7.03 Problem Set 1

Due before 5 PM on Thursday, September 23, 2004
Hand in answers to the appropriate slot in the box outside of 68-120.
Late problem Sets will NOT be accepted.

1. After a long summer of complete freedom, you return to your UROP in a yeast genetics lab. You examine some old plates and discover two different types of yeast colonies. Most of the colonies appear large as expected, but some colonies are very small.

For each small colony, you determine that the phenotype (which you call the puny phenotype) is due to a single mutation. You choose five puny strains, p1- p5 and perform a complementation test on these mutants.

In the table below, "normal" indicates that the diploid created produces a normal sized colony, and "puny" indicates that the diploid produces a small colony.

	p1	p2	p3	p4	p5
p1	puny	puny	puny	normal	normal
p2	puny	puny	puny	normal	normal
p3	puny	puny	puny	puny	puny
p4	normal	normal	puny	puny	puny
p5	normal	normal	puny	puny	puny

a) Assign the mutants p1-p5 into complementation groups.

You receive two characterized puny mutants from your advisor. One mutant*, pnyA*, has a mutation in gene A. The other mutant*, pnyB*, has a mutation in gene B. You want to use these to determine if any of your mutants p1-p5 have mutations in gene A or gene B.

b) You do a complementation test by mating mutant p1 to a *pnyA* strain. You demonstrate that the mutation in p1 is **not** in gene A. You do another complementation test by mating mutant p1 to a *pnyB* strain. You demonstrate that the mutation in p1 is in gene B. Complete the table below in a way that is consistent with this conclusion.

NOTE: For the tables below, use capital letters (A or B) to represent alleles that give the large colony phenotype. Use lowercase letters (a or b) to represent mutant alleles that give the recessive small colony phenotype. Use (A* or B*) to represent mutant alleles that give a dominant small colony phenotype. The first block is filled in for you as an example.

	pnyA yeast	*pnyB* yeast	Mutant p1 yeast	mutant p1 x *pnyA* yeast	mutant p1 x *pnyB* yeast
Genotype at the A locus and the B locus	aB				
Ploidy (haploid or diploid)					
Phenotype (puny or normal)					

c) You do a complementation test by crossing mutant p3 to a *pnyA* cell and conclude that mutant p3 has a mutation in gene A. Your roommate does a complementation test by crossing mutant p3 to a *pnyB* cell and concludes mutant p3 has a mutation in gene B. What mistake has been made?

d) By DNA sequencing it is determined that mutant p3 has a mutation in gene A. Complete the table below in a way that is consistent with both you and your roommate's observations.

	Mutant p3 yeast	mutant p3 x *pnyA* yeast	mutant p3 x *pnyB* yeast
Genotype at the A locus and the B locus			
Ploidy (haploid or diploid)			
Phenotype (puny or normal)			

2. In humans, blood type (type A, type B, type AB, or type O) is determined by three alleles at a single locus, allele I^A, allele I^B, and allele i. The genotypes that result in each blood type in humans are shown in the table below.

Blood type	Genotype(s)
A	$I^A I^A$ or I^A i
B	$I^B I^B$ or I^B i
AB	$I^A I^B$
O	i i

You are interested in determining whether blood type in a certain fast-breeding species of primate follows the same inheritance pattern as it does in humans. You cross a primate of blood type A to a primate of blood type B to obtain F1 primates that all have the same blood type. Crosses among these F1 primates yield 40 progeny; 6 with blood type A, 13 with blood type B, and 21 with blood type AB.

a) Could these results be consistent with a blood type inheritance pattern that is the same as that of humans?

b) You decide to use the chi-squared test to determine the statistical significance of your results. How many degrees of freedom will you use?

c) What is the chi-squared value you obtain?

d) Using the table below of chi-squared probabilities, can you conclude with confidence that blood type in this primate is determined by the inheritance pattern it follows in humans?

P value:	.995	.975	0.9	0.5	0.1	0.05	0.025	0.01	0.005
df = 1	.000	.000	.016	.46	2.7	3.8	5.0	6.6	7.9
df = 2	.01	.05	.21	1.4	4.6	6.0	7.4	9.2	10.6
df = 3	.07	.22	.58	2.4	6.3	7.8	9.3	11.3	12.8

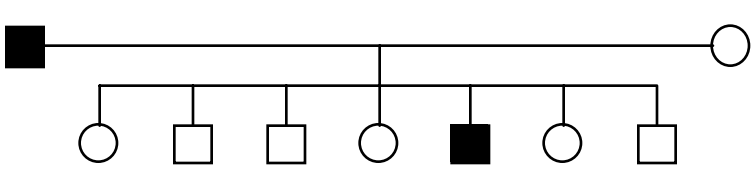
e) Say that you crossed a primate of blood type A to a primate with blood type A, and obtained progeny, , some of which have blood type A and others have blood type O. Would this result be consistent with the hypothesis that blood type in this primate follows the same mode of inheritance as it does in humans? **If so** give the genotypes of the parents.

f) Say that you crossed a primate of blood type A to a primate with blood type AB, and obtained progeny, some of which have blood type A, some have blood type B, and some have blood type AB. Would this result be consistent with the hypothesis that blood type in this primate follows the same mode of inheritance as it does in humans? **If so** give the genotypes of the parents.

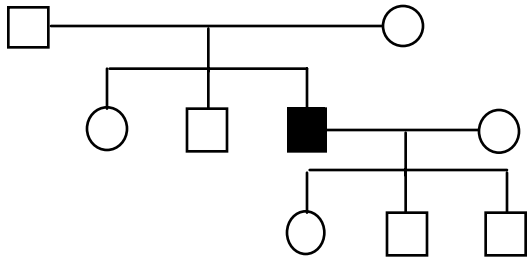
g) Say that you crossed a primate of blood type O to a primate with blood type AB, and obtained progeny, some of which have blood type A, some have blood type B, and some have blood type AB. Would this result be consistent with the hypothesis that blood type in this primate follows the same mode of inheritance as it does in humans? **If so** give the genotypes of the parents.

3. **a)** For each of the following pedigrees showing inheritance of a trait indicated by shaded circles and squares in the pedigree, list **all** possible modes of inheritance (the four options are: autosomal recessive, autosomal dominant, X-linked recessive, and X-linked dominant). Assume 100% penetrance and no spontaneous mutations. Also assume that the trait is rare and therefore any individuals that marry into the family after the starting generation do not possess any alleles that lead to the trait in question.

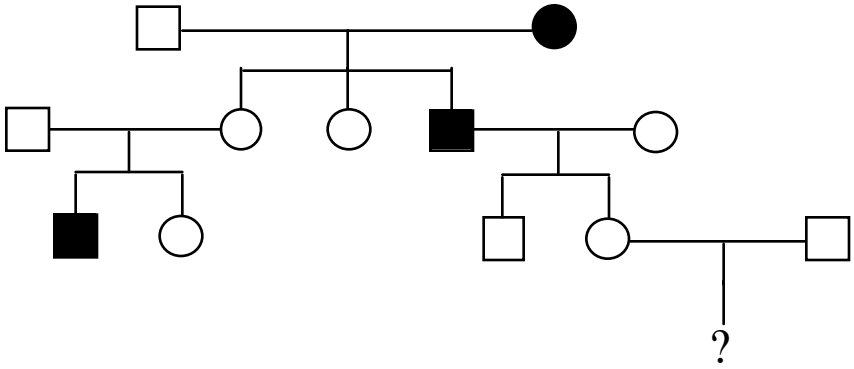
i)



ii)



iii)



b) For the pedigree in part **(iii)** consider the unborn child labeled with a question mark. For each possible mode of inheritance give the probability of that child will show the trait.

c) For the same pedigree assume that identical twin daughters are born, neither of which show the trait (identical twins arise from the same zygote). Give the probability that both twins are carriers for each possible mode of inheritance.

d) Now assume that fraternal twin daughters are born, neither of which show the trait (fraternal twins arise from different zygotes). Give the probability that both twins are carriers for each possible mode of inheritance.

7.03 Problem Set 1 - KEY

Due before 5 PM on Thursday, September 23, 2004
Hand in answers to the appropriate slot in the box outside of 68-120.

1. After a long summer of complete freedom, you return to your UROP in a yeast genetics lab. You examine some old plates and discover two different types of yeast colonies. Most of the colonies appear large as expected, but some colonies are very small.

For each small colony, you determine that the phenotype (which you call the puny phenotype) is due a single mutation. You choose five puny strains, p1-p5 and perform a complementation test on these mutants.

In the table below, “normal” indicates that the diploid created produces a normal sized colony, and “puny” indicates that the diploid produces a small colony.

	p1	p2	p3	p4	p5
p1	puny	puny	puny	normal	normal
p2	puny	puny	puny	normal	normal
p3	puny	puny	puny	puny	puny
p4	normal	normal	puny	puny	puny
p5	normal	normal	puny	puny	puny

a) Assign the mutants p1-p5 into complementation groups.
A complementation group is a group whose members do NOT complement each other. In other words, it's a group whose members all have mutations in the same gene. If you mate a yeast strain from one group to a yeast strain of another group, the strains should complement each other and result in a diploid cell with WT phenotype. If you mated two strains with mutations in the same gene the resulting diploid cell would have the mutant phenotype. Note that complementation tests only apply to recessive alleles. Here the complementation groups are the following:

(p1, p2)
(p4, p5)
p3 is behaving as if the allele gives a dominant phenoptye and can not be placed.

You receive two new puny mutants from your advisor. One mutant, *pnyA*, has a mutation in gene A. The other mutant, *pnyB*, has a mutation in gene B. You want to use these to determine if any of your mutants p1-p5 have mutations in gene A or gene B.

b) You do a complementation test by crossing mutant p1 to a *pnyA* cell. You demonstrate that the mutation in p1 is not in gene A. You do another complementation test by crossing mutant p1 to a *pnyB* cell. You demonstrate that the mutation in p1 is in gene B. Complete the table below in a way that is consistent with this conclusion.

NOTE: For the tables below, use capital letters (A or B) to represent alleles that give the large colony phenotype. Use lowercase letters (a or b) to represent mutant alleles that give the recessive small colony phenotype. Use (A* or B*) to represent mutant alleles that give a dominant small colony phenotype. The first block is filled in for you as an example.

	pnyA yeast	*pnyB* yeast	Mutant p1 yeast	mutant p1 x *pnyA* yeast	mutant p1 x *pnyB* yeast
Genotype at the A locus and the B locus	**aB**	**Ab**	**Ab**	**AaBb**	**AAbb**
Ploidy (haploid or diploid)	**haploid**	**haploid**	**haploid**	**diploid**	**diploid**
Phenotype (puny or normal)	**puny**	**puny**	**puny**	**normal**	**puny**

c) You do a complementation test by crossing mutant p3 to a *pnyA* cell and conclude that mutant p3 has a mutation in gene A. Your roommate does a complementation test by crossing mutant p3 to a *pnyB* cell and concludes mutant p3 has a mutation in gene B. What mistake has been made?

The p3 allele is behaving as a dominant allele and cannot be used in a complementation test.

d) By DNA sequencing it is determined that mutant p3 has a mutation in gene A. Complete the table below in a way that is consistent with both you and your roommate's observations.

	Mutant p3 yeast	mutant p3 x *pnyA* yeast	mutant p3 x *pnyB* yeast
Genotype at the A locus and the B locus	**A*B**	**A*aBB**	**A*ABb**
Ploidy (haploid or diploid)	**haploid**	**diploid**	**diploid**
Phenotype (puny or normal)	**puny**	**puny**	**puny**

2. In humans, blood type (type A, type B, type AB, or type O) is determined by three alleles at a single locus, allele I^A, allele I^B, and allele i. The genotypes that result in each blood type in humans are shown in the table below.

Blood type	Genotype(s)
A	$I^A I^A$ or I^A i
B	$I^B I^B$ or I^B i
AB	$I^A I^B$
O	i i

You are interested in determining whether blood type in a certain fast-breeding species of primate follows the same inheritance pattern as it does in humans. You cross a primate of blood type A to a primate of blood type B to obtain F1 primates that all have the same blood type. Crosses among these F1 primates yield 40 progeny; 6 with blood type A, 13 with blood type B, and 21 with blood type AB.

a) Could these results be consistent with a blood type inheritance pattern that is the same as that of humans?

Yes. Similar data could be generated if the following crosses were performed:

P **A: $I^A I^A$ x B: $I^B I^B$**

F1 **$I^A I^B$**

F2 **¼ $I^A I^A$ ¼ $I^B I^B$ ½ $I^A I^B$**

Homozygous type A and type B parents mate to create an F1 generation of all type AB. These type AB primates are then mated to produce an F2 generation in a 1:2:1 ratio (A : AB : B). Thus we approximate that one half of the 40 progeny will be of blood type AB, which is indeed the case. Here the exact number is 21. Similarly, we expect those with blood type A to constitute ¼ of the 40 progeny, and the same for Type B. Our results roughly approximate this expectation with 6 and 13 in each respective class.

b) You decide to use the chi-squared test to determine the statistical significance of your results. How many degrees of freedom will you use?

df = (number of phenotypic classes) – 1
Here the possible phenotypic classes are blood type A, B, and AB
df = 3 – 1 = 2

c) What is the chi-squared value you obtain?

$$X^2 = \sum \frac{(O-E)^2}{E} = \frac{(6-10)^2}{10} + \frac{(13-10)^2}{10} + \frac{(21-20)^2}{20} = 2.55$$

d) Using the table below of chi-squared probabilities, can you conclude with confidence that blood type in this primate is determined by the inheritance pattern it follows in humans?

P value:	.995	.975	0.9	0.5	0.1	0.05	0.025	0.01	0.005
df = 1	.000	.000	.016	.46	2.7	3.8	5.0	6.6	7.9
df = 2	.01	.05	.21	1.4	4.6	6.0	7.4	9.2	10.6
df = 3	.07	.22	.58	2.4	6.3	7.8	9.3	11.3	12.8

No. The p value corresponding to our chi-squared value lies between 0.5 and 0.1 This value is above the 0.05 cutoff and therefore this inheritance pattern is not rejected. We cannot, however, conclusively conclude that the inheritance pattern follows that of humans. Note that the chi-squared test is used to reject hypotheses, not to accept them.

e) Say that you crossed a primate of blood type A to a primate with blood type A, and obtained progeny, some of whom are blood type A and some of whom are blood type O. Would this result be consistent with the hypothesis that blood type in this primate follows the same mode of inheritance as it does in humans? **If so** give the genotypes of the parents.
Yes. The cross would look like the following:

P **A: $I^A i$ x $I^A i$**

F1 **A: $I^A I^A$ and $I^A i$**
O: ii

f) Say that you crossed a primate of blood type A to a primate with blood type AB, and obtained progeny, some of whom are blood type A, some of whom are blood type B, and some of whom are blood type AB. Would this result be consistent with the hypothesis that blood type in this primate follows the same mode of inheritance as it does in humans? **If so** give the genotypes of the parents.

Yes. The cross would look like the following:

P **A: $I^A i$ x AB: $I^A I^B$**

F1 **A: $I^A I^A$ and $I^A i$**
B: $I^B i$
AB: $I^A I^B$

g) Say that you crossed a primate of blood type O to a primate with blood type AB, and obtained progeny, some of whom are blood type A, some of whom are blood type B, and some of whom are blood type AB. Would this result be consistent with the hypothesis that blood type in this primate follows the same mode of inheritance as it does in humans? **If so** give the genotypes of the parents.

No. Blood type AB is not possible with this cross under the human pattern of inheritance, only blood types A and B are:

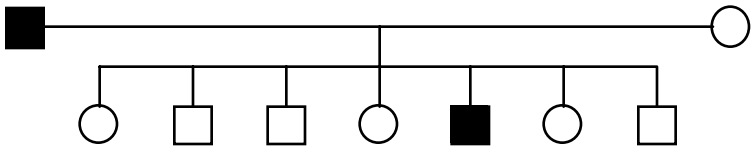
P **O: ii x AB: $I^A I^B$**

F1 **A: $I^A i$**
B: $I^B i$

3. a) For each of the following pedigrees, list **all** modes of inheritance (the four options are: autosomal recessive, autosomal dominant, X-linked recessive, and X-linked dominant) that are consistent with the trait indicated by shaded circles and squares in the pedigree. Assume 100% penetrance and no spontaneous mutations. Also assume that

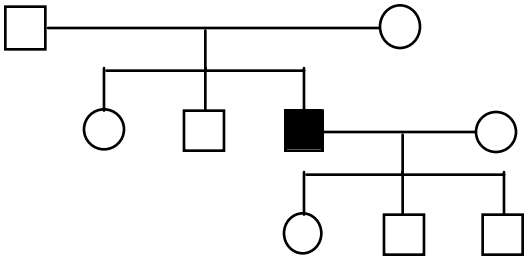
any individuals that marry into the family after the starting generation do not possess any alleles that lead to the trait in question.

i)



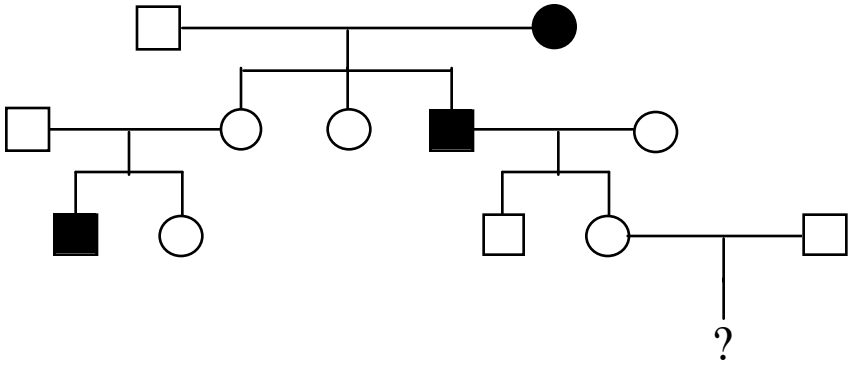
autosomal recessive
autosomal dominant
X-linked recessive

ii)



autosomal recessive
X-linked recessive

iii)



X-linked recessive

b) For each mode of inheritance that is consistent with the pedigree from part **(iii)**: if the couple in decides to have the child labeled with a question mark, what is the probability of that child showing the trait indicated by shaded circles and squares?

Because the father is an unaffected male, he must not have the mutant allele for this X-linked recessive trait. Therefore it is impossible for a daughter of this couple, who will receive an X chromosome with the WT allele from her father, to show this trait. The probability of a child of this couple showing the trait then becomes the probability of having a male child AND the mother passing her X chromosome with the mutant allele (which she received from her father) to her son. We know the mother is a carrier of the allele even though she is unaffected, because her father bore the trait, and passed his X chromosome to his daughter. Therefore her genotype is X^AX^a. The probability of the child having the trait is:

p = p(male child) x p(X^a allele from mother)
= ½ x ½
= ¼

c) For each mode of inheritance that is consistent with the pedigree from part **(iii):** if identical twin daughters who do not show the trait are born to this couple, what is the probability that both children are carriers?

Identical twins develop from a single zygote and therefore you must treat them as the result of a single meiosis. Note that the problem already gives you the condition that they are female, so you do not need to account for the probability of the couple having a female child. The twins will receive X^A from their father; thus the probability of the daughters being carriers is the probability that the heterozygous mother passes on the X^a allele.

p(twin carriers) = p(one carrier)
= p(X^a allele from mother)
= ½

d) For each mode of inheritance that is consistent with the pedigree from part **(iii):** if fraternal twin daughters who do not show the trait are born to this couple, what is the probability that both children are carriers?

