross between b & c confirms that b is unlinked to c

(v) analysis for markers b & d

$$\begin{aligned} b d &= 2 + 1 + 200 + 195 = 398 \\ ++ &= 2 + 3 + 208 + 205 = 418 \\ b + &= 11 + 41 + 9 + 36 = 97 \end{aligned}$$

$$+ d = 10 + 34 + 12 + 31 = 87$$

$$b \ \& \ d \text{ are linked } \Rightarrow \text{ distance} = 100 \times 184/1000 = 18.4 \text{ cM}$$

parental  $\Rightarrow$   $\begin{array}{c} b \ d \\ \text{-----} \\ + \ + \end{array}$

(vi) 2-factor cross between c & d confirms that they are unlinked.

genotype for F1 females (order not accurate at this point)

$$\begin{array}{c} a \ + \ + \\ + \ b \ d \end{array}$$

Therefore, one parent exhibited trait "a", and the other parent exhibited the "b" & "d" traits

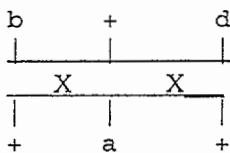
Only one parent exhibited the "c" trait, but it is unclear which one, since "c" is unlinked and segregates independently.

Part (b) (5 points)

We have three linked markers, so we can do a 3-factor cross to determine order:

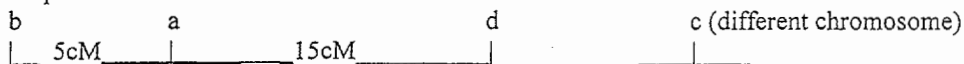
$$\begin{aligned} a \ b \ d &= 2 + 1 = 3 \\ + + + &= 2 + 3 = 5 \\ a \ b \ + &= 11 + 9 = 20 \\ + + \ d &= 10 + 12 = 22 \\ a \ + \ + &= 208 + 205 = 413 \text{---the fact that these classes are the highest confirms} \\ + \ b \ d &= 200 + 195 = 395 \text{---the chromosomal marker locations for the F1 females above} \\ a \ + \ d &= 34 + 31 = 65 \\ + \ b \ + &= 41 + 36 = 77 \end{aligned}$$

Because the abd and +++ reciprocal class is the smallest, a crossover on both sides of the middle marker would have to generate those recombinants. The order that fits this condition is:



so "a" is the middle marker.

Map:



The distance calculated with a two-factor cross between the outer markers will tend to be less than the sum of the internal intervals, because the calculation is not accounting for double crossovers that restore the parental gametes.

Also, preliminary appraisal of the data shows that there are four classes that have the same high frequency. This means that not all of the traits are linked, one of them is unlinked/independent. Because "c" segregates with b & d and with "a" at the same frequency, it must be in a different linkage group than a, b, & d. The 2-factor crosses confirm this.

## Problem 2

Part (a) (3 points)

Since physical distance in males and females is the same, then the # of chiasma per meiosis is twice as high in females than in males (ie. 100 chiasma per meiosis), because

genetic map distance/recombination rate = physical distance

Part (b) (4 points) (see next page for corrections!)

The 10 cM distance in this question was arrived at in the same way as the 8 cM distance in part c: by crosses in the random population, in which males are as equally likely to be the heterozygote as the females. This question is asking for the chiasma in the males only. First you should calculate the distance in just the males:

$y$  = male distance

$2y$  = female distance (since the genetic distance in females is twice that in males)

$$(y + 2y)/2 = 10$$

$$y = 6.666$$

$$2y = 13.333$$

Now we can use the male genetic distance of 6.666 cM to determine the number of chiasma:

$$6.666 \text{ cM} = 100 \times (\# \text{ of recombinants})/\text{total} \quad \# \text{ of recombinants} = 2 \times \# \text{ of chiasma}$$

$$6.666 \text{ cM} = 100 \times (2 \times \# \text{ of chiasma})/\text{total}$$

$$\# \text{ of chiasma}/\text{total} = 1/30$$

Part (c) (3 points)

One way to look at this problem is that the human equivalent of a test cross was done with homozygous males to check the passing of alleles from the mother, and also with homozygous females to check the recombination in the father. The data had a half-and-half distribution of the crosses, and the resulting distance was 8cM. The question is asking for the distance calculated if only the crosses with heterozygous males and homozygous females are used.

If  $y$  = distance using only male recombination data

then  $2y$  = distance using only female recombination data

$$(y + 2y)/2 = 8\text{cM}$$

$$y = 5.333 \text{ cM}$$

## Problem Set 2 Correction

2b.

Method 1

If you thought the 10 cM distance given in the problem was from both males and females, proceed as in the original solutions to get the male-only distance of 6.666 cM.

$$6.666 \text{ cM} = 100 \times \frac{6.666 \text{ recombinant gametes}}{100 \text{ total gametes}}$$

Now we can use a few conversion factors to "change units" from recombinant gametes/total gametes to chiasma/meiotic cells:

$$\frac{6.666 \text{ recombinant gametes}}{100 \text{ total gametes}} \times \frac{1 \text{ chiasma}}{2 \text{ recombinant gametes}} \times \frac{4 \text{ gametes}}{\text{meiotic cell}} = \frac{13.333 \text{ chiasma}}{100 \text{ meiotic cells}}$$

13.333% of the male meiotic cells you examined would have a chiasma between the two features.