Here the twins are not identical and therefore we must treat them as the result of two separate meioses.

p(twin carriers) = p(daughter 1 carrier) x p(daughter 2 carrier)
= p(X^a allele from mother) x p(X^a allele from mother)
= ½ x ½
= ¼

7.03 Problem Set 2

Due before 5 PM on Thursday, September 30, 2004
Hand in answers in recitation section or in the box outside of 68-120

1. You are studying three autosomal mutations in the fruitfly *Drosophila melanogaster*. The cw^- mutation gives the recessive "curly-wing" phenotype (wild-type flies (cw^+) are straight-winged), the sh^- mutation gives the recessive "short-bristled" phenotype (wild-type flies (sh^+) are long-bristled), and the nh^-mutation gives the recessive "no-bristled" phenotype (wild-type flies (nh^+) are long-bristled).

a) You mate a homozygous nh^- cw^+ sh^- fly (which is no-bristled and straight-winged) to a true-breeding long-bristled and curly-winged fly to obtain an F1 generation. **List all** phenotypic categories of the flies in the F1 generation **and** the ratio in which these phenotypic categories are found in the F1 generation. The possible phenotypic classes are: short-bristled and curly-winged, long-bristled and straight-winged, short-bristled and straight-winged, long-bristled and curly-winged, no-bristled and curly-winged, no-bristled and straight-winged.

b) To do a proper three-factor cross, you need to create an F2 generation by mating F1 females to a specific true-breeding male.

i) What is the **phenotype** of the true-breeding male fly you would choose as the parent for the F2 generation?

ii) What is the **diploid genotype** of the true-breeding male you would choose as the parent for the F2 generation?

You do a proper three-factor cross and, in the F2 generation, you obtain the following number of flies in each of the corresponding phenotypic classes:

short-bristled and curly-winged	(230 flies)
long-bristled and straight-winged	(10 flies)
short-bristled and straight-winged	(11 flies)
long-bristled and curly-winged	(238 flies)
no-bristled and curly-winged	(28 flies)
no-bristled and straight-winged	(483 flies)

c) All of the long-bristled and straight-winged flies have the alleles sh^+ nh^+ cw^+ on the chromosome they inherited from their mother. Using this notation, list all of the possible chromosomes that the no-bristled and curly-winged flies could have inherited from their mother.

d) Are sh^- and nh^- linked? **If so**, give the distance between them in cM.

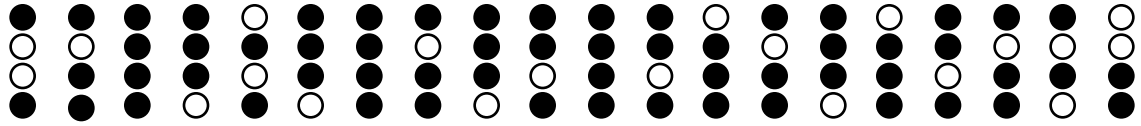
e) At this point, can you conclude whether sh^- and nh^- are in the same gene or not? **If so,** state whether sh^- and nh^- are in the same gene or in different genes.

f) Are sh^- and cw^- linked? **If so**, give the distance between them in cM.

g) Are nh^- and cw^- linked? **If so**, give the distance between them in cM.

2. You find a haploid yeast strain in the freezer that is marked "double mutant strain which contains $ade7^-$ and $ade8^-$." Like some other ade mutants you have studied before, this double mutant strain is red in color, as compared to wild-type yeast, which are white. You cannot find any other information regarding this strain or the mutations it contains.
a) Mating of the $ade7^-$ $ade8^-$ double mutant strain to wild type produces diploids that are white. What do these results tell you about the $ade7^-$ and $ade8^-$ mutations?

b) Sporulation of the diploids produced by the mating of $ade7^-$ $ade8^-$ double mutant yeast to wild-type yeast yields three different types of tetrads. Type One contains 4 red spores, Type Two contains 2 red spores and 2 white spores, and Type Three contains 3 red spores and 1 white spore. Below, all of the tetrads that you analyzed are drawn. (Each column below contains the four colonies that grew up from each tetrad you have analyzed. Dark colored circles indicate red yeast colonies, and white colored circles represent white yeast colonies.)



b) continued

i) How many of the tetrads that you analyzed are parental ditypes?

ii) How many are tetratypes?

iii) How many are nonparental ditypes?

c) Are ade7$^-$ and ade8$^-$ linked? **If so**, give the distance between them in cM.

d) At this point, can you conclude whether ade7$^-$ and ade8$^-$ are in the same gene or not? **If so,** state whether ade7$^-$ and ade8$^-$ are in the same gene or in different genes.

e) At this point, can you determine the single mutant phenotypes of ade7$^-$ yeast and ade8$^-$ yeast? **If so,** state whether ade7$^-$ single mutant yeast are red or white.

f) Below, the first two tetrads that you analyzed in part **(b)** are drawn. Some of the spores have their genotypes at the ade7 and ade8 loci written in for them. For the remaining spores (that is, those that are marked with question marks) give the genotypes at the ade7 and ade8 loci.

The first tetrad from part **(b)**:

- ● ?
- ○ ?
- ○ ?
- ● ?

The second tetrad from part **(b)**:

- ● ?
- ○ ?
- ● ade7$^-$ ade8$^-$
- ● ade7$^+$ ade8$^-$

g) You already had a mutant in your collection that gave a recessive red phenotype, called ade1⁻. You mate ade1⁻ yeast to ade7⁻ yeast, and sporulate the resulting diploid. Out of 40 tetrads, 39 contain four red spores, and 1 contains one white spore and three red spores. Are ade7⁻ and ade1⁻ linked? **If so**, give the distance between them in cM.

h) In yeast, 1 cM of genetic distance translates to 3,500 bases of physical distance. Average yeast genes are about 1,400 bases long, and the longest yeast gene is 14,700 bases long. Given this information, can you conclude whether ade7⁻ and ade1⁻ are in the same gene or not? **If so,** state whether ade7⁻ and ade1⁻ are in the same gene or in different genes.

i) The diploid yeast created in part **g)** are white. Given this information, can you conclude whether ade7⁻ and ade1⁻ are in the same gene or not? **If so,** state whether ade7⁻ and ade1⁻ are in the same gene or in different genes.

j) Given all of the above information, can you conclude whether ade8⁻ and ade1⁻ are in the same gene or not? **If so,** state whether ade8⁻ and ade1⁻ are in the same gene or in different genes.

3. You have isolated two mutations in phage λ. One phage mutant is called **r1⁻** and gives a rough plaque phenotype (wild-type phage are **r1⁺** and give smooth plaques). The other phage mutant is called **mi⁻** and gives a mini plaque phenotype (wild-type phage are **mi⁺** and give large plaques). You cross **r1⁻** phage to **mi⁻** phage by coinfecting *E. coli*. When the resulting lysate is plated, you count 1000 of the resulting plaques and record the numbers of plaques of each morphology. You conclude that **r1⁻** and **mi⁻** are 15 m.u. away from each other.

a) List all of the categories of plaque morphology that you observed in the 1000 plaques you counted.

b) Estimate how many plaques of each morphology you observed in the 1000 plaques you counted.

You next cross **r1⁻** phage to **r2⁻** phage that also carry a mutation that gives rough plaques. When the resulting lysate is plated out and 1000 plaques are examined, only 5 are smooth.

c) The smooth plaques must have resulted from phage of the genotype **r1$^+$ r2$^+$**. List all possible genotypes of the phage that resulted in rough plaques.

d) What is the distance between the **r1$^-$** and **r2$^-$** mutations in m.u.?

e) Draw one map of this phage **for each** possible relative order of **mi$^-$**, **r1$^-$** and **r2$^-$**.

f) You next cross **mi$^-$ r1$^-$** phage to **r2$^-$** phage. When the resulting lysate is plated out and 1000 plaques are examined, you find that you have isolated 13 smooth plaques, all of which are large. Given this new information, draw a map of this phage showing the relative order of **mi$^-$**, **r1$^-$** and **r2$^-$**.

Solutions to 7.03 Problem Set 2

1. You are studying three autosomal mutations in the fruitfly *Drosophila melanogaster*. The cw^- mutation gives the recessive "curly-wing" phenotype (wild-type flies (cw^+) are straight-winged), the sh^- mutation gives the recessive "short-bristled" phenotype (wild-type flies (sh^+) are long-bristled), and the nh^-mutation gives the recessive "no-bristled" phenotype (wild-type flies (nh^+) are long-bristled).

a) You mate a homozygous nh^- cw^+ sh^- fly (which is no-bristled and straight-winged) to a true-breeding long-bristled and curly-winged fly to obtain an F1 generation. **List all** phenotypic categories of the flies in the F1 generation **and** the ratio in which these phenotypic categories are found in the F1 generation. The possible phenotypic classes are: short-bristled and curly-winged, long-bristled and straight-winged, short-bristled and straight-winged, long-bristled and curly-winged, no-bristled and curly-winged, no-bristled and straight-winged.

Since you are mating one homozygous fly to another homozygous fly, you would expect <u>only one</u> phenotypic category:

$$\frac{nh\text{-}\ cw+\ sh-}{nh-\ cw+\ sh-} \times \frac{nh+\ cw-\ sh+}{nh+\ cw-\ sh+} = ALL\ \frac{nh-\ cw+\ sh-}{nh+\ cw-\ sh+}$$

And since nh-, cw-, and sh- are all recessive mutations, all F1 flies will be <u>long-bristled and straight-winged.</u>

b) To do a proper three-factor cross, you need to create an F2 generation by mating F1 females to a specific true-breeding male.

i) What is the **phenotype** of the true-breeding male fly you would choose as the parent for the F2 generation?

To do a proper three-factor cross, we need to be able to score all recombinant classes of the F2 directly. We do this by crossing the F1 fly to a fly that is homozygous recessive for all three traits. This fly would be <u>no-bristled and curly-winged.</u>

ii) What is the **diploid genotype** of the true-breeding male you would choose as the parent for the F2 generation?

$$\frac{nh-\ cw-\ sh-}{nh-\ cw-\ sh-}$$ ***(reasoning in part i)***

You do a proper three-factor cross and, in the F2 generation, you obtain the following number of flies in each of the corresponding phenotypic classes:

short-bristled and curly-winged	(230 flies)
long-bristled and straight-winged	(10 flies)
short-bristled and straight-winged	(11 flies)
long-bristled and curly-winged	(238 flies)
no-bristled and curly-winged	(28 flies)
no-bristled and straight-winged	(483 flies)

c) All of the long-bristled and straight-winged flies have the alleles sh^+ nh^+cw^+ on the chromosome they inherited from their mother. Using this notation, list all of the possible chromosomes that the no-bristled and curly-winged flies could have inherited from their mother.

We cannot tell by looking whether a no-bristled, curly-winged fly inherited sh- or sh+ from the mother (there are no bristles to be designated "long" or "short"). Therefore, the possible chromosomes that it could have inherited are the following:

sh+ cw- nh- OR sh- cw- nh-

d) Are sh^- and nh^- linked? **If so**, give the distance between them in cM.

In the above three-factor cross, we crossed a long-bristled, straight-winged fly to a no-bristled, curly-winged fly with genotypes, respectively:

sh- cw+ nh-		***sh- cw- nh-***
----------	***x***	***----------***
sh+ cw- nh+		***sh- cw- nh-***

We expect half of the offspring to have bristles and other half to be bristle-less. If we assume that nh- and sh- are unlinked, then we would also expect that of those with bristles, half are long and half are short. Examining the data, this is roughly what we see ➔ 2:1:1 ratio of no-bristle:long-bristle:short-bristle. Therefore, nh- and sh- are unlinked.

Additionally, by examining the no-bristled progeny and estimating the hidden genotypes by matching to the reciprocal classes (reciprocal classes should appear at approximately the same frequency, see table on following page), you could calculate a genetic distance of ~50cM. By this method, you would still conclude that nh- and sh- are unlinked.

Phenotype	Genotype			#	Parental (P) or recombinant (R) for…		
	sh	cw	nh		sh-cw	sh-nh	cw-nh
Short-bristled, curly-winged	-	-	+	230	R	R	P
No-Bristled, straight-winged	+	+	-	~230	R	R	P
Long-bristled, straight-winged	+	+	+	10	R	P	R
No-bristled, curly-winged	-	-	-	~10	R	P	R
Short-bristled, straight-winged	-	+	+	11	P	R	R
No-bristled, curly-winged	+	-	-	~18	P	R	R
Long-bristled, curly-winged	+	-	+	238	P	P	P
No-bristled, straight-winged	-	+	-	~253	P	P	P

e) At this point, can you conclude whether sh⁻ and nh⁻ are in the same gene or not? **If so,** state whether sh⁻ and nh⁻ are in the same gene or in different genes.
If nh- and sh- were mutations in the same gene, you would expect them to be linked. Since we have determined nh- and sh- to be unlinked, we can conclude that they are mutations in different genes.

f) Are sh⁻ and cw⁻ linked? **If so**, give the distance between them in cM.
sh- and cw- are unlinked. By examining the no-bristled progeny and estimating the hidden genotypes by matching to the reciprocal classes, you could calculate a genetic distance between sh- and cw- to be ~50cM ➔ unlinked.

g) Are nh⁻ and cw⁻ linked? **If so**, give the distance between them in cM.
Since the parental classes are nh- cw+ and nh+ cw- , the recombinant classes are nh- cw- (no-bristled, curly-winged) and nh+ cw+ (bristled, straight-winged). Counting those classes, we get the following:

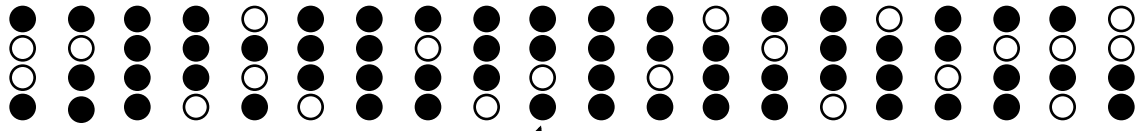
10 long-bristled, straight-winged
11 short-bristled, straight-winged
+ 28 no-bristled, curly-winged
---
49 recombinant gametes

Genetic distance = 100 x 49/1000 = 4.9 cM ➔ linked.

2. You find a haploid yeast strain in the freezer that is marked "double mutant strain which contains ade7$^-$ and ade8$^-$." Like some other ade mutants you have studied before, this double mutant strain is red in color, as compared to wild-type yeast, which are white. You cannot find any other information regarding this strain or the mutations it contains.

a) Mating of the ade7$^-$ ade8$^-$ double mutant strain to wild type produces diploids that are white. What do these results tell you about the ade7$^-$ and ade8$^-$ mutations?
This tells you that ade7- and ade8- are recessive mutations.

b) Sporulation of the diploids produced by the mating of ade7$^-$ ade8$^-$ double mutant yeast to wild-type yeast yields three different types of tetrads. Type One contains 4 red spores, Type Two contains 2 red spores and 2 white spores, and Type Three contains 3 red spores and 1 white spore. Below, all of the tetrads that you analyzed are drawn. (Each column below contains the four colonies that grew up from each tetrad you have analyzed. Dark colored circles indicate red yeast colonies, and white colored circles represent white yeast colonies.)



i) How many of the tetrads that you analyzed are parental ditypes?
PD = (2)ade7- ade8- spores and (2)ade7+ ade8+ spores. Therefore, we expect 2 white and 2 red. 4 PD tetrads.

ii) How many are tetratypes?
PD and NPD can not give a 3 to 1 phenotypic ratio (in this case 3 red : 1 white). There fore this must be a tetratype. TT = (1) ade7- ade8- spore, (1) ade7- ade8+ spore, (1) ade7+ ade8- spore, and (1) ade7+ ade8+ spore. 13 TT tetrads.

iii) How many are nonparental ditypes?
NPDs are the only remaining type of tetrad that we have not assigned. NPD (resulting from 2 crossovers) = (2) ade7- ade8+ spores and (2) ade7+ ade8- spores. 3 NPD tetrads.

c) Are ade7$^-$ and ade8$^-$ linked? **If so**, give the distance between them in cM.

ade7- and ade8- are <u>unlinked</u> because we see a ratio of approximately 1:4:1 of PD:TT:NPD (4:13:3).

d) At this point, can you conclude whether ade7$^-$ and ade8$^-$ are in the same gene or not? **If so,** state whether ade7$^-$ and ade8$^-$ are in the same gene or in different genes.
At this point, we can conclude that ade7- and ade8- are in different genes. If they were in the same gene, we would expect them to be linked. We established that they were unlinked in part c).

e) At this point, can you determine the single mutant phenotypes of ade7$^-$ yeast and ade8$^-$ yeast? **If so,** state whether ade7$^-$ single mutant yeast are red or white.
<u>Yes. Cells mutant in either ade7- or ade8- are red</u>. NPDs contain 4 spores, each containing one single mutant, and all four spores are red.

f) Below, the first two tetrads that you analyzed in part **(b)** are drawn. Some of the spores have their genotypes at the ade7 and ade8 loci written in for them. For the remaining spores (that is, those that are marked with question marks) give the genotypes at the ade7 and ade8 loci.

The first tetrad from part **(b)**:

- ● ? ***ade7-*** ***ade8-***
- ○ ? ***ade7+*** ***ade8+***
- ○ ? ***ade7+*** ***ade8+***
- ● ? ***ade7-*** ***ade8-***

The second tetrad from part **(b)**:

- ● ? ***ade7-*** ***ade8+***
- ○ ? ***ade7+*** ***ade8+***
- ● ade7$^-$ ade8$^-$
- ● ade7$^+$ ade8$^-$

g) You already had a mutant in your collection that gave a recessive red phenotype, called ade1$^-$. You mate ade1$^-$ yeast to ade7$^-$ yeast, and sporulate the resulting diploid. Out of 40 tetrads, 39 contain four red spores, and 1 contains one white spore and three red spores. Are ade7$^-$ and ade1$^-$ linked? **If so**, give the distance between them in cM.
In this mating, we get out only one tetratype. Since this does not fit the characteristic ratio of 1:4:1 of PD:TT:NPD (39:1:0 in this case) for unlinked mutations, we can say that ade1- and ade7- are linked.

Genetic distance = 100 x [(# tetratypes / 2) + 3 (# NPDs)] / total # tetrads
= 100 x [(1/2) + 3 (0)] / 40

= 100 x 1/80
= 100/80 = 1.25cM

h) In yeast, 1 cM of genetic distance translates to 3,500 bases of physical distance. Average yeast genes are about 1,400 bases long, and the longest yeast gene is 14,700 bases long. Given this information, can you conclude whether ade7$^-$ and ade1$^-$ are in the same gene or not? **If so,** state whether ade7$^-$ and ade1$^-$ are in the same gene or in different genes.
At this point, we cannot be certain. A genetic distance of 1.25cM between ade7- and ade1- corresponds to approximately 4,375 base pairs. Given the information above, it is possible that ade7- and ade1- are in the same gene.

i) The diploid yeast created in part **g)** are white. Given this information, can you conclude whether ade7$^-$ and ade1$^-$ are in the same gene or not? **If so,** state whether ade7$^-$ and ade1$^-$ are in the same gene or in different genes.
This is an example of the complementation test. Since the diploid is white and both ade1- and ade7- are recessive mutants, we can say that the two mutations complement each other and that they are in different genes.

j) Given all of the above information, can you conclude whether ade8$^-$ and ade1$^-$ are in the same gene or not? **If so,** state whether ade8$^-$ and ade1$^-$ are in the same gene or in different genes.
In the previous parts of this problem, we have determined 1) ade7- and ade8- are unlinked and 2) ade1- and ade7- are linked at 1.25 cM. We can conclude from this that ade1- and ade8- cannot be linked. Therefore, they are in different genes.

3. You have isolated two mutations in phage λ. One phage mutant is called **r1$^-$** and gives a rough plaque phenotype (wild-type phage are **r1$^+$** and give smooth plaques). The other phage mutant is called **mi$^-$** and gives a mini plaque phenotype (wild-type phage are **mi$^+$** and give large plaques). You cross **r1$^-$** phage to **mi$^-$** phage by coinfecting *E. coli*. When the resulting lysate is plated, you count 1000 of the resulting plaques and record the numbers of plaques of each morphology. You conclude that **r1$^-$** and **mi$^-$** are 15 m.u. away from each other.

a) List all of the categories of plaque morphology that you observed in the 1000 plaques you counted.

Here, we have the below cross:

-------I---15mu-------I------ x -------I---15mu--------I-------
mi- r1+ mi+ r1-

possible genotypes: ***mi- r1+ = small, smooth (parental)***
mi+ r1- = large, rough (parental)
mi- r1- = small, rough (recombinant)
mi+ r1+ = large, smooth (recombinant)

b) Estimate how many plaques of each morphology you observed in the 1000 plaques you counted.
Since we were given that r1- and mi- are 15 m.u. away, we can figure out how many recombinants we should expect.

15 = 100 x (# of recombinants)/1000 ➔ # of recombinant types = 150
# of parental types = 850

Reciprocal classes contain about the same number of members. Of the 150 recombinant types, we would expect 75 small, rough (mi- r1-) and 75 large, smooth (mi+ r1+). Similarly, of the 850 parental types, 425 are small, smooth (mi- r1+) and 425 large, rough (mi+ r1-).

You next cross **r1⁻** phage to **r2⁻** phage that also carry a mutation that gives rough plaques. When the resulting lysate is plated out and 1000 plaques are examined, only 5 are smooth.

c) The smooth plaques must have resulted from phage of the genotype **r1⁺ r2⁺**. List all possible genotypes of the phage that resulted in rough plaques.
All of the other possible genotypes would result in rough plaques:

r1- r2+ r1+ r2- r1- r2-

d) What is the distance between the **r1⁻** and **r2⁻** mutations in m.u.?
Since we have 5 smooth plaques (r1+ r2+), we need to take into account the reciprocal class of r1- r2-. So the total number of recombinants is 10 (5 r1+ r2+ and 5 r1- r2-).

Genetic Distance = 100 x 10/1000= <u>1 m.u.</u>

e) Draw one map of this phage **for each** possible relative order of **mi⁻**, **r1⁻** and **r2⁻**.
Known: distance from r1- to r2- = 1 m.u. and distance from r1- to mi- = 15 m.u. therefore, possible orders:

--mi----------r1--r2-- ***or*** ***--mi----------r2---r1***
15 m.u. 1 m.u. ***14 m.u. 1 m.u.***

f) You next cross **mi⁻ r1⁻** phage to **r2⁻** phage. When the resulting lysate is plated out and 1000 plaques are examined, you find that you have isolated 13 smooth plaques, all of which are large. Given this new information, draw a map of this phage showing the relative order of **mi⁻**, **r1⁻** and **r2⁻**.

Option 1 ***Option 2***
--m1$^-$----------r1$^-$--r2$^+$-- ***or*** ***--m1$^-$----------r2$^+$--r1$^-$--***
--m1$^+$----------r1$^+$--r2$^-$-- ***or*** ***--m1$^+$----------r2$^-$--r1$^+$--***
15 m.u. 1 m.u. ***14 m.u. 1 m.u.***

m 1	r1	r2	P or R	Phenotype		m1	r2	r1	P or R	Phenotype
+	+	-	P	Rough, Large		-	+	-	P	Rough, Small
-	-	+	P	Rough, Small		+	-	+	P	Rough, Large
+	-	+	R	Rough, Large		-	-	+	R	Rough, Small
-	+	-	R	Rough, Small		+	+	-	R	Rough, Large
+	+	+	R*	Smooth, Large		-	+	+	R*	Smooth, Small
-	-	-	R*	Rough, Small		+	-	-	R*	Rough, Large
+	-	-	R (d)	Rough, Large		-	-	-	R (d)	Rough, Small
-	+	+	R (d)	Smooth, Small		+	+	+	R (d)	Smooth, Large

R = X-over within the larger region, R = X-over within the 1 mu region, R (d) = double X-over*

We see 13/1000 smooth, large colonies. If option 2 is correct, then the smooth large colonies arise from a double cross-over event. If this is so, then you would expect to see many more than 13 colonies that are smooth and small because smooth and small plaques result from a single cross over and single crossovers are more frequent than double crossovers. Option 2 seems unlikely. If Option 1 is correct then the smooth large colonies arise from single cross-over events. Smooth, small colonies would arise from double cross-over events, which in this case do not appear because we only examined 1000 plaques. <u>Option 1 is a better fit for the data.</u>

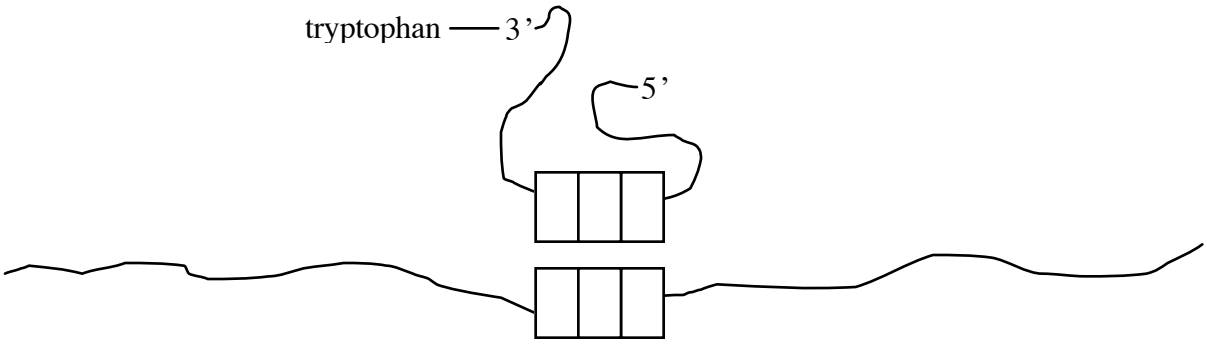
7.03 Problem Set 3

Due before 5 PM on Thursday, October 21, 2004
Hand in answers in recitation section or in the box outside of 68-120.
PLEASE WRITE YOUR ANSWERS ON THE PROBLEM SET!

1. You are studying a bacterial species whose genome normally contains one gene that encodes a tryptophan tRNA. The wild-type sequence of this gene, called *trnA-trp*, is shown below. The portion of the gene that encodes the anticodon of the tRNA is boxed.

5′GTACCTGCACTGCATGCCTAGCTAGCCCTAG**CCA**GCCTAGCTAGCTAGCACCAA3′
3′CATGGACGTGACGTACGGATCGATCGGGATC**GGT**CGGATCGATCGATCGTGGTT5′

(a) Below is drawn the folded tryptophan-tRNA (produced from the *trnA-trp* gene) base-pairing with an mRNA containing the codon that this tryptophan-tRNA recognizes. **Fill in** the three boxes on the tRNA with the correct nucleotide sequence of its anticodon. Then **fill in** the three boxes on the mRNA (that is in the process of being translated) with the correct nucleotide sequence of the codon currently being read. Be sure to **label the ends** of the mRNA to show directionality.



(b) Which strand of the double-stranded *trnA-trp* gene shown at the top of the page is used as a template when the tryptophan tRNA is transcribed, the upper strand or the lower strand? Remember that tRNAs are transcribed directly from genes – there is no mRNA intermediate made during the production of a tRNA from its DNA sequence.

(c) You isolate a mutation in the *trnA-trp* gene which you call *trnA-trp*.* *trnA-trp** encodes a suppressor tRNA, and its sequence is shown below, with the mutation underlined. **Draw** the mutant tRNA that will be transcribed from the *trnA-trp** gene, using the drawing of the wild-type tryptophan trnA from part **(a)** as a model. Be sure to write out the sequence of the anticodon, label the ends of the tRNA to show directionality, and include the amino acid to which this tRNA would be covalently bound.

5′GTACCTGCACTGCATGCCTAGCTAGCCCTAGTCAGCCTAGCTAGCTAGCACCAA3′
3′CATGGACGTGACGTACGGATCGATCGGGATCAGTCGGATCGATCGATCGTGGTT5′

(d) Which stop codon (5'-UAG-3', 5'-UGA-3', or 5'-UAA-3') in an mRNA will be recognized by the mutant tRNA produced from the *trnA-trp** gene?

You are working on another project in which you are studying the gene that gives this bacterial species its wild-type "shiny" colony morphology; you name the gene *dulA*. The small polypeptide DulA is the product of the *dulA* gene. You isolate four mutations in the *dulA* gene, and you call the mutations *dulA1*, *dulA2*, *dulA3,* and *dulA4.* Bacterial strains containing one of the mutations *dulA1*, *dulA3* or *dulA4* have the mutant "dull" colony morphology. You are particularly intrigued by the conditional mutant *dulA2* because *dulA2* only gives a dull colony phenotype when in a *trnA-trp** background (that is, the *dulA2* single mutant strain is shiny like wild-type).

You sequence the different wild-type and mutant forms of the *dulA* gene and deduce the following information about the sequence of the *dulA* mRNA. The mutations are underlined.

The mRNA produced from the wild-type *dulA* gene:
5′GAACUAUGGGAAUACCGUACUCAAUCUGCCGUAUCUAAUUGCUAUAAACG3′

The mRNA produced from the mutant *dulA1* gene:
5′GAACUAUGGGAAUACCGUACUAAAUCUGCCGUAUCUAAUUGCUAUAAACG3′

The mRNA produced from the mutant *dulA2* gene:
5′GAACUAUGGGAAUACCGUACUCAAUCUGCCGUAUCUGAUUGCUAUAAACG3′

The mRNA produced from the mutant *dulA3* gene:
5′GAACUAUGGGAAUACCGUACUCAAUCUGCCGUAUCCAAUUGCUAUAAACG3′

The mRNA produced from the mutant *dulA4* gene:
5′GAACUAUGGGAAUACCGUACUCAAUCUGACGUAUCUAAUUGCUAUAAACG3′

(e) Fill in the chart below for each of the following bacterial strains by predicting how many amino acids long each form of the DulA polypeptide will be, and what the colony phenotype of each strain will be. Some boxes are already filled in for you correctly as examples. [NOTE: Assume that the suppressor tRNA is **100% efficient** when filling in this chart (but keep in mind that suppressor tRNAs are actually rather inefficient).]

Genotype of the bacterial strain	Number of amino acids found in the protein product of the *dulA* gene in this strain	Phenotype of the bacterial colonies (dull, shiny, or "cannot conclude")
dulA trnA-trp (wild-type)	10	shiny
*dulA trnA-trp** (single mutant)		
dulA1 trnA-trp (single mutant)		
*dulA1 trnA-trp** (double mutant)		
dulA2 trnA-trp (single mutant)		shiny
*dulA2 trnA-trp** (double mutant)		dull
dulA3 trnA-trp (single mutant)		
*dulA3 trnA-trp** (double mutant)		
dulA4 trnA-trp (single mutant)		
*dulA4 trnA-trp** (double mutant)		

2. You are studying the ability of a certain bacterial species to form biofilms, which are thick layers of bacteria that can adhere tightly to and colonize a surface. You isolate a mutation, *flm1*, which results in the inability to produce of biofilms.

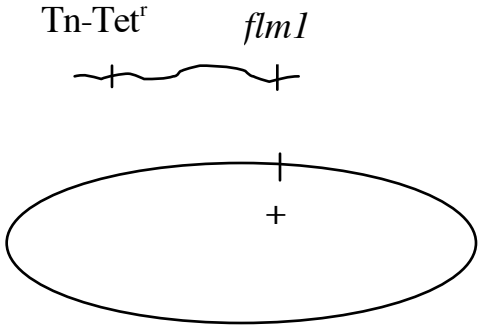
You cannot select for the ability or inability to produce biofilms; you can only screen for it. Thus you decide to isolate a transposon insertion that is near to the *flm1* mutation and that carries a gene allowing for tetracycline resistance (Tet^r). You are given three candidate transposon insertions (Insertion One, Insertion Two, and Insertion Three) by your advisor, and you create three strains that each have a single transposon insertion and the *flm1* mutation.

For your *flm1* strain containing Insertion One, you grow P1 phage on this *flm1* Tet^r strain and use the resulting phage lysate to infect a wild-type strain, selecting for tetracycline resistance (Tet^r). Among 100 Tet^r transductants, you find that 75 can form biofilms and 25 cannot form biofilms.

For your *flm1* strain containing Insertion Two, you grow P1 phage on this *flm1* Tet^r strain and use the resulting phage lysate to infect a wild-type strain, selecting for tetracycline resistance (Tet^r). Among 100 Tet^r transductants, you find that 25 can form biofilms and 75 cannot form biofilms.

For your *flm1* strain containing Insertion Three, you grow P1 phage on this *flm1* Tet^r strain and use the resulting phage lysate to infect a wild-type strain, selecting for tetracycline resistance (Tet^r). Among 100 Tet^r transductants, you find that none can form biofilms.

(a) For Insertion One, we have drawn the entering DNA from the phage transduction lining up and recombining with the homologous portion of the bacterial chromosome in the recipient cell. Draw in the set of crossover events that result in the cotransduction of Tet^r and *flm1*.



(b) Express the distance between Insertion One and the *flm1* mutation as a cotransduction frequency.

(c) Express the distance between Insertion Three and the *flm1* mutation as a cotransduction frequency.

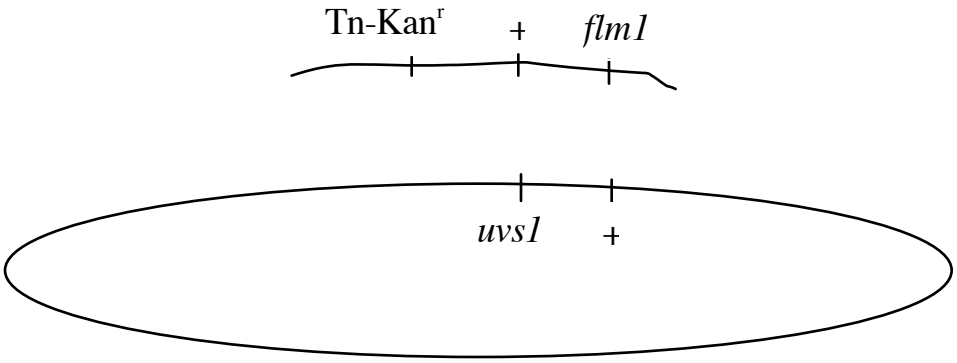
(d) Which insertion is physically located closer to *flm1* -- Insertion One or Insertion Two?

(e) You examine the strain carrying *flm1* and Insertion three more closely. You grow P1 phage on this strain and use the resulting phage lysate to infect a wild-type strain. After selecting for tetracycline resistance (Tet^r), you now screen 1000 Tet^r transductants, and you find that none can form biofilms. Propose two possible explanations for this result.

You isolate another mutation, *uvs1*, that results in sensitivity to UV irradiation. You also isolate a bacterial strain containing a transposon insertion that carries a gene allowing for kanamycin resistance. In this strain, the kanamycin resistance gene is cotransduced with *uvs1* with a cotransduction frequency of 50%. Preliminary P1 transduction experiments indicate that *flm1* is linked to this same Kan^r transposon insertion. You set up the following cross to map the *flm* locus relative to the *uvs* locus:

You grow P1 phage on a Kan^r strain that contains this transposon insertion and the *flm1* mutation, and use the resulting phage lysate to infect a *uvs1* strain. You select for kanamycin resistance (Kan^r), and among 100 Kan^r transductants, you find that 39 can form UV-resistant biofilms, 55 can form UV-sensitive biofilms, and 6 cannot form biofilms and are UV-resistant.

(f) Using the **putative** order in which ***uvs* is closer** to the transposon than *flm*, we have drawn the entering DNA from the phage transduction lining up and recombining with the homologous portion of the bacterial chromosome. Draw in the set of crossover events that yield Kan^r bacteria that are UV-resistant and can form biofilms.



(g) Following the model from part **(f)**, draw the entering DNA from the phage transduction lining up and recombining with the homologous portion of the bacterial chromosome for the **putative** order in which ***uvs* is closer** to the transposon than *flm*. Draw in the set of crossover events that yield Kan^r bacteria that are UV-sensitive and cannot form biofilms.

(h) Following the model from part **(f)**, draw the entering DNA from the phage transduction lining up and recombining with the homologous portion of the bacterial chromosome for the **putative** order in which ***flm* is closer** to the transposon than *uvs* (ie. the *flm* locus is in between the transposon and the *uvs* locus). Draw in the set of crossover events that yield Kan^r bacteria that are UV-resistant and can form biofilms.

(i) Following the model from part **(f)**, draw the entering DNA from the phage transduction lining up and recombining with the homologous portion of the bacterial chromosome for the **putative** order in which ***flm* is closer** to the transposon than *uvs*. Draw in the set of crossover events that yield Kan^r bacteria that are UV-sensitive and cannot form biofilms.

(j) Which putative gene order is more likely to be correct? Your choices are:

flm is closer to the transposon than *uvs* OR *uvs* is closer to the transposon than *flm*

You isolate a triple mutant strain which you call *uvs1 sup* Kan^r. This triple mutant contains your Kan^r encoding transposon insertion linked to the *uvs1* mutation, but this strain is now UV-resistant because of the *sup* mutation. You grow P1 phage on this *uvs1 sup* Kan^r strain and use the resulting phage lysate to infect a *sup* single mutant strain, selecting for kanamycin resistance (Kan^r). Among 100 Kan^r transductants, you find that all 100 are UV-resistant.

(k) Can you conclude if *sup* is linked to *uvs1*? **If so,** state whether they are very tightly linked, loosely linked, or unlinked.

You grow P1 phage on the *uvs1 sup* Kan^r triple mutant strain and use the resulting phage lysate to infect a *uvs1* single mutant strain, selecting for kanamycin resistance (Kan^r). Among 100 Kan^r transductants, you find that all 100 are UV-sensitive.

(l) Can you conclude if *sup* is an intragenic suppressor or an extragenic suppressor of *uvs1*? **If so**, state whether it is intragenic or extragenic.

3. You isolate a single mutation (*genA1*) in an *E. coli* gene, *genA*. *E. coli* that contain the *genA1* mutation have the following properties:

P1 phage grown on an *E. coli* strain that carries a Tn5 insertion in the chromosome (Tn5 has a gene for kanamycin resistance) will transduce wild type *E. coli* to kanamycin resistance at a frequency of about 10^{-4}, whereas *genA1* mutants are transduced to kanamycin resistance at a frequency less than 10^{-9}.

An Hfr strain that contains a Tn5 insertion as an early marker when mated to wild type *E. coli* will produce kanamycin resistant exconjugants at a frequency of about 10^{-2} , whereas a *genA1* mutant will become kanamycin resistance at a frequency less than 10^{-9}.

An *E. coli* donor stain that carries an F' plasmid carrying a Tn5 insertion when mated to either wild type or a *genA1* mutant will produce kanamycin resistant exconjugants at the same frequency (of about 10^{-1}).