Method 2

If you thought the 10 cM distance given in the problem was from males only, use that distance in the same way as above.

$$10 \text{ cM} = 100 \times \frac{10 \text{ recombinant gametes}}{100 \text{ total gametes}}$$

$$\frac{10 \text{ recombinant gametes}}{100 \text{ total gametes}} \times \frac{1 \text{ chiasma}}{2 \text{ recombinant gametes}} \times \frac{4 \text{ gametes}}{\text{meiotic cell}} = \frac{20 \text{ chiasma}}{100 \text{ meiotic cells}}$$

20% of the male meiotic cells you examined would have a chiasma between the two features.

## 7.03 Problem Set 3

due before 5 PM on Thursday, October 15

Hand in answers in recitation section or in the box outside the class

**Note: Question #1 was purposely omitted.**

2. From the table for the genetic code shown on p. 397 of Griffiths et al., you can see that of the 64 possible triplet codons, 61 code for amino acids and only 3 code for chain termination signals. Of course, at the time that the experiments to show that the genetic code was read in triplets were carried out, it was not known whether any more than 20 of the triplets would specify amino acids. Nevertheless, the authors of the triplet code paper were able to deduce that most triplets must code for one or another amino acid and not a stop, simply because so many of the reciprocal frame shift mutant pairs that were examined actually gave a functional protein. To understand this logic, assume that there are two frame shift mutations (one + and the other -), that are 60 nucleotides apart. Knowing that only 3/64 of the codons specify a stop, calculate the probability of not encountering a stop codon in the out of frame sequence that lies between the two frame shift mutations. Next, assume that only 20 of the triplets code for amino acids and the rest specify a stop, recalculate the probability of not encountering a stop codon in the out of frame sequence that lies between the frame shift mutations.

3. Consider the *E. coli* genes **pyrF** and **trpA** which are linked by a cotransduction frequency of 50%. You have isolated two different **trpA** mutations, designated **trpA-1** and **trpA-2**. These two mutations show linkage to each other by a cotransduction frequency of 95%. You would like to determine the order of **trpA-1** and **trpA-2** relative to **pyrF**. Clearly, three factor crosses are the way to do this but there are a number of different ways that a three factor cross could be set up.

(a) One way to do a three factor cross would be to grow phage P1 on a **pyrF<sup>+</sup> trpA-1<sup>-</sup>** host and then infect a **pyrF<sup>-</sup> trpA-2<sup>-</sup>** recipient, selecting for **pyrF<sup>+</sup>**. Then the fraction of **trpA<sup>+</sup>** transductants can be counted to determine the order. Draw out the two possible orders and estimate the frequency of **pyrF<sup>+</sup> that will also be trpA<sup>+</sup>** for each order. Because the transductional mapping function is complex, you will not be able to calculate these frequencies exactly - just make the best estimate that you can.

(b) You will see that the differences in transduction frequency for the two orders will not be so great that you would want to draw a conclusion from the a single absolute determination of cotransduction frequency. Usually the most reliable way to make such determinations of order would be to compare the relative frequencies of two related cotransduction experiments. Propose another three factor cross that you could do to give a cotransduction frequency to which the frequency in part a could be usefully compared.

(c) Another way to perform a three factor cross would be to set up the same transduction experiments as described in parts **a** and **b**, but instead to select for **trpA<sup>+</sup>** and then to determine the fraction of these transductants that are **pyrF<sup>+</sup>**. Describe how you would do each of two relevant three factor crosses. Draw out the two possible orders and then estimate the relative frequency of **pyrF<sup>+</sup>** transductants that would be expected for each order, in each of the two three factor crosses.

## Problem Set 3 Solutions

2.

probability of not encountering a stop codon in the 20 codons = probability that each codon is a coding codon =  $(61/64)^{20} = 0.38$

probability of above if only 20 triplets code for amino acids =  $(20/64)^{20} = 7.9 \times 10^{-11}$

3.

a) Assuming that the given percentages of cotransduction frequencies incorporate probabilities of P1 phage heads packaging the gene markers together and also of a cross over not occurring between the two markers, the percentages can be used to denote approximate distances between markers.

experiment 1:

donor genotype-  $\text{pyrF}^+$ ,  $\text{trpA-1}^-$ ,  $\text{trpA2}^+$

recipient genotype-  $\text{pyrF}^-$ ,  $\text{trpA-1}^+$ ,  $\text{trpA2}^-$

### model 1

```

---pyrF+----- trpA2+--- trpA-1----
  X                   X
-----pyrF------ trpA2---- trpA-1+----- → wt via double crossover
  