(a) In what cellular process does the protein product of the *genA* gene function?

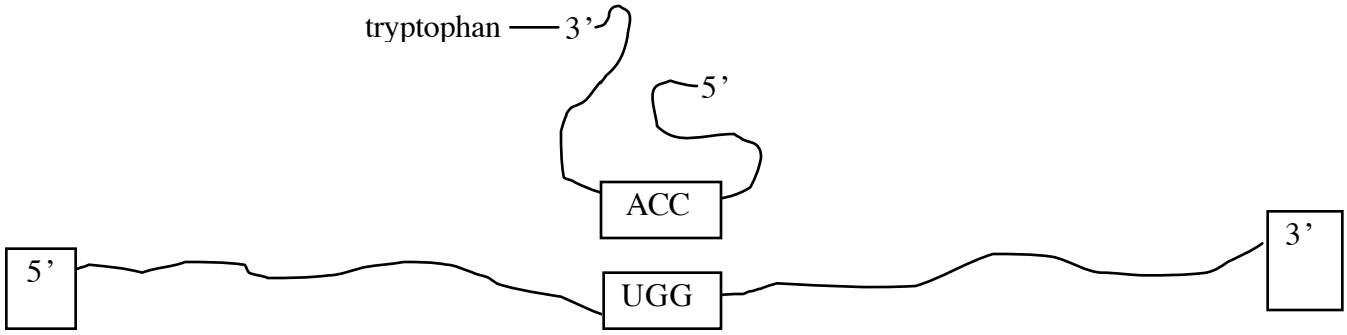
(b) You find that *genA1 E. coli* can become kanamycin resistant when transduced with int-P_{amber} λ phage that carries Tn5 at the same frequency as wild type(~10^5). In one sentence, what does this result tell you about the process of transposition given what you know about the function of the *genA* gene?

Solutions to 7.03 Problem Set 3

1. You are studying a bacterial species whose genome normally contains one gene that encodes a tryptophan tRNA. The wild-type sequence of this gene, called *trnA-trp*, is shown below. The portion of the gene that encodes the anticodon of the tRNA is boxed.

5′GTACCTGCACTGCATGCCTAGCTAGCCCTAGCCAGCCTAGCTAGCTAGCACCAA3′
3′CATGGACGTGACGTACGGATCGATCGGGATCGGTCGGATCGATCGATCGTGGTT5′

(a) Below is drawn the folded tryptophan-tRNA (produced from the *trnA-trp* gene) base-pairing with an mRNA containing the codon that this tryptophan-tRNA recognizes. **Fill in** the three boxes on the tRNA with the correct nucleotide sequence of its anticodon. Then **fill in** the three boxes on the mRNA (that is in the process of being translated) with the correct nucleotide sequence of the codon currently being read. Be sure to **label the ends** of the mRNA to show directionality.



5′ UGG 3′ is the codon for tryptophan. The anticodon 5′ CCA 3′ in the tRNA recognizes this sequence in the mRNA.

(b) Which strand of the double-stranded *trnA-trp* gene shown at the top of the page is used as a template when the tryptophan tRNA is transcribed, the upper strand or the lower strand? Remember that tRNAs are transcribed directly from genes – there is no mRNA intermediate made during the production of a tRNA from its DNA sequence.

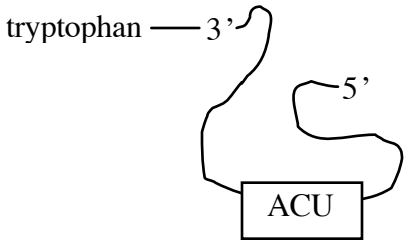
The lower strand is used as the template strand. When transcribed, the sequence for the anticodon in the tRNA should be 5′CCA3′. Therefore, the lower strand must be used as the template:

3′ …GGT… 5′ used as template
5′ …CCA… 3′ tRNA

(c) You isolate a mutation in the *trnA-trp* gene which you call *trnA-trp**. *trnA-trp** encodes a suppressor tRNA, and its sequence is shown below, with the mutation underlined. **Draw** the mutant tRNA that will be transcribed from the *trnA-trp** gene, using the drawing of the wild-type tryptophan trnA from part **(a)** as a model. Be sure to write out the sequence of the anticodon, label the ends of the tRNA to show directionality, and include the amino acid to which this tRNA would be covalently bound.

5′GTACCTGCACTGCATGCCTAGCTAGCCCTAGTCAGCCTAGCTAGCTAGCACCAA3′
3′CATGGACGTGACGTACGGATCGATCGGGATCAGTCGGATCGATCGATCGTGGTT5′

The mutation changed the sequence encoding the anticodon from 3′ *GG*T 5′ to 3′ *AG*T 5′. The mutation in the sequence for the anticodon won't alter the amino acid with which the tRNA is charged.



(d) Which stop codon (5'-UAG-3', 5'-UGA-3', or 5'-UAA-3') in an mRNA will be recognized by the mutant tRNA produced from the *trnA-trp** gene?

The anticodon of the mutant tRNA, 3′ *ACU* 5′ will base pair with the mRNA sequence 5′ *UGA* 3′, mistakenly recognizing that stop codon.

You are working on another project in which you are studying the gene that gives this bacterial species its wild-type "shiny" colony morphology; you name the gene *dulA*. The small polypeptide DulA is the product of the *dulA* gene. You isolate four mutations in the *dulA* gene, and you call the mutations *dulA1*, *dulA2*, *dulA3,* and *dulA4.* Bacterial strains containing one of the mutations *dulA1*, *dulA3* or *dulA4* have the mutant "dull" colony morphology. You are particularly intrigued by the conditional mutant *dulA2* because *dulA2* only gives a dull colony phenotype when in a *trnA-trp** background (that is, the *dulA2* single mutant strain is shiny like wild-type).

You sequence the different wild-type and mutant forms of the *dulA* gene and deduce the following information about the sequence of the *dulA* mRNA. The mutations are underlined.

The mRNA produced from the wild-type *dulA* gene:

5′GAACU AUG GGA AUA CCG UAC UCA AUC UGC CGU AUC UAA UUGCUAUAAACG3′

1-2. Met Gly Ile Pro Tyr Ser Ile Cys Arg Ile stop

The mRNA produced from the mutant *dulA1* gene:

5′GAACU AUG GGA AUA CCG UAC U<u>A</u>A AUCUGCCGUAUCUAAUUGCUAUAAACG3′

3-4. Met Gly Ile Pro Tyr stop

The mRNA produced from the mutant *dulA2* gene:

5′GAACU AUG GGA AUA CCG UAC UCA AUC UGC CGU AUC U<u>G</u>A UUG CUA UAA ACG3′

5. Met Gly Ile Pro Tyr Ser Ile Cys Arg Ile stop
6. Met Gly Ile Pro Tyr Ser Ile Cys Arg Ile Trp Leu Leu stop

The mRNA produced from the mutant *dulA3* gene:

5′GAACU AUG GGA AUA CCG UAC UCA AUC UGC CGU AUC <u>C</u>AA UUG CUA UAA ACG3′

7-8. Met Gly Ile Pro Tyr Ser Ile Cys Arg Ile Gln Leu Leu stop

The mRNA produced from the mutant *dulA4* gene:

5′GAACU AUG GGA AUA CCG UAC UCA AUC UG<u>A</u> CGU AUC UAA UUGCUAUAAACG3′

9. Met Gly Ile Pro Tyr Ser Ile stop
10. Met Gly Ile Pro Tyr Ser Ile Trp Arg Ile stop

(e) Fill in the chart below for each of the following bacterial strains by predicting how many amino acids long each form of the DulA polypeptide will be, and what the colony phenotype of each strain will be. Some boxes are already filled in for you correctly as examples. [<u>NOTE</u>: Assume that the suppressor tRNA is **100% efficient** when filling in this chart (but keep in mind that suppressor tRNAs are actually rather inefficient).]

Each line of the table is labeled with a number. The protein product for each number is shown on the previous page. The intro to the problem states that dulA1, dulA3, and dulA4 give the mutant "dull" colony morphology. Since trnA-trp* has no effect on the protein product formed by dulA1 and dulA3, those double mutants should show the dull morphology. For the dulA4 mutant, the trnA-trp* will restore the wild type length of the protein product, but will cause a Cys → Trp mutation, the effect of which we cannot predict. Therefore, the morphology of this double mutant is unknown.

	Genotype of the bacterial strain	Number of amino acids found in the protein product of the *dulA* gene in this strain	Phenotype of the bacterial colonies (dull, shiny, or "cannot conclude")
1	*dulA trnA-trp* (wild-type)	10	shiny
2	*dulA trnA-trp** (single mutant)	10	shiny
3	*dulA1 trnA-trp* (single mutant)	5	dull
4	*dulA1 trnA-trp** (double mutant)	5	dull
5	*dulA2 trnA-trp* (single mutant)	10	shiny
6	*dulA2 trnA-trp** (double mutant)	13	dull
7	*dulA3 trnA-trp* (single mutant)	13	dull
8	*dulA3 trnA-trp** (double mutant)	13	dull
9	*dulA4 trnA-trp* (single mutant)	7	dull
10	*dulA4 trnA-trp** (double mutant)	10	Can't conclude

2. You are studying the ability of a certain bacterial species to form biofilms, which are thick layers of bacteria that can adhere tightly to and colonize a surface. You isolate a mutation, *flm1*, which results in the inability to produce of biofilms.

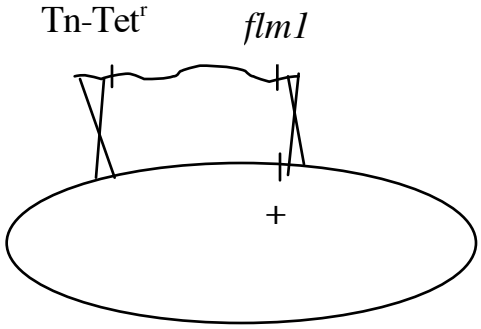
You cannot select for the ability or inability to produce biofilms; you can only screen for it. Thus you decide to isolate a transposon insertion that is near to the *flm1* mutation and that carries a gene allowing for tetracycline resistance (Tet^r). You are given three candidate transposon insertions (Insertion One, Insertion Two, and Insertion Three) by your advisor, and you create three strains that each have a single transposon insertion and the *flm1* mutation.

For your *flm1* strain containing Insertion One, you grow P1 phage on this *flm1* Tet^r strain and use the resulting phage lysate to infect a wild-type strain, selecting for tetracycline resistance (Tet^r). Among 100 Tet^r transductants, you find that 75 can form biofilms and 25 cannot form biofilms.

For your *flm1* strain containing Insertion Two, you grow P1 phage on this *flm1* Tet^r strain and use the resulting phage lysate to infect a wild-type strain, selecting for tetracycline resistance (Tet^r). Among 100 Tet^r transductants, you find that 25 can form biofilms and 75 cannot form biofilms.

For your *flm1* strain containing Insertion Three, you grow P1 phage on this *flm1* Tet^r strain and use the resulting phage lysate to infect a wild-type strain, selecting for tetracycline resistance (Tet^r). Among 100 Tet^r transductants, you find that none can form biofilms.

(a) For Insertion One, we have drawn the entering DNA from the phage transduction lining up and recombining with the homologous portion of the bacterial chromosome in the recipient cell. Draw in the set of crossover events that result in the cotransduction of Tet^r and *flm1*.



For both Tet^r and flm 1 to be brought into the chromosome, or cotransduced, the crossover events must flank them, as shown.

(b) Express the distance between Insertion One and the *flm1* mutation as a cotransduction frequency.

When Tn-Tet^r and flm1 are cotransduced as shown in part (a), the tetracycline-resistant transductants will contain the flm1 mutation, and won't be able to form biofilms. In the intro about Insertion One, we're told that of 100 Tet^r transductants, 25 cannot form biofilms. Therefore Insertion One and flm1 are cotransduced at a frequency of 25%.

(c) Express the distance between Insertion Three and the *flm1* mutation as a cotransduction frequency.

If Insertion Three and flm1 are cotransduced, Tet^r transductants should contain the flm1 mutation and should be unable to form biofilms. The intro states that of 100 Tet^r transductants, none can form biofilms. Therefore Insertion Three and flm1 are cotransduced at a frequency of 100%.

(d) Which insertion is physically located closer to *flm1* -- Insertion One or Insertion Two?

For Insertion Two, the cotransduction frequency with Tet^r is 75%, whereas for Insertion One, the cotransduction frequency (calculated in part (b)) was only 25%. Since Insertion Two and flm1 are cotransduced more often, they must be closer together than Insertion One and flm1.

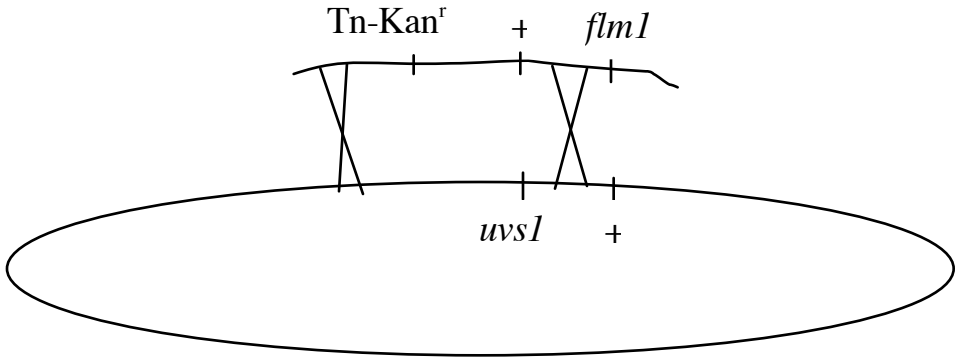
(e) You examine the strain carrying *flm1* and Insertion three more closely. You grow P1 phage on this strain and use the resulting phage lysate to infect a wild-type strain. After selecting for tetracycline resistance (Tet^r), you now screen 1000 Tet^r transductants, and you find that none can form biofilms. Propose two possible explanations for this result.

1. The transposon could've inserted directly into flm, disrupting the gene. Then all Tet^r transductants would be unable to form biofilms.
2. The transposon could've inserted into and disrupted another gene required for biofilm formation.
3. The transposon insertion could be so close to the flm 1 mutation that among 1000 transductants, none were the result of a crossover between flm1 and the transposon.

You isolate another mutation, *uvs1*, that results in sensitivity to UV irradiation. You also isolate a bacterial strain containing a transposon insertion that carries a gene allowing for kanamycin resistance. In this strain, the kanamycin resistance gene is cotransduced with *uvs1* with a cotransduction frequency of 50%. Preliminary P1 transduction experiments indicate that *flm1* is linked to this same Kan^r transposon insertion. You set up the following cross to map the *flm* locus relative to the *uvs* locus:

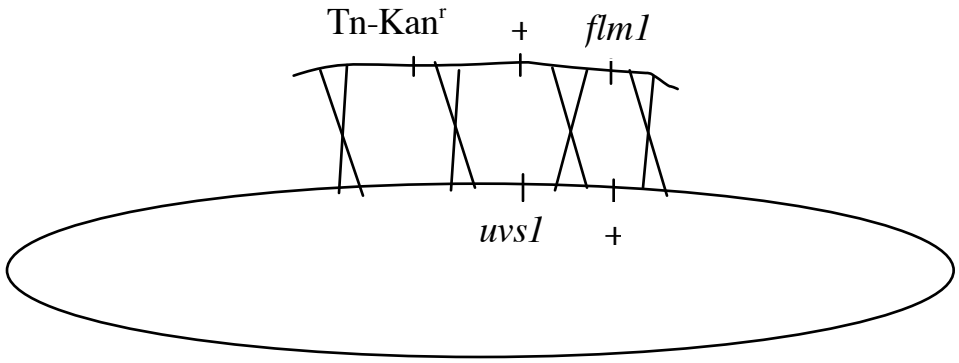
You grow P1 phage on a Kan^r strain that contains this transposon insertion and the *flm1* mutation, and use the resulting phage lysate to infect a *uvs1* strain. You select for kanamycin resistance (Kan^r), and among 100 Kan^r transductants, you find that 39 can form UV-resistant biofilms, 55 can form UV-sensitive biofilms, and 6 cannot form biofilms and are UV-resistant.

(f) Using the **putative** order in which ***uvs* is closer** to the transposon than *flm*, we have drawn the entering DNA from the phage transduction lining up and recombining with the homologous portion of the bacterial chromosome. Draw in the set of crossover events that yield Kan^r bacteria that are UV-resistant and can form biofilms.



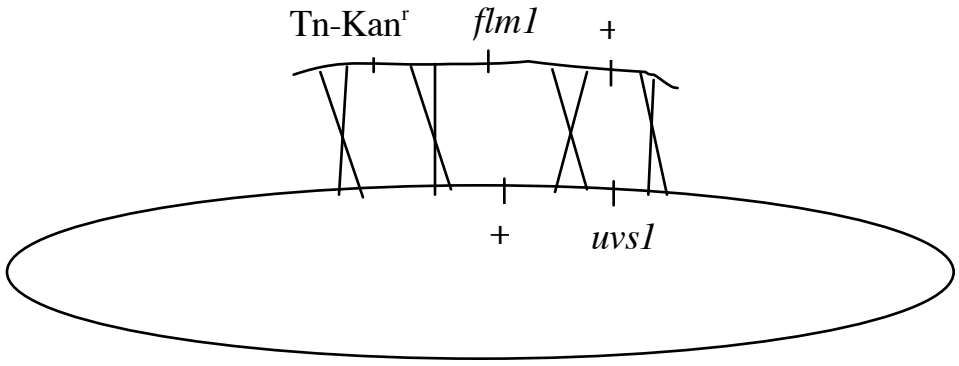
In order for the transductants to be Kan^r, UV-resistant, and able to form biofilms, Tn-Kan^r and wild type uvs$^+$ must recombine into the chromosome, but the flm1 mutation must not. This would happen by the crossovers shown. The resulting bacterial chromosome would have the genotype: Tn- Kan^r uvs$^+$ flm$^+$.

(g) Following the model from part **(f)**, draw the entering DNA from the phage transduction lining up and recombining with the homologous portion of the bacterial chromosome for the **putative** order in which ***uvs* is closer** to the transposon than *flm*. Draw in the set of crossover events that yield Kan^r bacteria that are UV-sensitive and cannot form biofilms.



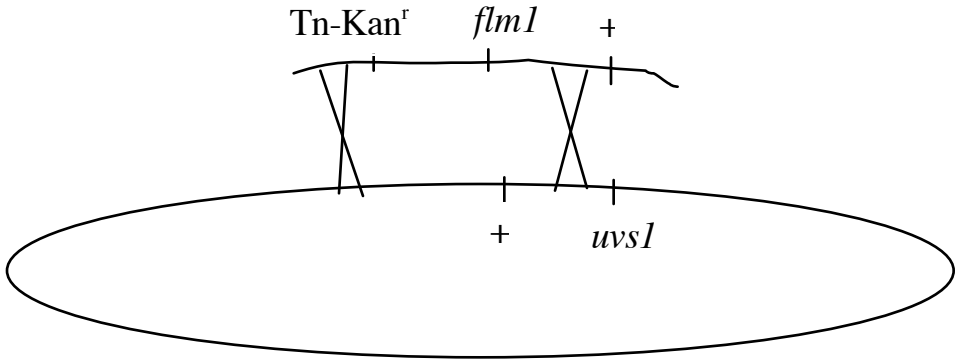
In order for the transductants to be Kan^r, UV-sensitive and unable to form biofilms, Tn-Kan^r and flm1, but not wild type uvs, must recombine into the chromosome, as shown. Thus this class results from a quadruple crossover event.

(h) Following the model from part **(f)**, draw the entering DNA from the phage transduction lining up and recombining with the homologous portion of the bacterial chromosome for the **putative** order in which ***flm* is closer** to the transposon than *uvs* (ie. the *flm* locus is in between the transposon and the *uvs* locus). Draw in the set of crossover events that yield Kan^r bacteria that are UV-resistant and can form biofilms.



In order to be Kan^r, UV-resistant, and able to form biofilms, Tn- Kan^r and wildtype uvs$^+$, but not the flm1 mutation, must recombine in to the chromosome.

(i) Following the model from part **(f)**, draw the entering DNA from the phage transduction lining up and recombining with the homologous portion of the bacterial chromosome for the **putative** order in which ***flm* is closer** to the transposon than *uvs*. Draw in the set of crossover events that yield Kan^r bacteria that are UV-sensitive and cannot form biofilms.



In order for the transductants to be Kan^r, UV-sensitive, and unable to form biofilms, Tn-Kan^r and flm1, but not wild type uvs$^+$ must recombine into the chromosome.

(j) Which putative gene order is more likely to be correct? Your choices are:

flm is closer to the transposon than *uvs* OR (*uvs* is closer to the transposon than *flm*)

According to the introduction to the problem, among 100 Kan^r transductants,

39 can form UV-resistant biofilms (uvs+, flm+)

55 can form UV-sensitive biofilms (uvs1, flm+)

6 are UV-resistant and cannot form biofilms (uvs+, flm1)

Quadruple crossover (QC) events are always less likely than the double crossover (DC) events.

If uvs is closer to the transposon, then the QC class- uvs1 flm1 Kan^r (see part g) should occur less frequently than DC class- uvs+ flm+ Kan^r (see part f).

If flm is closer to the transposon, then the QC class- uvs+ flm+ Kan^r (see part h) should occur less frequently than DC class- uvs1 flm1 Kan^r (see part i).

You isolate a triple mutant strain which you call *uvs1 sup* Kan^r. This triple mutant contains your Kan^r encoding transposon insertion linked to the *uvs1* mutation, but this strain is now UV-resistant because of the *sup* mutation. You grow P1 phage on this *uvs1 sup* Kan^r strain and use the resulting phage lysate to infect a *sup* single mutant strain, selecting for kanamycin resistance (Kan^r). Among 100 Kan^r transductants, you find that all 100 are UV-resistant.

(k) Can you conclude if *sup* is linked to *uvs1*? **If so,** state whether they are very tightly linked, loosely linked, or unlinked.
Since the donor and recipient strain both have the sup mutation, you cannot track its cotransduction frequency with Kan^r. Thus, you cannot conclude whether sup is linked to uvs1.

You grow P1 phage on the *uvs1 sup* Kan^r triple mutant strain and use the resulting phage lysate to infect a *uvs1* single mutant strain, selecting for kanamycin resistance (Kan^r). Among 100 Kan^r transductants, you find that all 100 are UV-sensitive.

(l) Can you conclude if *sup* is an intragenic suppressor or an extragenic suppressor of *uvs1*? **If so**, state whether it is intragenic or extragenic.
Because all the transductants are UV-sensitive, you can conclude that the sup mutation was never cotransduced with Kan^r (unlike uvs1 which was stated as being linked to Tn- Kan^r). Therefore, uvs1 and sup must be in different genes; the suppressor is extragenic.

3. You isolate a single mutation (*genA1*) in an *E. coli* gene, *genA*. *E. coli* that contain the *genA1* mutation have the following properties:

P1 phage grown on an *E. coli* strain that carries a Tn5 insertion in the chromosome (Tn5 has a gene for kanamycin resistance) will transduce wild type *E. coli* to kanamycin resistance at a frequency of about 10^{-4}, whereas *genA1* mutants are transduced to kanamycin resistance at a frequency less than 10^{-9}.

An Hfr strain that contains a Tn5 insertion as an early marker when mated to wild type *E. coli* will produce kanamycin resistant exconjugants at a frequency of about 10^{-2}, whereas a *genA1* mutant will become kanamycin resistance at a frequency less than 10^{-9}.

An *E. coli* donor stain that carries an F' plasmid carrying a Tn5 insertion when mated to either wild type or a *genA1* mutant will produce kanamycin resistant exconjugants at the same frequency (of about 10^{-1}).

(a) In what cellular process does the protein product of the *genA* gene function?

The gen A1 mutation has no effect on the frequency with which kanamycin-resistant exconjugants are formed when the Tn5 is brought into the recipient cell on an F' plasmid. In contrast, the gen A1 mutation does affect the number of Kan^r exconjugants obtained when the transposon is carried in by P1 phage or an Hfr. The F' is distinguished from the others in that it does not need to integrate into the chromosome of the recipient cell to be stably maintained. *E. Coli* chromosomal DNA from P1 phage and Hfrs must integrate by homologous recombination in the recipient to be stably maintained. If the gen A1 mutation disrupted homologous recombination, this would account for the data.

(b) You find that *genA1 E. coli* can become kanamycin resistant when transduced with int-P_{amber} λ phage that carries Tn5 at the same frequency as wild type(~10^5). In one sentence, what does this result tell you about the process of transposition given what you know about the function of the *genA* gene?

The int-P_{amber} λ phage can't replicate and it can't integrate into the chromosome of a recipient cell. Thus the only way the kanamycin resistance gene can be transferred to the chromosome is by transposition. Transposition requires transposase, an enzyme normally encoded in the transposon itself, and does not require host machinery (*i.e.* resident *E. Coli* proteins). Therefore, transposition is not affected by the genA1 mutation.

7.03 Problem Set 4

Due before 5 PM on Thursday, October 28, 2004
Hand in answers in recitation section or in the box outside of 68-120

1. In a genetic screen using an uncharacterized bacterial species, you identify a gene called *osmA,* which is induced by osmotic stress. You do not have an easy way of detecting OsmA protein or its activity. Therefore, you use a version of the *lacZ* gene (which encodes the enzyme ß -galactosidase) that lacks its own promoter as a reporter for *osmA* expression. You construct a strain in which *lacZ* is inserted into the *osmA* gene so that *lacZ* is under the control of the *osmA* promoter and other regulatory DNA sequences (P_{osmA}-*lacZ*). The result is an engineered bacterial strain that expresses ß-galactosidase from the P_{osmA}-*lacZ* construct whenever OsmA should normally be produced.

	Expression of the P_{osmA}-*lacZ* construct	
	no osmotic stress	osmotic stress
P_{osmA}-*lacZ*	–	+

You now do a second genetic screen in which you look for mutations that disrupt the normal regulation of your P_{osmA}-*lacZ* fusion construct. In this screen, you isolate two mutations, which you call *osmB1* and *osmC1*. You find that both loci (*osmB* and *osmC*) are closely linked to P_{osmA}-*lacZ*. Making use of F' plasmids that carry the *osm* genes, you construct P_{osmA}-*lacZ* strains that you test for ß-galactosidase activity with results shown below.

	Expression of the P_{osmA}-*lacZ* construct	
	no osmotic stress	osmotic stress
P_{osmA}-*lacZ*	-	+
osmB1 P_{osmA}-*lacZ*	+	+
osmC1 P_{osmA}-*lacZ*	-	-
osmB$^+$/ F' *osmB1* P_{osmA}-*lacZ*	+	+
osmC$^+$/ F' *osmC1* P_{osmA}-*lacZ*	-	-
osmB$^+$ P_{osmA}-*lacZ* / F' *osmB1*	-	+
osmB1 P_{osmA}-*lacZ* / F' *osmB*$^+$	+	+
osmC1 P_{osmA}-*lacZ* / F' *osmC*$^+$	-	-
osmC$^+$ P_{osmA}-*lacZ* / F' *osmC1*	-	-

You next construct an *osmB1 osmC1* double mutant, which gives the following behavior:

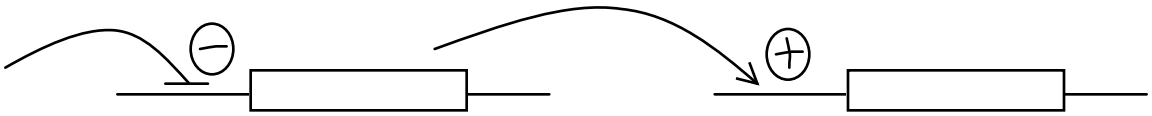
	Expression of the P_{osmA}-*lacZ* construct	
	no osmotic stress	osmotic stress
P_{osmA}-*lacZ*	–	+
osmB1 osmC1 P_{osmA}-*lacZ*	+	+

(a) Classify the *osmB1* and *osmC1* mutations as: constitutive OR uninducible, dominant OR recessive, **and** cis-acting OR trans-acting.

(b) Classify the *osmB1* and *osmC1* mutations by their possible identities with respect to *osmA* (your choices are: repressor –, activator –, promoter –, operator –, super activator, super repressor, dominant negative repressor, dominant negative activator).

(c) Classify wild-type *osmB* and *osmC* by their possible wild-type functions with respect to *osmA* (your choices are: negative regulator, positive regulator, promoter, operator). For each (that is, *osmB* and *osmC*), also state whether it functions as a regulatory DNA sequence that does not encode a protein product, or whether it functions to encode a regulatory protein.

(d) Diagram **three possible models** for regulatory pathways for *osmA* that can explain the behavior of the *osmB1* and *osmC1* mutations. (Please diagram only linear models in which each gene is controlled by no more than one regulator. Also, please invoke no more than one unknown trans-acting regulator in each model.) For each model, include the inducer (osmotic stress), the reporter P_{osmA}-*lacZ,* and wild-type *osmB* and *osmC*. For each model, be sure to use the proper notation of:



to indicate negative regulation by a trans-acting factor

to indicate positive regulation by a trans-acting factor

Next, you isolate a different mutation in the *osmC* gene, called *osmC2*. Genetic tests of the *osmC2* mutation yield the following:

	Expression of the P_{osmA}-*lacZ* construct	
	no osmotic stress	osmotic stress
P_{osmA}-*lacZ*	–	+
osmC2 P_{osmA}-*lacZ*	–	–
osmC2 P_{osmA}-*lacZ* / F' *osmC*$^+$	–	+

(e) Does this new data help you to narrow down which of your possible models from part **(d)** could be the correct model? **If so**, which of your possible models is/are consistent with this new data?

2. You are studying the regulation of a bacterial gene that is required for synthesis of the amino acid threonine. This gene is called *thrA*.

(a) Circle **one answer in each** of the two blanks below to complete this sentence:

It makes the most sense for this bacterium to regulate the expression of *thrA* by turning ON / OFF the expression of *thrA* when the bacterium is growing in the absence of threonine, and by turning ON / OFF the expression of *thrA* when the bacterium is growing in the presence of threonine.

In the bacterial strain you are studying, there is a monomeric protein called ThrG that binds to sequences between the promoter and the start site of transcription of *thrA.* When ThrG is bound upstream of *thrA*, *thrA* expression is blocked. You isolate a transposon insertion in *thrG* (*thrG*::Tn5) which results in a loss of ThrG function. Using an F' form of the F factor, you perform several genetic tests listed below. These tests are not necessarily presented in the order that you would have done them. **For each test:**

i. State which test you are performing.
   (ie. dominant/recessive, constitutive/uninducible, cis, trans, OR epistasis)
ii. State the conclusion you would expect to arrive at based on the test.
   (eg. If you select dom./rec. test, would you conclude dominant or recessive?)
   (eg. If you select trans test, does the locus act in cis or in trans?)
   (eg. If you select epistasis test, which gene do you conclude acts earlier?)
iii. Fill in the chart showing whether *thrA* would be expressed ("+") or not ("–").

(b)

i. _____

ii. _____

iii.

	thrA activity – threonine	*thrA* activity + threonine
wild type (*thrA*$^+$ *thrG*$^+$)		
thrA$^+$ *thrG*::Tn5 / F' *thrA*$^-$ *thrG*$^+$		

(c)

i. _____

ii. _____

iii.

	thrA activity – threonine	*thrA* activity + threonine
wild type (*thrA*$^+$ *thrG*$^+$*)*		
thrA$^+$ *thrG*::Tn5 / F' *thrA*$^+$ *thrG*$^+$		

(d)

i. _____

ii. _____

iii.

	thrA activity – threonine	*thrA* activity + threonine
wild type (*thrA*$^+$ *thrG*$^+$*)*		
thrA$^+$ *thrG*$^+$ / F' *thrA*$^-$ *thrG*::Tn5		

In the bacterial strain you are studying, there is a monomeric protein called ThrD that binds to sequences near to the promoter of the *thrG* gene whenever threonine is absent. When ThrD is bound upstream of *thrG*, *thrG* expression is repressed. You isolate an ochre mutation in *thrD* (*thrD*–) that results in a loss of ThrD function. Using an F' form of the F factor, you perform the genetic tests listed below. These tests are not necessarily presented in the order that you would have done them.

As above, for each test:

i. State which test you are performing.
   (ie. dominant/recessive, constitutive/uninducible, cis, trans, OR epistasis)

ii. State the conclusion you would expect to arrive at based on the test.
   (eg. If you select dom./rec. test, would you conclude dominant or recessive?)
   (eg. If you select trans test, does the locus act in cis or in trans?)
   (eg. If you select epistasis test, which gene do you conclude acts earlier?)

iii. Fill in the chart showing whether *thrA* would be expressed ("+") or not ("–").

(e)

i. _____

ii. _____

iii.

	thrA activity – threonine	*thrA* activity + threonine
wild type (*thrA*$^+$ *thrD*$^+$ *thrG*$^+$)		
thrA$^+$ *thrD*– *thrG*::Tn5		

(f)

i. _____

ii. _____

iii.

	thrA activity – threonine	*thrA* activity + threonine
wild type (*thrA*$^+$ *thrD*$^+$)		
thrA$^+$ *thrD*–		

(g)

i. _____

ii. _____

iii.

	thrA activity – threonine	*thrA* activity + threonine
wild type (*thrA*$^+$ *thrD*$^+$)		
thrA$^+$ *thrD*$^+$ / F' *thrA*$^-$ *thrD*–		

(h) Circle **one answer in each** of the four blanks below to complete these sentences:

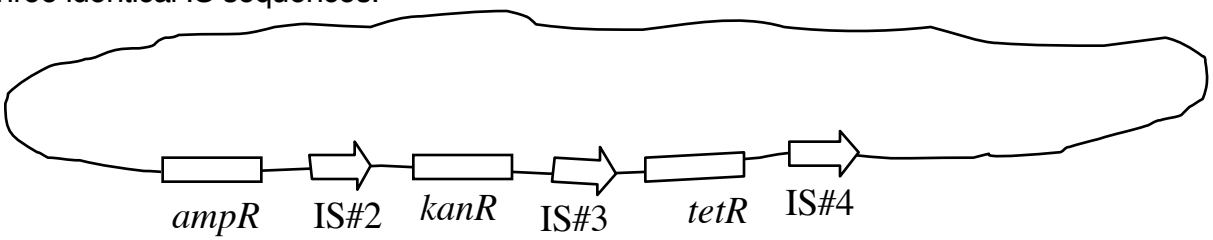
When the amino acid threonine is present and thus can bind to the ThrD protein, ThrD BINDS / DOESN'T BIND to sequences in front of the *thrG* gene, and expression of the *thrG* gene is thereby turned ON / OFF. Thus ThrG BINDS / DOESN'T BIND to sequences in front of the *thrA* gene, and expression of *thrA* is thereby turned ON / OFF.

(i) Diagram a model for the regulatory pathway for *thrA* that is consistent with all of the information you have about *thrA, thrD*, and *thrG*. In your model, include the amino acid threonine, and wild-type *thrA*, *thrD*, and *thrG*. Be sure to use the proper notation depicted by the drawings in Question #1 part **(d)**.

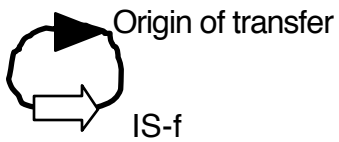
(j) Circle **one answer in each** of the two blanks below to complete these sentences:

The ThrD protein is a net POSITIVE / NEGATIVE regulator with respect to the *thrG* gene.
The ThrD protein is a net POSITIVE / NEGATIVE regulator with respect to the *thrA* gene.

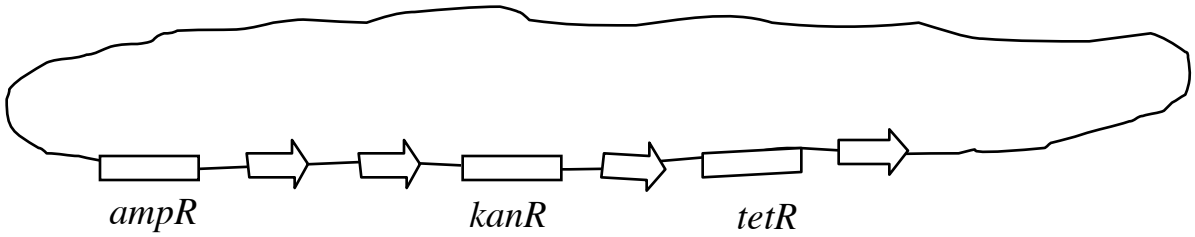
3. You are studying a strain of *E. coli* whose chromosome is diagrammed below. This strain's chromosome contains a small cluster of three genes encoding drug resistances and three identical IS sequences.



You are utilizing a form of the F factor that has a single IS sequence and a single origin of transfer, as indicated:



In a strain containing this chromosome and this F factor, a single recombination event can occur that leads to the formation of the following Hfr chromosome:



(a) In the above diagram, **box** the DNA that came from the F factor, and draw in the origin of transfer in the position and orientation in which it must be located using the symbol ▶.

(b) Which two insertion sequences must have recombined together to form the Hfr shown above? (In your answer, refer to the IS's as they were labeled in the original diagram.)

(c) In the Hfr shown above, would *ampR* be transferred early or late?

(d) In the Hfr shown above, would *kanR* be transferred early or late?

(e) In the Hfr shown above, would *tetR* be transferred early or late?

(f) If an F' plasmid formed from the above Hfr, and that F' carried the gene for kanamycin resistance, would it definitely, definitely not, **OR** maybe carry the gene for ampicillin resistance?

(g) If an F' plasmid formed from the above Hfr, and that F' carried the gene for kanamycin resistance, would it definitely, definitely not, **OR** maybe carry the gene for tetracycline resistance?

(h) If an F' plasmid formed from the above Hfr, and that F' carried the gene for tetracycline resistance, would it definitely, definitely not, **OR** maybe carry the gene for kanamycin resistance?

Solutions to 7.03 Problem Set 4

1. In a genetic screen using an uncharacterized bacterial species, you identify a gene called *osmA,* which is induced by osmotic stress. You do not have an easy way of detecting OsmA protein or its activity. Therefore, you use a version of the *lacZ* gene (which encodes the enzyme ß -galactosidase) that lacks its own promoter as a reporter for *osmA* expression. You construct a strain in which *lacZ* is inserted into the *osmA* gene so that *lacZ* is under the control of the *osmA* promoter and other regulatory DNA sequences (P_{osmA}-*lacZ*). The result is an engineered bacterial strain that expresses ß-galactosidase from the P_{osmA}-*lacZ* construct whenever OsmA should normally be produced.