```

$\text{pyrF}^+$  is selected for, so all surviving recipients will have received the  $\text{pyrF}^+$ . The cotransduction frequency for  $\text{pyrF}$  and  $\text{trpA}$  is 50%, so in effect, if a recipient becomes  $\text{pyrF}^+$ , 50% of these will have a crossover between  $\text{pyrF}$  and  $\text{trpA}$  markers while 50% will not. The cotransduction frequency of  $\text{trpA-1}$  and  $\text{trpA2}$  is 95%, so 5% of recipients receiving one allele of  $\text{trpA}$  from transduction do not receive the other allele because of a crossover. If the gene order is  $\text{pyrF}$ ,  $\text{trpA1}$ ,  $\text{trpA2}$ , then the probability of a transduction event occurring with the required crossovers to generate a wildtype ( $\text{pyrF}^+$  and  $\text{trpA}^+$ ) is approximated by the following equation:

$$50/100 * 5/100 = 2.5\%$$

### model 2

```

---pyrF+----- trpA-1---- trpA2+---
  X           X           X           X
-----pyrF------ trpA-1+--- trpA2------ → wt via quadruple crossover
  
```

If the gene order is *pyrF*, *trpA1*, *trpA2*, then four crossover events must occur to render the recipient *pyrF*<sup>+</sup> and *trpA*<sup>+</sup>. The probability of the first cross over and the second crossover occurring is 2.5% again, but there is also another crossover that must occur after *trpA2*, and hence the probability is actually less than 2.5%. Hence, this experiment does not resolve the ambiguity of the gene order between the positions of the two *trpA* mutations.

b) If another experiment is set up such that relative frequencies between the two experiments can be compared, it becomes possible to determine the gene orders

experiment 2:

donor genotype- *pyrF*<sup>+</sup>, *trpA*-1<sup>+</sup>, *trpA2*<sup>-</sup>

recipient genotype- *pyrF*<sup>-</sup>, *trpA1*<sup>-</sup>, *trpA2*<sup>+</sup>

model 1

```

---pyrF+-----trpA2----trpA-1+---
X           X           X           X
-----pyrF------trpA2+---trpA-1------ → wt via quadruple crossover
  
```

model 2

```

---pyrF+-----trpA-1+---trpA2----
X                           X
-----pyrF------trpA-1----trpA2+----- → wt via double crossover
  
```

(ie. if experiment 1 has a lower frequency of wt recombinants, then model 2 is likely; if experiment 2 has a lower frequency of wt recombinants, then model 1 is likely)

c) The same experiment can be set up with a screen for *trpA* instead of *pyrF*, and a similar logic can be applied to determine gene orders but the expected frequencies will be different from the above experiments.

experiment 1:

donor genotype- *pyrF*<sup>+</sup>, *trpA*-1<sup>-</sup>, *trpA2*<sup>+</sup>

recipient genotype- *pyrF*<sup>-</sup>, *trpA*-1<sup>+</sup>, *trpA2*<sup>-</sup>

model 1

```

---pyrF+-----trpA2+---trpA-1----
X                           X
-----pyrF------trpA2----trpA-1+----- → wt via double crossover
  
```



If the gene order is pyrF, trpA2, trpA1, then ~ 50% of recombinant trpA<sup>+</sup> transductants will be pyrF<sup>+</sup> due to cotransduction.

model 2

```

---pyrF+----- trpA-1---- trpA2+---
X*           X*           X           X
-----pyrF------ trpA-1+--- trpA2------ → wt via quadruple crossover
  
```

If the gene order is as per model 2, then <50% of recombinant trpA<sup>+</sup> transductants will be pyrF<sup>+</sup> because most of the trpA cells will not have undergone a second double crossover\* to receive the pyrF<sup>+</sup> allele. Though the actual percentages are difficult to derive beyond approximation without actually performing the experiment, a separate experiment done in parallel can be used to compare relative frequencies of pyrF<sup>+</sup> transductants to deduce the correct model.

experiment 2:

donor genotype- pyrF<sup>+</sup>, trpA-1<sup>+</sup>, trpA2<sup>-</sup>  
 recipient genotype- pyrF<sup>-</sup>, trpA1<sup>-</sup>, trpA2<sup>+</sup>

model 1

```

---pyrF+----- trpA2----trpA-1+---
X           X           X           X
-----pyrF------ trpA2+---trpA-1------ → wt via quadruple crossover
  
```

model 2

```

---pyrF+-----trpA-1+---trpA2----
X           X
-----pyrF------ trpA-1----trpA2+----- → wt via double crossover
  
```

Hence, comparing the frequency of pyrF<sup>+</sup> transductants from the two experiments, if the frequency is higher for experiment 1, then model 1 is correct. If the frequency is higher for experiment 2, then model 2 is correct.