	Expression of the P_{osmA}-*lacZ* construct	
	no osmotic stress	osmotic stress
P_{osmA}-*lacZ*	–	+

You now do a second genetic screen in which you look for mutations that disrupt the normal regulation of your P_{osmA}-*lacZ* fusion construct. In this screen, you isolate two mutations, which you call *osmB1* and *osmC1*. You find that both loci (*osmB* and *osmC*) are closely linked to P_{osmA}-*lacZ*. Making use of F' plasmids that carry the *osm* genes, you construct P_{osmA}-*lacZ* strains that you test for ß-galactosidase activity with results shown below.

		Expression of the P_{osmA}-*lacZ*	
		no osmotic stress	osmotic stress
1	P_{osmA}-*lacZ*	-	+
2	*osmB1* P_{osmA}-*lacZ*	+	+
3	*osmC1* P_{osmA}-*lacZ*	-	-
4	*osmB*$^+$/ F' *osmB1* P_{osmA}-*lacZ*	+	+
5	*osmC*$^+$/ F' *osmC1* P_{osmA}-*lacZ*	-	-
6	*osmB*$^+$ P_{osmA}-*lacZ* / F' *osmB1*	-	+
7	*osmB1* P_{osmA}-*lacZ* / F' *osmB*$^+$	+	+
8	*osmC1* P_{osmA}-*lacZ* / F' *osmC*$^+$	-	-
9	*osmC*$^+$ P_{osmA}-*lacZ* / F' *osmC1*	-	-

You next construct an *osmB1 osmC1* double mutant, which gives the following behavior:

	Expression of the P_{osmA}-*lacZ* construct	
	no osmotic stress	osmotic stress
P_{osmA}-*lacZ*	–	+
osmB1 osmC1 P_{osmA}-*lacZ*	+	+

(a) Classify the *osmB1* and *osmC1* mutations as: constitutive OR uninducible, dominant OR recessive, **and** cis-acting OR trans-acting.

The osmB1 mutant is constitutive, dominant, and cis-acting.

- *In the osmB1 mutant (row 2), the reporter is expressed with or with out osmotic stress, thus osmB1is constitutive.*
- *If a merodiploid possessing one mutant allele and one wild-type allele (row 4) has the same phenotype as the mutant, the mutation is dominant.*
- *osmB+ P_{osmA}-lacZ / F' osmB1 (row 6) has the dominant allele of osmB (osmB1) on a different DNA molecule than the good reporter (i.e. in the trans configuration). This strain shows the recessive phenotype so the osmB locus fails the trans test.*

The osmC1 mutant is uninducible, dominant, and trans-acting.

- *In the osmC1 mutant (row 3), the reporter is never expressed, thus osmC1is uninducible.*
- *If a merodiploid possessing one mutant allele and one wild-type allele (row 5) has the same phenotype as the mutant, the mutation is dominant.*
- *osmC+ P_{osmA}-lacZ / F' osmC1 (row 9) has the dominant allele of osmC (osmC1) on a different DNA molecule than the good reporter (i.e. in the trans configuration). This strain shows the dominant phenotype so the osmC locus passes the trans test.*

(b) Classify the *osmB1* and *osmC1* mutations by their possible identities with respect to *osmA* (your choices are: repressor –, activator –, promoter –, operator –, super activator, super repressor, dominant negative repressor, dominant negative activator).

B1 is an operator- mutation, because it is constitutive, dominant and cis.

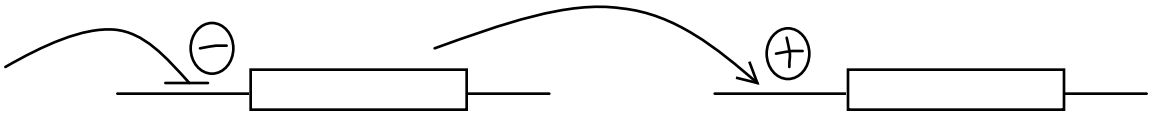
C1 is either a super repressor or a dominant negative activator mutation because it is uninducible, dominant and trans.

(c) Classify wild-type *osmB* and *osmC* by their possible wild-type functions with respect to *osmA* (your choices are: negative regulator, positive regulator, promoter, operator). For each (that is, *osmB* and *osmC*), also state whether it functions as a regulatory DNA sequence that does not encode a protein product, or whether it functions to encode a regulatory protein.

B+ is an operator, which is a regulatory DNA sequence that does not encode a protein. As the B1 mutation is an operator- mutation, B+ must be an operator.

C+ is either a positive or a negative regulator, and C+ must encode a trans-acting regulatory protein. As the C1 mutation is a super repressor or a dominant negative activator mutation, then wild type C+ must either function as a positive regulator or a negative regulator. Positive and negative regulators are proteins.

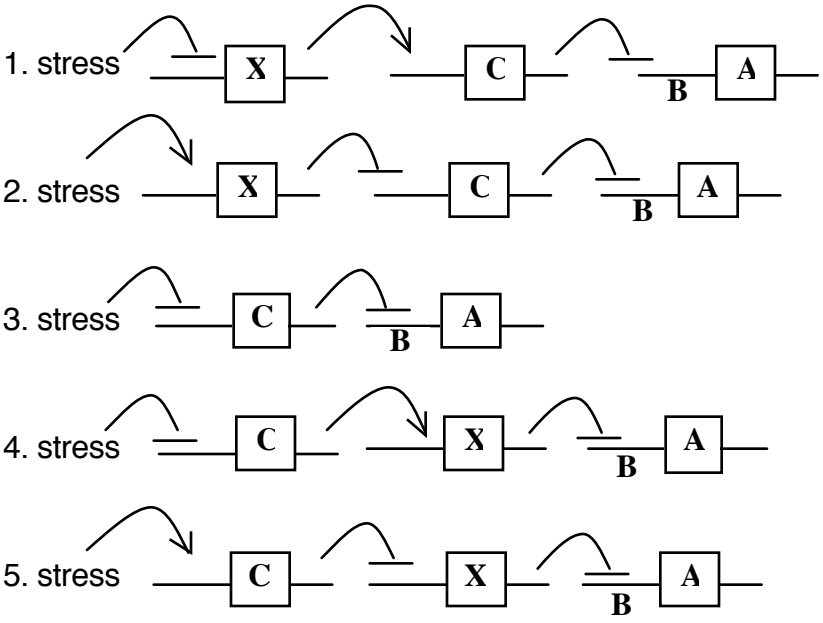
(d) Diagram **three possible models** for regulatory pathways for *osmA* that can explain the behavior of the *osmB1* and *osmC1* mutations. (Please diagram only linear models in which each gene is controlled by no more than one regulator. Also, please invoke no more than one unknown trans-acting regulator in each model.) For each model, include the inducer (osmotic stress), the reporter P_{osmA}-*lacZ,* and wild-type *osmB* and *osmC*. For each model, be sure to use the proper notation of:



to indicate negative regulation by a trans-acting factor

to indicate positive regulation by a trans-acting factor

There are 5 possible models, although you only needed to provide 3 of them in this question. Each model fulfills a set of three criteria that are dictated by the information in the above problem.
Criterion One: In each model, osmB is a cis-acing operator sequence in front of osmA.
Criterion Two: In each model, a repressor acts directly on osmA, because an operator sequence is a binding site for a repressor. Also the problem states that each gene can only be regulated by a single regulatory protein. Thus an activator cannot bind to osmA.
Criterion Three: In each model, osmotic stress has a net positive effect on osmA. All 5 models below fit these 3 criteria.



Next, you isolate a different mutation in the *osmC* gene, called *osmC2*. Genetic tests of the *osmC2* mutation yield the following:

	Expression of the P_{osmA}-*lacZ* construct	
	no osmotic stress	osmotic stress
P_{osmA}-*lacZ*	–	+
osmC2 P_{osmA}-*lacZ*	–	–
osmC2 P_{osmA}-*lacZ* / F' *osmC*$^+$	–	+

(e) Does this new data help you to narrow down which of your possible models from part **(d)** could be the correct model? **If so**, which of your possible models is/are consistent with this new data?

The C2 mutation is uninducible and recessive. We already know that C works in trans from our tests with osmC1. Therefore, C2 is an activator- mutant. Therefore C+ must be a net activator of the system. The only one of the 5 models where C acts as a net activator is #5 (it negatively regulates X which negatively regulates osmA, and two negatives make a positive in regulatory circuits). Therefore #5 is the only model from part (d) that is consistent with the new data. If you didn't have #5 as one of your 3 models, the new info would tell you that there are more possible models that you didn't consider that are in fact consistent with all of the data up through part (d).

2. You are studying the regulation of a bacterial gene that is required for synthesis of the amino acid threonine. This gene is called *thrA*.

(a) Circle **one answer in each** of the two blanks below to complete this sentence:

It makes the most sense for this bacterium to regulate the expression of *thrA* by turning (**ON**) / OFF the expression of *thrA* when the bacterium is growing in the absence of threonine, and by turning ON / (**OFF**) the expression of *thrA* when the bacterium is growing in the presence of threonine.
It would make sense for a bacterium to make threonine only when it's not provided in the environment. thrA is required to make threonine, therefore thrA should be turned on in the absence of threonine, and should be turned off in the presence of threonine.

In the bacterial strain you are studying, there is a monomeric protein called ThrG that binds to sequences between the promoter and the start site of transcription of *thrA.* When ThrG is bound upstream of *thrA*, *thrA* expression is blocked. You isolate a transposon insertion in *thrG* (*thrG*::Tn5) which results in a loss of ThrG function. Using an F' form of the F factor, you perform several genetic tests listed below. These tests are not necessarily presented in the order that you would have done them. **For each test:**

i. State which test you are performing.
   (ie. dominant/recessive, constitutive/uninducible, cis, trans, OR epistasis)
ii. State the conclusion you would expect to arrive at based on the test.
   (eg. If you select dom./rec. test, would you conclude dominant or recessive?)
   (eg. If you select trans test, does the locus act in cis or in trans?)
   (eg. If you select epistasis test, which gene do you conclude acts earlier?)
iii. Fill in the chart showing whether *thrA* would be expressed ("+") or not ("–").

(b)

i. *trans*

ii. *trans*

iii.

	thrA activity – threonine	*thrA* activity + threonine
wild type (*thrA*$^+$ *thrG*$^+$)	+	-
thrA$^+$ *thrG*::Tn5 / F' *thrA*$^-$ *thrG*$^+$	+	-

In this test, thrA+ and thrG- are on the bacterial chromosome, and thrA- and thrG+ are on the F' plasmid. This is a trans test because the two alleles that give the dominant phenotypes are in trans to (or far from) each other. thrG+ is dominant over thrG-, and thrA+ is dominant over thrA. Only loci that encode proteins that can float around in the cell and act on DNA sequences found anywhere can pass the trans test. ThrG is such a trans-acting regulatory protein, and thus the thrG locus will pass the trans test. Passing the trans test means that the strain will display the dominant phenotype, and since thrG+ is dominant over thrG-, the dominant phenotype is the phenotype that is displayed by a thrG+ cell, which is inducible.

(c)

i. *dom/rec*

ii. *recessive*

iii.

	thrA activity – threonine	*thrA* activity + threonine
wild type (*thrA*$^+$ *thrG*$^+$)	+	-
thrA$^+$ *thrG*::Tn5 / F' *thrA*$^+$ *thrG*$^+$	+	-

In this test, the bacteria are merodiploid for thrG. There is a mutated thrG::Tn5 allele on the chromosome, and a wild type thrG+ allele on the F' plasmid. Looking at the phenotype of the merodiploid possessing one mutant allele and one wild-type allele is how you test for dominance/recessivity. We know that thrG::Tn5 must be recessive because it is a loss of function allele of a gene encoding a monomeric protein. When a mutation is recessive, the phenotype of a merodiploid possessing one mutant allele and one wild-type allele is the same as the phenotype of the wild type, as shown in the table.

(d)

i. *cis*

ii. *thrG passes the cis test. However, thrG is trans-acting so it would also pass the trans test.*

iii.

	thrA activity – threonine	*thrA* activity + threonine
wild type (*thrA*$^+$ *thrG*$^+$)	+	-
thrA$^+$ *thrG*$^+$ / F' *thrA*$^-$ *thrG*::Tn5	+	-

In this test, there are the wild type alleles of thrA and thrG on the bacterial chromosome, and the mutant alleles of thrA and thrG on the F' plasmid. This is a cis test because the two alleles that give the dominant phenotypes are in cis to (near to) each other. thrG+ is dominant over thrG-, and thrA+ is dominant over thrA-. All loci pass the cis test, and thus all strains constructed for use in cis tests display the dominant phenotype.

In the bacterial strain you are studying, there is a monomeric protein called ThrD that binds to sequences near to the promoter of the *thrG* gene whenever threonine is absent. When ThrD is bound upstream of *thrG*, *thrG* expression is repressed. You isolate an ochre mutation in *thrD* (*thrD*–) that results in a loss of ThrD function. Using an F' form of the F factor, you perform the genetic tests listed below. These tests are not necessarily presented in the order that you would have done them.

As above, for each test:

i. State which test you are performing.
   (ie. dominant/recessive, constitutive/uninducible, cis, trans, OR epistasis)
ii. State the conclusion you would expect to arrive at based on the test.
   (eg. If you select dom./rec. test, would you conclude dominant or recessive?)
   (eg. If you select trans test, does the locus act in cis or in trans?)
   (eg. If you select epistasis test, which gene do you conclude acts earlier?)
iii. Fill in the chart showing whether *thrA* would be expressed ("+") or not ("–").

(e)

i. *epistasis*

ii. *thrD acts earlier than thrG*

iii.

	thrA activity – threonine	*thrA* activity + threonine
wild type (*thrA*$^+$ *thrD*$^+$ *thrG*$^+$)	+	-
thrA$^+$ *thrD*– *thrG*::Tn5	+	+

In this test there is a double mutant (thrD- and thrG::Tn5) on the bacterial chromosome and no F' plasmid. This is an epistasis test. We know that thrD represses thrG which represses thrA. Therefore, the phenotype of a single thrD- loss of function mutant is uninducible. The phenotype of a single thrG- loss of function mutant is constitutive. Since ThrD acts before ThrG, the phenotype of the double mutant would be the same as the phenotype of the thrG- single mutant, which is constitutive.

(f)

i. *constitutive/uninducible*

ii. *uninducible*

iii.

	thrA activity – threonine	*thrA* activity + threonine
wild type (*thrA*$^+$ *thrD*$^+$)	+	-
thrA$^+$ *thrD*–	-	-

This is a test of the phenotype of a strain containing a single mutation. It is a constitutive/uninducible test. Since thrD represses thrG which represses thrA, a loss of thrD function would result in an unregulated thrG which would always repress thrA. thrA would therefore be uninducible.

(g)

i. *cis*

ii. *thrD passes the cis test. However, thrD is trans-acting so it would also pass the trans test.*

iii.

	thrA activity – threonine	*thrA* activity + threonine
wild type (*thrA*$^+$ *thrD*$^+$)	+	-
thrA$^+$ *thrD*$^+$ / F' *thrA*$^-$ *thrD*–	+	-

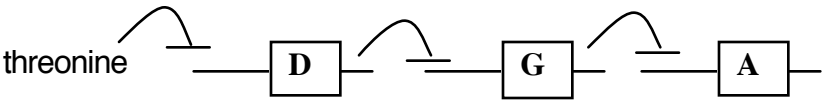
In this test, there are the wild type alleles of thrA and thrD on the bacterial chromosome, and the mutant alleles of thrA and thrD on the F' plasmid. This is a cis test because the two alleles that give the dominant phenotypes are in cis to (near to) each other. thrD+ is dominant over thrD-, and thrA+ is dominant over thrA-. All loci pass the cis test, and thus all strains constructed for use in cis tests display the dominant phenotype.

(h) Circle **one answer in each** of the four blanks below to complete these sentences:

When the amino acid threonine is present and thus can bind to the ThrD protein, ThrD BINDS / **DOESN'T BIND** to sequences in front of the *thrG* gene, and expression of the *thrG* gene is thereby turned **ON** / OFF. Thus ThrG **BINDS** / DOESN'T BIND to sequences in front of the *thrA* gene, and expression of *thrA* is thereby turned ON / **OFF**.

According to the question, ThrD protein binds to the regulatory sequences in front of thrG when threonine is absent. Therefore, when threonine is present, ThrD DOESN'T BIND to sequences in front of the thrG gene. The question tells us that ThrD represses thrG, therefore when ThrD doesn't bind thrG, thrG is turned ON. The question tells us that ThrG represses thrA, so when ThrG is produced, it BINDS the promoter of thrA and turns it OFF.

(i) Diagram a model for the regulatory pathway for *thrA* that is consistent with all of the information you have about *thrA, thrD*, and *thrG*. In your model, include the amino acid threonine, and wild-type *thrA*, *thrD*, and *thrG*. Be sure to use the proper notation depicted by the drawings in Question #1 part **(d)**.



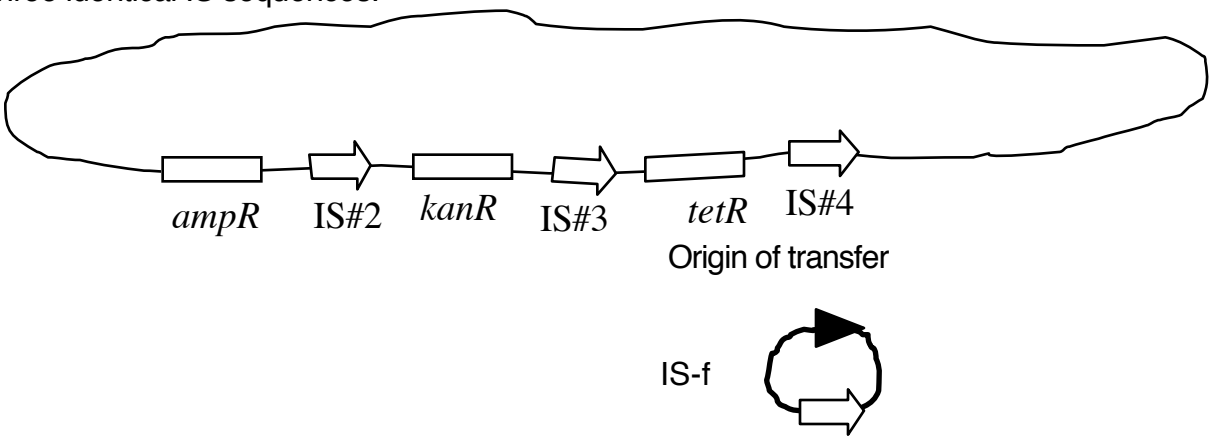
We know that threonine prevents ThrD from functioning (threonine negatively regulates thrD). We know that thrD represses thrG. We know that thrG represses thrA.

(j) Circle **one answer in each** of the two blanks below to complete these sentences:

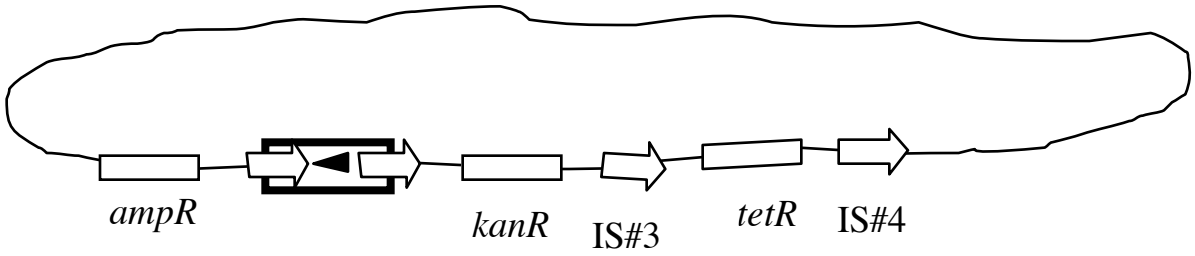
The ThrD protein is a net POSITIVE / **NEGATIVE** regulator with respect to the *thrG* gene. The ThrD protein is a net **POSITIVE** / NEGATIVE regulator with respect to the *thrA* gene.

ThrD represses thrG, therefore it is a negative regulator of thrG. ThrD represses thrG, which represses thrA. Two negatives make a positive, therefore ThrD is a net positive regulator of thrA.

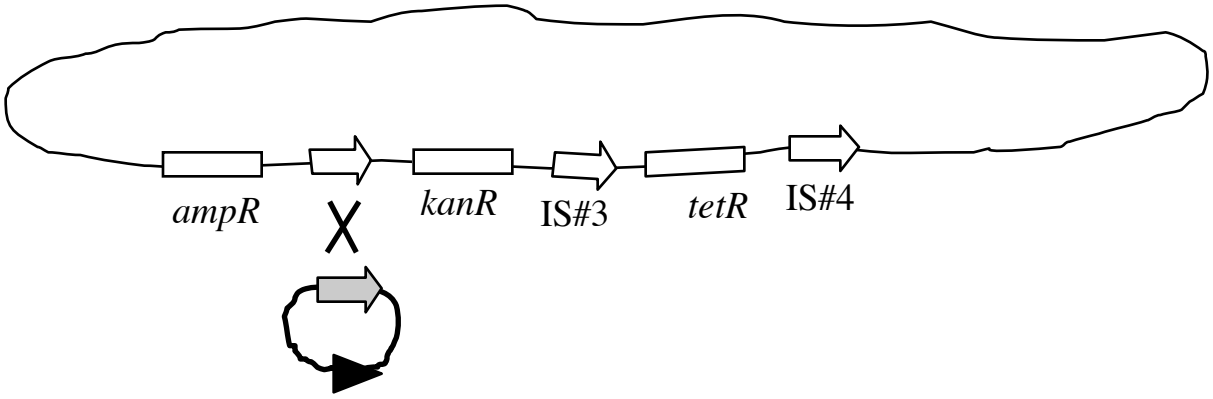
3. You are studying a strain of *E. coli* whose chromosome is diagrammed below. This strain's chromosome contains a small cluster of three genes encoding drug resistances and three identical IS sequences.



You are utilizing a form of the F factor that has a single IS sequence and a single origin of transfer, as indicated. In a strain containing this chromosome and this F factor, a single recombination event can occur that leads to the formation of the following Hfr chromosome:



(a) In the above diagram, **box** the DNA that came from the F factor, and draw in the origin of transfer in the position and orientation in which it must be located using the symbol ▶ .



The recombination event must have taken place as shown above. To get a final result of two IS sequences between ampR and kanR, the recombination must have taken place between IS2 and IS-f. These two insertion sequences must have been oriented

in the same direction for the recombination to happen. To draw the product of this recombination event, trace a "figure-eight" from the left side of the F factor, up through the recombination event, around to the right side of the chromosome, around to the left side of the chromosome, back down to the right side of the F factor.

NOTE: As the recombination event happened in the middle of the two IS sequences, the right half of the leftmost IS in the Hfr (and the left half of the neighboring IS in the Hfr) will be from the F factor. (See where the box is drawn above for a visual representation of this information.)

(b) Which two insertion sequences must have recombined together to form the Hfr shown above? (In your answer, refer to the IS's as they were labeled in the original diagram.)

IS-f from the F plasmid and IS-2 from the chromosome must have recombined as explained above.

(c) In the Hfr shown above, would *ampR* be transferred early or late?

The Hfr will transfer in the direction that the origin of transfer is pointing (ie left), taking sequences from behind the blunt side of the arrow first. Thus kanR would be the first gene to be transferred, and ampR would be the last gene to be transferred (the latest gene to be transferred).

(d) In the Hfr shown above, would *kanR* be transferred early or late?
Early. See part (c).

(e) In the Hfr shown above, would *tetR* be transferred early or late?
Early. See part (c).

(f) If an F' plasmid formed from the above Hfr, and that F' carried the gene for kanamycin resistance, would it definitely, definitely not, **OR** maybe carry the gene for ampicillin resistance?

Definitely not. F' plasmids form between two IS sequences on an Hfr, but must contain the origin of transfer to be considered a form of the F factor. If an F' contained kanamycin, it couldn't contain ampicillin, because there are no IS sequences to the left of ampicillin resistance that would allow ampR to be brought onto an F' through recombination.

(g) If an F' plasmid formed from the above Hfr, and that F' carried the gene for kanamycin resistance, would it definitely, definitely not, **OR** maybe carry the gene for tetracycline resistance?

Maybe. If the leftmost IS sequence on the Hfr recombined with IS#4, the F' plasmid would contain kanR and tetR. However, if the leftmost IS sequence on the Hfr recombined with IS#3, the F' plasmid would contain kanR but not tetR.

(h) If an F' plasmid formed from the above Hfr, and that F' carried the gene for tetracycline resistance, would it definitely, definitely not, **OR** maybe carry the gene for kanamycin resistance?

Definitely. The only way that tetR will be included on an F' factor is if the leftmost IS sequence on the Hfr recombined with IS4. If the leftmost IS sequence on the Hfr recombined with IS#4, the F' plasmid would contain both kanR and tetR.

7.03 Problem Set 5

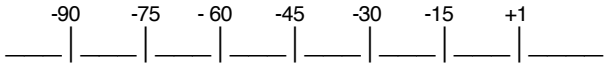
Due before 5 PM on Thursday, November 18, 2004
Hand in answers in recitation section or in the box outside of 68-120

1. You are studying the regulation of a yeast enzyme “convertase” (encoded by the Con1 gene). This enzyme converts the compound lifanone into the compound livanic. The regulation of the Con1 gene is designed to minimize the amount of energy that the yeast must expend to ensure that livanic is available to the cell. Yeast cells require livanic for growth and it can be obtained in the following three ways (which are listed in order from least amount of energy required, to most energy required):

1) Livanic can be provided in the medium.
2) Lifanone can be supplied in the medium and then converted to livanic.
3) Livanic can be synthesized *de novo*. (This pathway **does not** include the reaction in which lifanone is coverted to livanic.)

To understand how convertase expression is regulated, you fuse a small region of the yeast genome containing only the Con1 regulatory sequences to the *lacZ* coding sequence and then place this hybrid gene on an appropriate yeast plasmid. Wild-type cells carrying this plasmid express ß -galactosidase under the same conditions in which convertase is expressed in wild-type cells (see line 1 below).

You create a series of different 15 base-pair long deletions in the Con1 regulatory region. You examine the phenotype of wild-type yeast cells which harbor plasmids containing *lacZ* fused to each of the following versions of the Con1 regulatory region:



		Color when grown on:		
		NEITHER	LIFANONE	LIVANIC
1)	wt ——————————— *lacZ*	white	blue	white
2)	A- —— ——————————— *lacZ*	white	white	white
3)	B- ———— ——————————— *lacZ*	white	blue	white
4)	C- ————— ——————————— *lacZ*	white	blue	blue
5)	D- ——————— ———— *lacZ*	white	blue	white
6)	E- ——————————— ——— *lacZ*	white	white	white
7)	F- ——————————— *lacZ*	white	blue	white

KEY TO THE PLATES ON PAGE 1:

NEITHER = a plate that contains X-gal but does not contain livanic or lifanone

LIFANONE = a plate that contains X-gal and lifanone, but no livanic.

LIVANIC = a plate that contains X-gal and livanic. The results of this plate are the same whether or not it contains lifanone.

(a) State whether A^-, C^-, and E^- are cis-acting or trans-acting with respect to Con1. Briefly justify your answers.

(b) Predict the phenotypes of cells containing each of the following pairs of plasmids. Briefly justify your predictions.

	The two plasmids contained in each cell	Colony color when grown on plates containing:		
		NEITHER	LIFANONE	LIVANIC
i)	A^- P_{con1}-*lacZ* and A^+ P_{con1}-*lacZ*. Justify:			
ii)	C^- P_{con1}-*lacZ* and C^+ P_{con1}-*lacZ*. Justify:			
iii)	E^- P_{con1}-*lacZ* and E^+ P_{con1}-*lacZ*. Justify:			

(c) Classify the A^- and C^- mutations by their possible identities with respect to Con1 (your choices are: repressor –, activator –, UAS–, URS–, super activator, super repressor, dominant negative repressor, dominant negative activator).

(d) You find that a protein Reg1 binds to the –45 to –60 part of the Con1 regulatory region, and that a protein Reg2 binds to the –75 to –90 part of the Con1 regulatory region. Diagram **two possible models** for the regulatory pathway for Con1 that can explain the behavior of the A^- and C^- mutations. For each model, include **only** lifanone, livanic, Reg1, Reg2, the reporter P_{con1}-*lacZ,* and the wild-type A^+ and C^+ loci. (Please have livanic and lifanone acting through different trans-acting regulators in your model.)

(e) Circle the one correct choice in each of the two blanks in the following sentence:

Livanic is a net INHIBITOR / INDUCER with respect to the Con1 gene. Lifanone is a net INHIBITOR / INDUCER with respect to the Con1 gene.

2. You are studying the regulation of a yeast enzyme Cal1, whose activity you can measure directly. Expression of the gene encoding Cal1 is inhibited by the presence of calcium ions.

You perform a genetic screen looking for mutants that do not properly regulate expression of this enzyme. In this screen, you isolate a mutant that you call "$Cal2^-$." You analyze the following strains, with the following results:

	Units of Cal1 enzyme activity	
	- Ca	+Ca
$Cal1^+$ $Cal2^+$	200	3
$Cal1^-$ $Cal2^+$	0	0
$Cal1^+$ $Cal2^-$	3	3
$Cal1^-$ $Cal2^+$ / $Cal1^+$ $Cal2^+$	200	3
$Cal1^+$ $Cal2^-$ / $Cal1^+$ $Cal2^+$	200	3

Sporulation of diploids of the genotype Cal1$^-$ Cal2$^+$ / Cal1$^+$ Cal2$^-$ yields only one type of tetrad. All 40 tetrads that you analyze look as follows. Each column below contains the four colonies that grew up from each tetrad you have analyzed. Dark colored circles would indicate yeast colonies that are strongly expressing Cal1.

(a) Categorize this type of tetrad as PD, NPD, or TT.

(b) Can you conclude whether the Cal1 and Cal2 loci are linked? **If you can conclude**, state whether they are tightly linked, loosely linked, or unlinked.

(c) Classify Cal2$^-$ as constitutive or uninducible.

(d) Can you conclude if Cal2$^-$ is dominant or recessive? Why or why not?
If you can conclude, state whether Cal2$^-$ is dominant or recessive.

(e) Can you conclude if Cal2$^-$ is cis-acting or trans-acting? Why or why not?
If you can conclude, state whether Cal2$^-$ is cis-acting or trans-acting.

(f) List **all** of the possible kinds of mutations that Cal2$^-$ could be with respect to Cal1 (your choices are: repressor –, activator –, UAS–, URS–, super activator, super repressor, dominant negative repressor, dominant negative activator).

(g) State the genotype of one strain that you could make to distinguish between all of the models that you listed in part **(f)** for the kind of mutation that Cal2^{-} is. **For each model**, state what phenotype (uninducible, constitutive, OR regulated normally) that strain would display.

(h) The strain that you described in part **(g)** actually allows you to perform two different genetic tests at once, using only one strain. What are the two genetic tests that you simultaneously performed using the strain from part **(g)**?

3. The scenarios on the next page ask biological questions that can be addressed by creating genetically engineered mice. When creating engineered mice, the following 9 steps need to be considered. **For each mouse you make** in parts **(a)** and **(b)**, please state:

i) whether you are using pronuclear injection or gene targeting techniques

ii) what DNA you would introduce into the mouse cells (also draw the piece of DNA)

iii) whether you would put the DNA into a fertilized egg or ES cells

iv) the genotype of the mouse you would get the fertilized egg or ES cells from

v) where in the mouse genome the DNA you introduced would integrate

vi) how you will ensure that your DNA got into the mouse cells (for *gene targeting* only)

vii) whether creating your mouse should involve the generation of a chimera or not

viii) which additional breeding steps you would do to make the mouse you wanted

ix) two possible phenotypic results you could get from the newly made mice, and the corresponding conclusions you would make based on each result

(a) You are studying an autosomal gene in mice called *m-Per* whose function is necessary to properly control the mouse's circadian rhythm. You want to know whether half the amount of the protein product of that gene is sufficient to properly control circadian rhythms.

(b) You have mice in your lab that are homozygous for a loss of function mutation in the *Bonappetit* gene, a gene that is necessary for proper appetite control. These mice eat like a horse! You then find a mouse on the street that eats like a horse, and you know that this mutant phenotype is recessive. You want to know if the street mouse is a *Bonappetit* mutant or is a mutant in some other gene without doing any breeding or DNA sequencing.

(c) You give your street mouse and your lab mice to your friend, and challenge her to determine whether the street mouse has a mutation in the *Bonappetit* gene using only breeding. (For this part, you need only describe parts (viii) and (ix) from the instructions to this question, as your friend will not be making any new genetically engineered mice.)

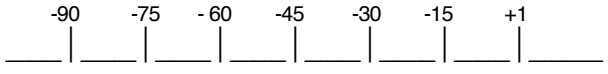
Solutions to 7.03 Problem Set 5

1. You are studying the regulation of a yeast enzyme "convertase" (encoded by the Con1 gene). This enzyme converts the compound lifanone into the compound livanic. The regulation of the Con1 gene is designed to minimize the amount of energy that the yeast must expend to ensure that livanic is available to the cell. Yeast cells require livanic for growth and it can be obtained in the following three ways (which are listed in order from least amount of energy required, to most energy required):

1) Livanic can be provided in the medium.
2) Lifanone can be supplied in the medium and then converted to livanic.
3) Livanic can be synthesized *de novo*. (This pathway **does not** include the reaction in which lifanone is coverted to livanic.)

To understand how convertase expression is regulated, you fuse a small region of the yeast genome containing only the Con1 regulatory sequences to the *lacZ* coding sequence and then place this hybrid gene on an appropriate yeast plasmid. Wild-type cells carrying this plasmid express ß -galactosidase under the same conditions in which convertase is expressed in wild-type cells (see line 1 below).

You create a series of different 15 base-pair long deletions in the Con1 regulatory region. You examine the phenotype of wild-type yeast cells which harbor plasmids containing *lacZ* fused to each of the following versions of the Con1 regulatory region:



		Color when grown on:		
		NEITHER	LIFANONE	LIVANIC
1)	wt ———————————— *lacZ*	white	blue	white
2)	A- ____ ———————————— *lacZ*	white	white	white
3)	B- _____ ———————— *lacZ*	white	blue	white
4)	C- _____ _____ *lacZ*	white	blue	blue
5)	D- _____ _____ *lacZ*	white	blue	white
6)	E- _____ _____ *lacZ*	white	white	white
7)	F- ———————————— *lacZ*	white	blue	white

KEY TO THE PLATES ON PAGE 1:

NEITHER = a plate that contains X-gal but does not contain livanic or lifanone

LIFANONE = a plate that contains X-gal and lifanone, but no livanic.

LIVANIC = a plate that contains X-gal and livanic. The results of this plate are the same whether or not it contains lifanone.

(a) State whether A^-, C^-, and E^- are cis-acting or trans-acting with respect to Con1. Briefly justify your answers.

We are told in the question that the deletions are in the regulatory region of the Con1 gene. The regulatory region consists of *sequences* responsible for the transcription of Con1. Therefore, A-, C-, and E- are all cis-acting mutations with respect to Con1.

(b) Predict the phenotypes of cells containing the following two plasmids. Briefly justify your predictions.

	Genotype of cell	Colony color when grown on plates containing:		
		NEITHER	LIFANONE	LIVANIC
i)	A^- P_{con1}-*lacZ* and A^+ P_{con1}-*lacZ*. Justify: A+ causes inducible expression of P_{con1}-*lacZ* in the presence of lifanone. This expression of the reporter gene hides the uninducible phenotype resulting from the A- mutant. Therefore, A- is recessive to A+.	white	blue	white
ii)	C^- P_{con1}-*lacZ* and C^+ P_{con1}-*lacZ*. Justify: C- causes constitutive expression of P_{con1}-*lacZ*. This masks the C+ inducible phenotype by resulting in expression of P_{con1}-*lacZ* even in the presence of livanic, which normally acts as an inhibitor of Con1. Therefore, C- is dominant to C+.	white	blue	blue
iii)	E^- P_{con1}-*lacZ* and E^+ P_{con1}-*lacZ*. Justify: E+ causes inducible expression of P_{con1}-*lacZ* in the presence of lifanone. This expression of the reporter gene hides the uninducible phenotype resulting from the E- mutant. Therefore, E- is recessive to E+.	white	blue	white

(c) Classify the A⁻ and C⁻ mutations by their possible identities with respect to Con1 (your choices are: repressor –, activator –, UAS–, URS–, super activator, super repressor, dominant negative repressor, dominant negative activator).
Both A- and C- have been identified in a) as cis-acting mutations. Therefore, we can eliminate all repressor and activator mutations, since they are trans-acting. This leaves us with UAS- and URS- as our two options.

We also know from the table in the beginning and from part b) that A- is uninducible and recessive. This implies that A+ is needed for expression of the P_{con1}-*lacZ* reporter, meaning it binds an activator of Con1. Therefore, A- is a UAS- mutation.

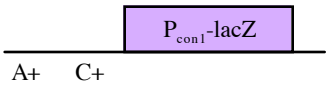
Finally, we know from the table in the beginning and from part b) that C- is constitutive and dominant. This means that C+ normally inhibits expression of the P_{con1}-*lacZ* reporter. Therefore, C+ must bind a repressor of Con1. This makes C- a URS- mutation.

(d) You find that a protein Reg1 binds to the –45 to –60 part of the Con1 regulatory region, and that a protein Reg2 binds to the –75 to –90 part of the Con1 regulatory region. Diagram **two possible models** for the regulatory pathway for Con1 that can explain the behavior of the A⁻ and C⁻ mutations. For each model, include **only** lifanone, livanic, Reg1, Reg2, the reporter P_{con1}-*lacZ,* and the wild-type A⁺ and C⁺ loci. (Please have livanic and lifanone acting through different trans-acting regulators in your model.)

This problem was intended to be solved in the following way:

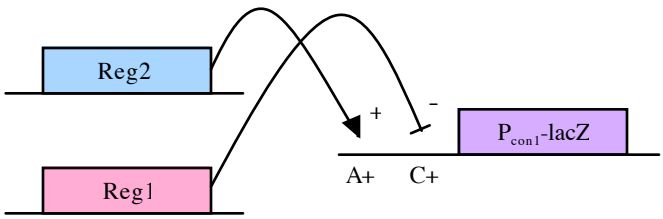
Step 1:

From the table in the beginning, we can draw a schematic of the P_{con1}-*lacZ* reporter gene and the location of the A+ and C+ regulatory sequences:



Step 2:

The question stem in d) tells us that Reg1 binds to the -45 to -60 part of the regulatory sequence, which is the location of C+. C+ is a URS (part c), so Reg1 must be a repressor of P_{con1}-*lacZ*.
Reg2 binds to the -75 to -90 part of the regulatory sequence, which is the location of A+. A+ is a UAS (part c), so Reg2 must be an activator of P_{con1}-*lacZ*.