## 7.03 Problem Set 4

due before 5 PM on Thursday, October 22

Hand in answers in recitation section or in the box outside the class

1. A region of the *E. coli* chromosome contains three markers in the following order:

ArgA — ThyA — LysA. You have isolated an Hfr that transfers LysA within the first 10 minutes of an interrupted mating experiment, but transfers ArgA or ThyA only after more than 90 minutes of mating.

a) Draw a diagram showing how the F factor integrated into the bacterial chromosome. Show the orientation of the origin of transfer and the location of ArgA, ThyA, and LysA.

b) In an  $F^-$  strain, the markers ArgA and ThyA show linkage of 50% cotransduction. Given that the length of the F plasmid and the P1 phage are about the same, what will the cotransduction frequency of ArgA and ThyA be in the Hfr that was isolated above?

c) Describe how you would isolate an  $F'$  that carries the ThyA and LysA genes (include complete genotypes of the strains that you would use, including any relevant markers such as streptomycin resistance). Will every  $F'$  that carries ThyA also necessarily carry LysA? And will every  $F'$  that carries LysA also necessarily carry ThyA?

2. The region of the *E. coli* chromosome surrounding the Lac operon contains the markers PhoA — Lac I — LacZ,Y,A — ProB. Lac and phoA are about one minute apart and Lac and ProB are about two minutes apart. You have isolated an  $F'$  from a  $LacZ^-$  strain that contains the entire Lac region but does not contain either PhoA or ProB. In order to identify Hfrs that come from the  $F'$  strain you mate 100 individual colonies of a  $PhoA^+ LacZ^+ ProB^+ / F' LacZ^- Str^s$  strain to an  $F^- PhoA^- LacZ^- ProB^- Str^r$  recipient, selecting for  $Str^r$  and then screening for  $PhoA^+$  or  $ProB^+$ . Out of the 100 colonies tested, 3 have the ability to transfer  $ProB^+$  efficiently but none have the ability to transfer  $PhoA^+$  efficiently.

a) Draw the structure of the  $F'$  showing the orientation of the origin of transfer relative to the orientation of the Lac operon.

b) One of the three Hfr strains isolated above transfers  $Lac^+$  efficiently and early. Draw a diagram of the recombination event that converted the  $F'$  strain into to this Hfr strain. Be sure to include the positions of the relevant PhoA, ProB, and LacZ alleles.

c) You next move the  $F'$  into a strain that is  $F^- PhoA^+ LacO^c ProB^+$ . Because of the  $LacO^c$  mutation this new  $F'$  strain expresses  $\beta$ -galactosidase constitutively. From this strain you isolate a rare derivative that shows normal inducible regulation of  $\beta$ -galactosidase. Will this derivative transfer the inducible Lac operon early or late? To explain your answer, diagram the recombination event that gives the normally regulated  $\beta$ -galactosidase.

**3** Consider a hypothetical operon that controls the ability of *E. coli* to utilize sucrose by controlling the synthesis of a sucrose permease and the enzyme sucrose. When glucose is available as a carbon source, the synthesis of both the permease and sucrose is shut off:

	- glucose	+ glucose
sucrose	high	low
permease	high	low

There are three linked mutations ( $A^-$ ,  $B^-$ , and  $C^-$ ) which alter the expression of sucrose activity in response to glucose. You construct the following bacterial strains and analyze the regulation of sucrose and permease.

Genotype	Sucrose activity		Permease activity	
	- glucose	+ glucose	- glucose	+ glucose
$A^- B^+ C^+$	low	low	high	low
$A^+ B^- C^+$	low	low	low	low
$A^+ B^+ C^-$	high	high	high	high
$A^- B^+ C^+ / F' A^+ B^+ C^+$	high	low	high	low
$A^+ B^- C^+ / F' A^+ B^+ C^+$	low	low	low	low
$A^+ B^+ C^- / F' A^+ B^+ C^+$	high	high	high	high
$A^- B^- C^+ / F' A^+ B^+ C^+$	low	low	low	low
$A^- B^+ C^+ / F' A^+ B^- C^+$	low	low	low	low
$A^- B^+ C^- / F' A^+ B^+ C^+$	high	low	high	high
$A^- B^+ C^+ / F' A^+ B^+ C^-$	high	high	high	high

a) Propose a molecular model for sucrose regulation that explains the nature of the A, B, and C mutations and how glucose shuts off sucrose expression. Be as specific as you can.

b) Based on your model from part (a), how would you expect sucrose expression in a strain with the genotype  $A^+ B^- C^-$  to respond to glucose?

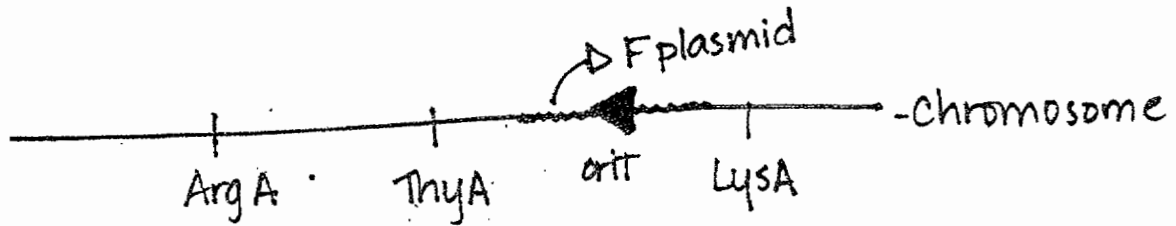
You have isolated a dominant mutation,  $D^d$ , involved in regulation of sucrose activity, that is unlinked to the A, B, or C mutations. You construct the following strains to analyze how this mutation affects sucrose regulation:

Strains	Levels of sucrose activity	
	- glucose	+ glucose
$A^+ B^+ C^+ D^+$	high	low
$A^+ B^+ C^+ D^d$	high	high
$A^+ B^- C^+ D^d$	low	low

c) Propose a mechanism to explain how the  $D^d$  mutation fits into the model that you proposed above. Also propose new possible mechanisms for the  $B^-$  mutation and the action of glucose in light of the existence of the  $D^d$  mutation.

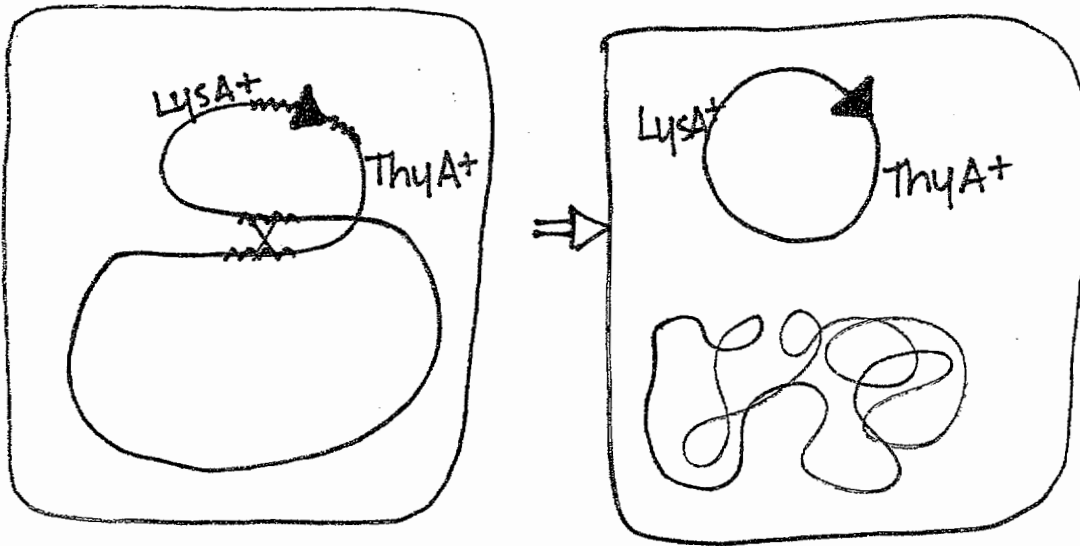
## PROBLEM SET 4 SOLUTIONS

1. a) The F plasmid has integrated between the Thy A gene and the LysA gene, oriented such that LysA will be transferred early and ThyA late.

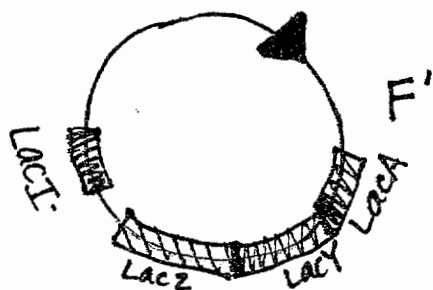


- b) Since the F plasmid has not changed the distance between the ArgA gene and the ThyA gene, their cotransduction frequency will remain 50% in the Hfr.

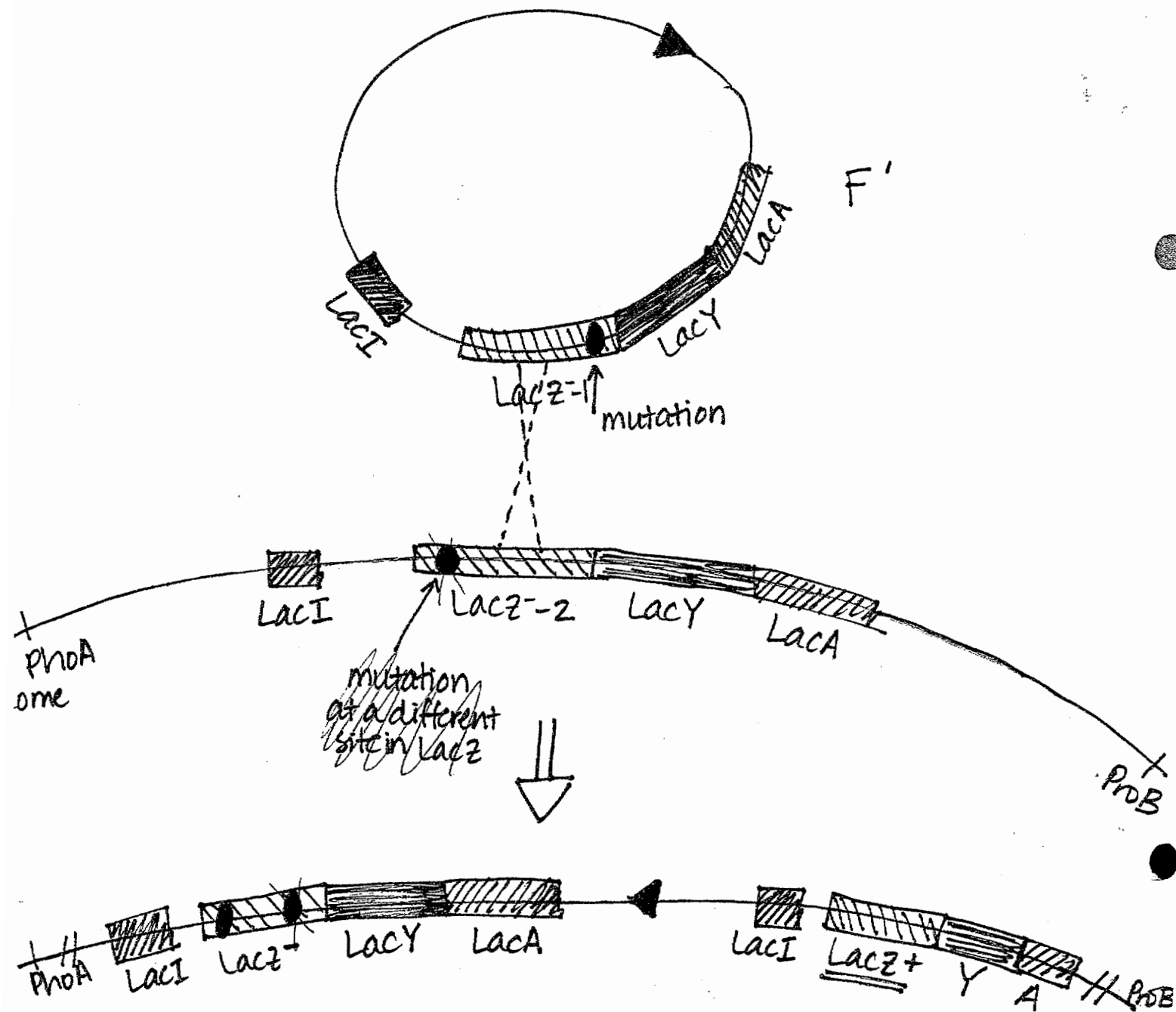
c.) The object here is to find a rare, specific excision event among a population of Hfr cells (see below). This F' can be isolated by mating the above Hfr population (that is streptomycin sensitive) with a streptomycin resistant, ThyA<sup>-</sup>, LysA<sup>-</sup>, F<sup>-</sup> strain. After a brief mating, select for str<sup>r</sup>, LysA<sup>+</sup>, ThyA<sup>+</sup>. Only those cells that have received an F' containing LysA<sup>+</sup> ThyA<sup>+</sup> will be able to grow and form colonies. When the other Hfr cells (which do not contain an F') transfer, they will transfer LysA<sup>+</sup> early, but the mating will be disrupted before ThyA<sup>+</sup> is transferred.



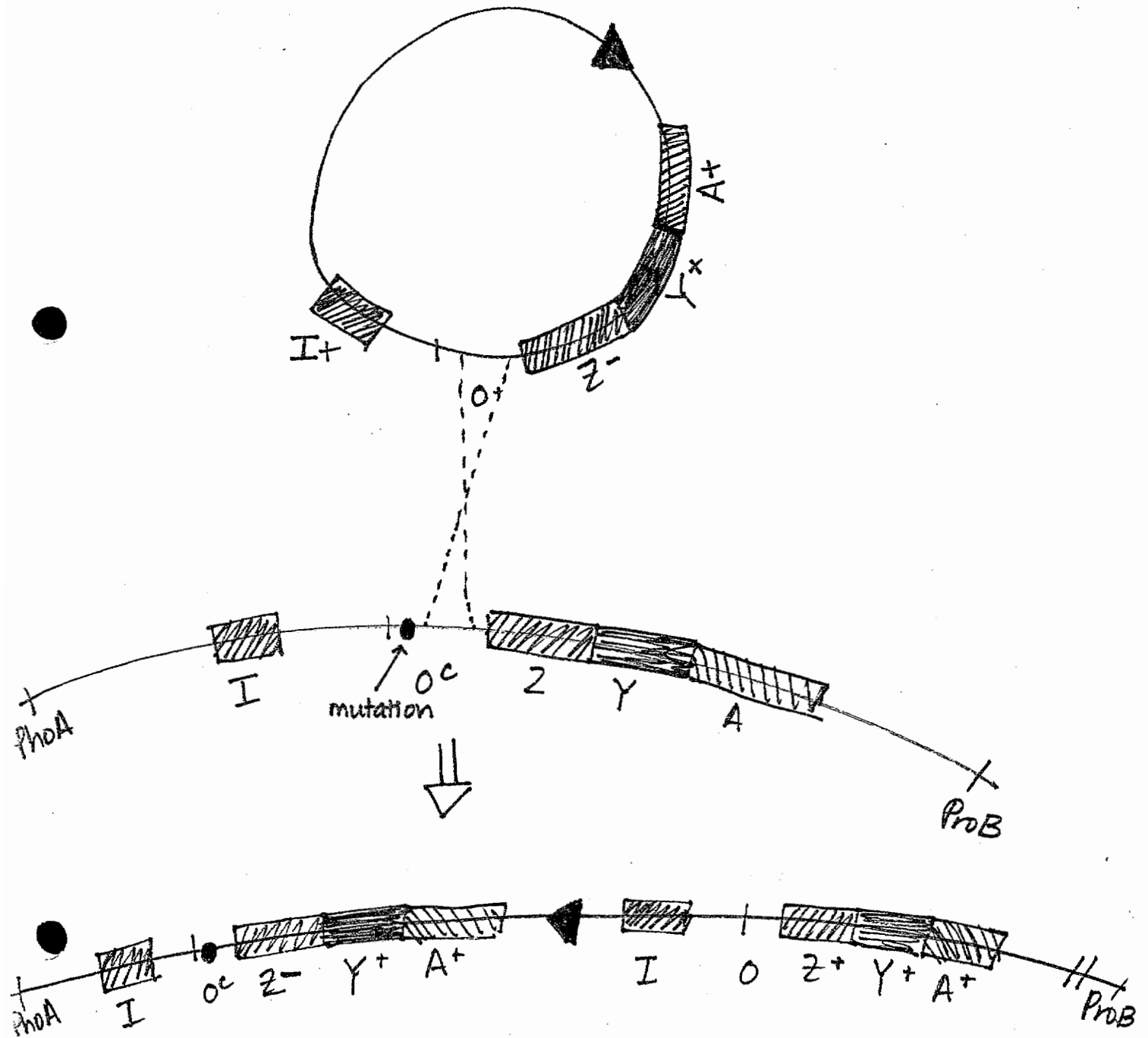
2.a) The structure of the F' is seen here:



b) Although the LacZ gene is defective on the F' and on the chromosome, the two LacZ- alleles may not be due to the same mutation. If this is the case, then an integration/recombination event could have occurred as diagrammed below. Note, this event happens after the transfer of the F' to the F-.



2. c) If the integration recombination event occurs as diagrammed below, then the  $O^c$  mutation will be in *cis* to the LacZ- mutation, while the wildtype operator will be adjacent to the wildtype copy of LacZ. This will result in normal, inducible expression of LacZ.



3. a. The **A** mutation selectively affects sucrase function without disturbing permease function (line 1). From the data in line 4, we see that the **A** mutation is recessive to wildtype, since the mero-diploid is normally induced. A simple explanation for these data is that the mutation is a loss of function mutation in the sucrase gene itself.

The **B** mutation is uninducible since both sucrase and permease activities are low even in the absence of glucose. From the data in line 5, we see that the **B** mutation is dominant, and from the data in lines 7 and 8, we see that the **B** mutation is *trans*-acting. This can be explained if the **B** mutation is in the repressor for the sucrose operon, creating a super-repressor.

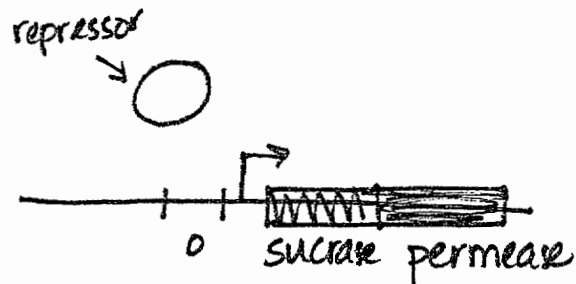
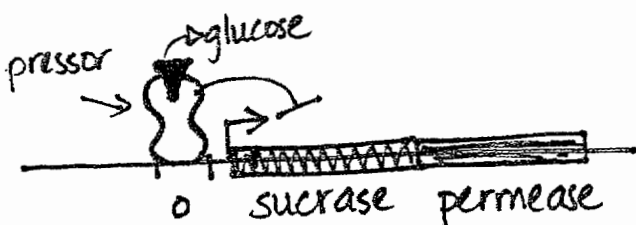
The **C** mutation is constitutive, since the sucrase and permease activities are high even in the presence of glucose. From the data in line 6, we see that the **C** mutation is dominant, and from the data in lines 9-10, we see that the **C** mutation is *cis*-acting. (We can only see the constitutive phenotype of the **C** mutation when it is on the same piece of DNA as a wildtype copy of the **A** gene). This suggests that the **C** mutation is in the operator, the binding site for the repressor. The repressor cannot bind the mutant operator.

Model:

In the presence of glucose, glucose binds the repressor, enabling it to bind the operator and prevent expression of the sucrose operon. In the absence of glucose, the repressor can no longer bind the operator, and expression of the sucrose operon is induced. The **B** mutation enables the repressor to bind the operator even in the absence of glucose.

+ glucose

- glucose



b.) In strain with the genotype  $A^+B^-C^-$ , the operator mutation would prevent binding by the mutant repressor. Therefore, sucrase expression would be constitutive in this strain, and the **C** mutation is epistatic to **B**. (Since the phenotype of the double mutant is constitutive, **C** must be acting downstream of **B**).



Note: Problem Set #5

was purposely omitted.

## PROBLEM SET 6 ANSWERS

1. You are studying the regulation of the mitose metabolism. Mitose is a fictional sugar which is metabolized by the MIT10 gene. MIT10 gene expression requires the MIT1 and MIT2 genes and is induced by the presence of mitose. MIT2 is an activator of MIT10 gene expression and has two domains, a DNA-binding domain and an activation domain. MIT1 is a repressor of MIT10 gene expression, which binds to the activation domain of MIT2 while MIT2 is on the DNA. You collect 3 mutants which are defective in MIT10 expression and want to learn more about them. You do the following experiments:

	MIT10 levels	
	<u>+mitose</u>	<u>-mitose</u>
wt	200	1
A-	200	200
B-	1	1
C-	1	1
mit1-1	200	200
mit2-1	1	1
A-/A+	200	200
B-/B+	200	1
C-/C+	1	1
mit1-1/MIT1	200	1
mit2-1/MIT2	200	1

a.) Describe the nature of the A, B, and C mutants and how they affect the regulation of MIT10 expression.

**A: constitutive, dominant**

**B: uninducible, recessive**

**C: uninducible, dominant**

You construct the following strains and measure their MIT10 levels:

	MIT10 levels	
	<u>+mitose</u>	<u>-mitose</u>
A-/mit2-1	200	200
B-/mit2-1	1	1
C-/mit2-1	1	1

b.) Are any of the A, B, C mutants in MIT2? If so, which one(s)?

**Yes, B- is a mutation in the MIT2 gene. Since mit2-1 in the diploid strain is not able to complement the B- mutation. (You can only do complementation testing with recessive mutants.)**

(Actually, A- is in the MIT2 gene, but it is not able to be determined from this data, as you cannot do complementation testing with dominant mutants)

You mate the A- mutant to the mit1-1 mutant. You take the resulting diploid, and put it under starvation conditions so that it sporulates. When you dissect 300 tetrads and test the spores for MIT10 levels, you find 3 different classes of tetrads. 48 are of class I, 51 are of class II, and 201 are of class III.

When you mate the A- mutant to the mit2-1 mutant and then sporulate the diploid, you find that all 300 tetrads which you dissect are of class IV.

	MIT10 levels	
	<u>+mitose</u>	<u>-mitose</u>
class I: all 4 spores	200	200
class II: 2 spores	200	200
2 spores	200	1
class III: 3 spores	200	200
1 spore	200	1
class IV: 2 spores	200	200
2 spores	1	1

c.) In which gene is the A- mutation and what is the nature of the mutation? Explain.

**A- is in the MIT2 mutation. When the A- mutation is mated to the mit2-1 strain, only parental ditype tetrads are found, indicating that the A- mutation is completely linked to the mit2-1 mutation (therefore, they are in the same gene.)**

**Since the A- mutation leads to constitutive expression of MIT10, and the A- mutation is in MIT2, the mutation most likely results in the inability of MIT2 to bind MIT1 (the repressor), but still allows MIT2 to activate MIT10 expression.**

You construct a fusion protein (MIT2-VP16) which has the DNA binding domain of MIT2 fused to the viral activation domain of the VP16 protein. You introduce it into the following strains and measure MIT10 levels:

	MIT10 levels	
	<u>+mitose</u>	<u>-mitose</u>

wt, MIT2-VP16	200	200
C-, MIT2-VP16	200	200

d.) In which gene is the C- mutation and what is the nature of the mutation?

**The C- mutation is in the MIT1 gene, and results in its acting as a super-repressor. C- makes the MIT1 gene product insensitive to mitose levels and therefore MIT1 always binds to the activation domain of MIT2. In the MIT2-VP16 experiment, there is no activation domain to bind to, so expression is constitutive as MIT2-VP16 can never bind MIT1 (regardless of the presence of the C- mutation).**

**Note: Question #2 was purposely omitted.**

3. You are interested in studying the secretion of invertase in *S. cerevisiae*. Yeast cells must secrete invertase in order to grow on raffinose. You have isolated a few mutants that cannot secrete invertase and thus cannot grow on raffinose. You have chosen one of these mutants to study in depth which you call *isd1* for invertase secretion defect. In the course of your study, you have isolated several different mutant alleles of *ISD1*: *isd1-1*, *isd1-2*, and *isd1-3*.

Strains containing each of these alleles, when mated to wildtype strains, produce the following results when tetrads are dissected and grown on raffinose:

2 alive:2 dead 100%

To study *ISD1* more in-depth, you decide to conduct suppression analysis. You mutagenize a strain containing *isd1-1* and isolate several suppressors which are now able to grow on raffinose.

a.) When you cross one of the suppressor strains to a wildtype strain you get the following result when the tetrads are grown on raffinose:

4 alive: 0 dead 353 (100%)

Propose two models for the nature of this suppressor, and describe an experiment that would distinguish the two models.

**Only a single tetrad class is represented, which is the parental ditype class. Thus, the two mutations (the suppressor mutation and the original *isd1-1* mutation) are closely linked. Thus, the suppressor is most likely intragenic or a true revertant. To distinguish these**