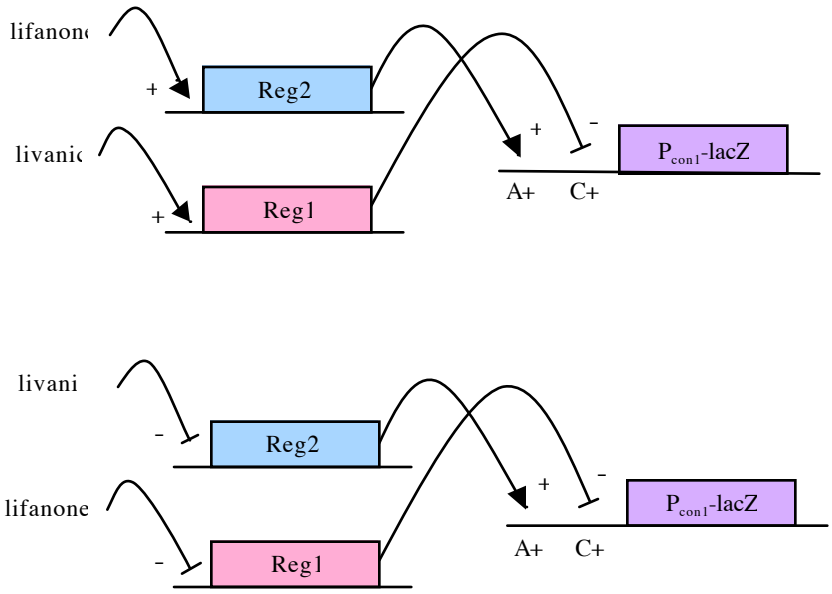


Step 3:

Finally, we have to figure out how to include lifanone and livanic in the pathway. We are told that lifanone and livanic must act through different trans-acting regulators. Our model only has two trans-acting regulators: Reg1 and Reg2. One model is that livanic acts on Reg1 while lifanone acts on Reg2. The other model is that lifanone acts on Reg1 while livanic acts on Reg2.

We also know from the first table that lifanone has a net positive effect on P_{con1}-*lacZ* expression and livanic has a net negative effect.

We can draw only two models that satisfy these constraints:



Unfortunately, if one now takes these models and goes back and looks at the data retrospectively, these models are not entirely consistent with the C- phenotype as shown in the table. For the upper model, a C- deletion would give the pattern WHITE BLUE BLUE in the table only if the last column was plates containing both livanic and lifanone (and not livanic only). The bottom model would give the pattern BLUE BLUE WHITE in the table.

Thus this part of Question 1 was not graded.

(e) Circle the one correct choice in each of the two blanks in the following sentence:

Livanic is a net INHIBITOR / INDUCER with respect to the Con1 gene. Lifanone is a net INHIBITOR / INDUCER with respect to the Con1 gene.

Livanic is a net INHIBITOR with respect to the Con1 gene – In the first table, row 1), the wild-type case, we see that cells grown in the presence of livanic are white, meaning that livanic INHIBITS the expression of P_{con1}-*lacZ*.

Lifanone is a net INDUCER with respect to the Con1 gene. – In the first table, row 1), the wild-type case, we see that cells grown in the presence of lifanone are blue, meaning that lifanone INDUCES the expression of P_{con1}-*lacZ*.

2. You are studying the regulation of a yeast enzyme Cal1, whose activity you can measure directly. Expression of the gene encoding Cal1 is inhibited by the presence of calcium ions.

You perform a genetic screen looking for mutants that do not properly regulate expression of this enzyme. In this screen, you isolate a mutant that you call "Cal2⁻." You analyze the following strains, with the following results:

	Units of Cal1 enzyme activity	
	- Ca	+Ca
Cal1⁺ Cal2⁺	200	3
Cal1⁻ Cal2⁺	0	0
Cal1⁺ Cal2⁻	3	3
Cal1⁻ Cal2⁺ / Cal1⁺ Cal2⁺	200	3
Cal1⁺ Cal2⁻ / Cal1⁺ Cal2⁺	200	3

Sporulation of diploids of the genotype Cal1⁻ Cal2⁺ / Cal1⁺ Cal2⁻ yields only one type of tetrad. All 40 tetrads that you analyze look as follows. Each column below contains the four colonies that grew up from each tetrad you have analyzed. Dark colored circles would indicate yeast colonies that are strongly expressing Cal1.

- Ca	+Ca
○	○
○	○
○	○
○	○

(a) Categorize this type of tetrad as PD, NPD, or TT.

If this were a PD tetrad, all the spores would be Cal1-Cal2+ or Cal1+Cal2-. The Cal1-Cal2+ spores would show uninducible expression, because there is no wild-type copy of Cal1 to be expressed. The Cal1+Cal2- spores would also be uninducible, as can be seen by looking at the table above.

If this were an NPD tetrad, all the spores would be Cal1-Cal2- or Cal1+Cal2+. The Cal1-Cal2- spores would be uninducible because they would not have either a good copy of the Cal1 reporter or the Cal2 regulator. The Cal1+Cal2+ spores would be inducible because they are wild-type for both loci.

If this were a TT tetrad, each spore would have a different genotype, with one each of Cal1-Cal2+, Cal1+Cal2-, Cal1-Cal2-, and Cal1+Cal2+. This would result in 3 uninducible spores and one inducible spore per tetrad.
In this case, we see that all four spores in the tetrads are uninducible, indicating that the tetrads are all PD.

(b) Can you conclude whether the Cal1 and Cal2 loci are linked? **If you can conclude**, state whether they are tightly linked, loosely linked, or unlinked.
Since all the tetrads are PD, no cross-overs have occurred, indicating that the Cal1 and Cal2 loci are tightly linked

(c) Classify Cal2$^-$ as constitutive or uninducible.
Cal2- is uninducible: The third row of the table in the beginning of the question is a phenotype test for the Cal2- mutation because it has a mutant copy of Cal2 and a wild-type reporter. Cal1 expression is not induced in this strain, even in the absence of Ca+.

(d) Can you conclude if Cal2$^-$ is dominant or recessive? Why or why not?
If you can conclude, state whether Cal2$^-$ is dominant or recessive.
Yes, recessive: The fifth row of the table is a dominant/recessive test. Both a mutant copy of Cal2 and a wild-type copy of Cal2 are present in this strain, along with wild-type copies of the reporter. The phenotype of this strain is inducible, indicating that Cal2+ is dominant to Cal2-.

(e) Can you conclude if Cal2$^-$ is cis-acting or trans-acting? Why or why not?
If you can conclude, state whether Cal2$^-$ is cis-acting or trans-acting.
No, we cannot conclude: We didn't conduct a cis/trans test. While we know that Cal1 and Cal2 are tightly linked, we cannot tell whether Cal2 is a regulatory sequence of Cal1 (cis-acting) or a regulatory protein whose encoding gene is at a nearby locus (trans-acting).

(f) List **all** of the possible kinds of mutations that Cal2$^-$ could be with respect to Cal1 (your choices are: repressor –, activator –, UAS–, URS–, super activator, super repressor, dominant negative repressor, dominant negative activator).
We know that Cal2- is recessive, which rules our super activator, super repressor, dominant negative repressor, and dominant negative activator.

We also know Cal2- is uninducible with respect to Cal1, meaning Cal2 is necessary for expression of Cal1. This means Cal2 must somehow be responsible for activating transcription of Cal1. This rules out repressor- and URS-.

This leaves activator- and UAS-.
If Cal2 is Cis with respect to Cal1, Cal2- is a UAS- mutation.
If Cal2 is Trans with respect to Cal1, Cal2- is an activator- mutation

(g) State the genotype of one strain that you could make to distinguish between all of the models that you listed in part **(f)** for the kind of mutation that Cal2⁻ is. **For each model**, state what phenotype (uninducible, constitutive, OR regulated normally) that strain would display.

We need to do a trans test to determine whether Cal2 is cis- or trans-acting with respect to Cal1. For a trans test, make a strain with the dominant regulator (Cal2+) in trans with the wild-type reporter (Cal1+):

Cal1+ Cal2- / Cal1- Cal2+

If Cis-acting, we will see the recessive phenotype: uninducible

If Trans-acting, we will see the dominant phenotype: inducible

(h) The strain that you described in part **(g)** actually allows you to perform two different genetic tests at once, using only one strain. What are the two genetic tests that you simultaneously performed using the strain from part **(g)**?

The test we performed in part g) was the trans test.

But we also performed a Complementation test. If the Cal1- and Cal2- mutations are in the same gene *(ie. if Cal2- is a mutation in a regulatory sequence upstream of the Cal1+ open reading frame),* the uninducible phenotype will not be rescued. If the Cal1- and Cal2- mutations are in different genes, and Cal2 is trans-acting with respect to Cal1, Cal2+ can rescue the uninducible phenotype.

3. The scenarios on the next page ask biological questions that can be addressed by creating genetically engineered mice. When creating engineered mice, the following 9 steps need to be considered. **For each mouse you make** in parts **(a)** and **(b)**, please state:

i) whether you are using pronuclear injection or gene targeting techniques

ii) what DNA you would introduce into the mouse cells (also draw the piece of DNA)

iii) whether you would put the DNA into a fertilized egg or ES cells

iv) what mouse you would get the fertilized egg or ES cells from

v) where in the mouse genome the DNA you introduced would integrate

vi) For *gene targeting*, how you will ensure that your DNA got into the mouse cells

vii) whether creating your mouse should involve the generation of a chimera or not

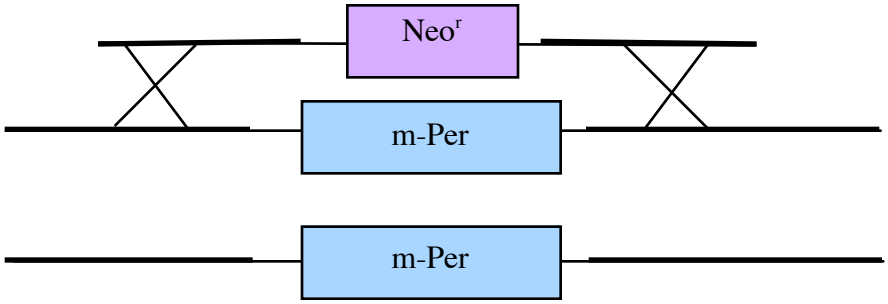
viii) which additional breeding steps you would do to make the mouse you wanted

ix) two possible phenotypic results you could get from the newly made mice, and the corresponding conclusions you would make based on each result

(a) You are studying an autosomal gene in mice called *m-Per* whose function is necessary to properly control the mouse's circadian rhythm. You want to know whether half the amount of the protein product of that gene is sufficient to properly control circadian rhythms.

i. Gene targeting: You want to see what the effect of removing one copy of m-Per is on the mouse's circadian rhythm. This involves creating a knock-out heterozygote that only has one functional copy of m-Per.

ii. The bold lines flanking the Neo^r construct and the m-Per genes are homologous sequences, which allow for homologous recombination and site-specific insertion of the targeting construct.



iii. ES cells. Homologous recombination in mice is very inefficient and thus we must start with many cells to ensure that we'll get what we want.

iv. m-Per homozygote: We will most likely get only one insertion in a given cell. A homozygote with one insertion will become heterozygous, with one copy of m-Per and one copy of the targeting construct.

v. m-Per locus, due to homologous sequences, which allow for site-specific insertion.

vi. Select for Neomycin resistance: Only cells that take up and integrate the targeting construct will contain the gene for neomycin resistance.

vii. Gene targeting involves the formation of a Chimera – A mouse with four parents: two for the donor ES cells and two for the host blastocyst.

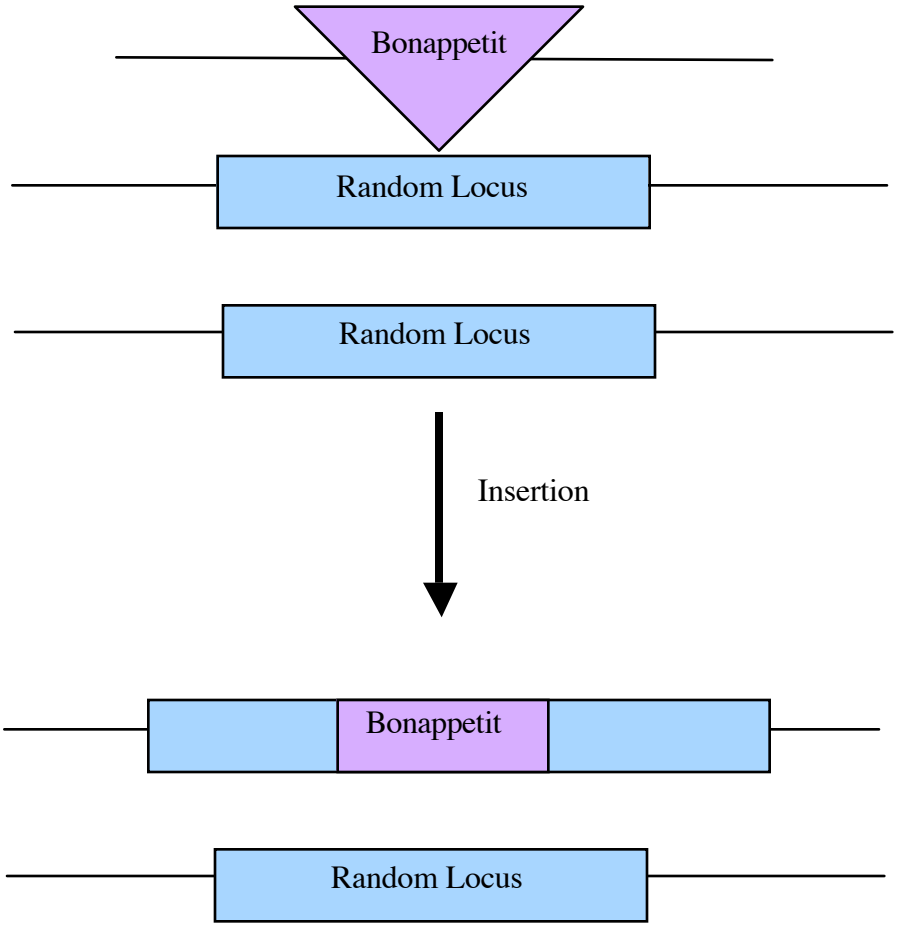
viii. Inject ES cells into a wild-type (m-Per/m-Per) blastocyst. Grow up chimeras. The chimeras will have some m-Per/m-Per host cells and some m-Per/Neo^r donor cells. We want mice whose cells are ALL m-Per/Neo^r, so breed the chimeras to wildtype (m-Per/m-Per) mice. Screen for offspring who are m-Per/Neo^r. Use these heterozygotes for further analysis.

ix. Two possible phenotypes: circadian rhythm is disrupted in heterozygotes or is normal in heterozygotes. If disrupted, _ the amount of m-Per protein is not sufficient to maintain a normal circadian rhythm. If normal, _ the amount of protein is sufficient to maintain a normal circadian rhythm.

(b) You have mice in your lab that are homozygous for a loss of function mutation in the *Bonappetit* gene, a gene that is necessary for proper appetite control. These mice eat like a horse! You then find a mouse on the street that eats like a horse, and you know that this mutant phenotype is recessive. You want to know if the street mouse is a *Bonappetit* mutant or is mutant is some other gene without doing any breeding or DNA sequencing.

i. Pronuclear injection, because additional breeding (after the initial formation of a fertilized egg) is not necessary. We want to insert the Bonappetit gene into a street mutant fertilized egg, and see if it can rescue the mutant phenotype.

ii.



iii. Fertilized egg

iv. The fertilized egg will come from a street mutant female mouse, who has mated with a street mutant male mouse, to form a fertilized egg that is homozygous for the street mouse recessive mutation.

v. Random location -- transgenes do not integrate in a site-specific manner because non-homologous recombination is much, much more efficient than homologous recombination in mice.
vi. Since we are using pronuclear injection and not gene-targeting, we do not have to ensure that the DNA entered the cell: It happens at a very high frequency.
vii. Chimeras are not made in pronuclear injection. The whole mouse is made from one fertilized egg, which contains the transgenic construct. This means that all the cells if the resulting mouse contain the trangene, as well.
viii. Allow your transgenic eggs to mature in a foster mother and use the resulting mice for further analysis.
ix. Two possible phenotypes: Normal appetites and huge appetites. If those expressing the transgene have normal appetites, the mutation is in Bonappetit, because insertion of wild-type Bonappetit into the mutant mice was able to rescue the mutant phenotype. If those expressing the transgene have huge appetites, the mutation is not in Bonappetit, because the wild-type Bonappetit gene was not sufficient to rescue the mutant phenotype.

(c) You give your street mouse and your lab mice to your friend, and challenge her to determine whether the street mouse has a mutation in the *Bonappetit* gene using only breeding. (For this part, you need only describe parts (viii) and (ix) from the instructions to this question, as your friend will not be making any new genetically engineered mice.)

viii. Cross the mutant street mouse and the mutant lab mouse.
ix. Both mutations are recessive. If the offspring are normal, the street mouse does not have a mutation in the Bonappetit gene. If the offspring have huge appetites, the street mouse has a mutation in the Bonappetit gene. (Complementation test)

7.03 Problem Set 6

Due before 5 PM on **WEDNESDAY**, November 24, 2004.
Turn answers in to the box outside of 68-120 (as there are NO SECTIONS 11/24 or 11/25).

1. You are studying an autosomal recessive trait in the fruit fly *Drosophila* called "vestigial wings," which is caused by the "vg" allele. A wild-type fly is "vg+" and has long wings. When answering the following parts, show all of your calculations.

A geneticist crosses wild-type males to vestigial-winged females (the P generation) to get the F1 generation. The geneticist then crosses the F1 flies together to get the F2 generation. The scientist removes the vestigial-winged flies from the F2 generation, and breeds the remaining flies together to get the F3 generation.

(a) What are the phenotypic frequencies of the two kinds of flies in the F3 generation?

(b) The scientist then removes the vestigial-winged flies from the F3 generation, and breeds the remaining flies together to get the F4 generation. What are the allele frequencies of the two kinds of alleles in the F4 generation?

(c) The scientist removes the vestigial-winged flies from the F4 generation, and breeds the remaining flies together to get the F5 generation. What are the genotype frequencies of the three kinds of genotypes in the F5 generation?

(d) Is the fly population described in this problem at Hardy-Weinberg equilibrium? Justify your answer in one sentence.

2. Consider a small village in which 100 individuals are of blood group "M," 1300 individuals are of blood group "N," and 100 individuals are of blood group "MN." When answering the following parts, show all of your calculations.

(a) Is this small village at Hardy-Weinberg equilibrium? Justify your answer.

(b) If the M and N blood groups were not co-dominant, would it be possible to determine whether the population is in Hardy-Weinberg equilibrium? Explain why or why not?

(c) After one generation of random mating in this village, if the new generation has 2000 people in it, how many will be of each blood type ("M," "N," and "MN")?

(d) In the generation of 2000 people from part **(c)**, how many copies of each allele ("L^M" and "L^N") will be present in the population?

(e) Fill in the following chart to show the number of each type of couple in the generation from part (c), where 2000 people = 1000 couples. Assume that each person mates with one (and only one) other person, and mating choice is random with respect to blood group.

One parent + Other parent	Number of this sort of **couple** in the generation of 1000 couples from part **(c)**
$L^M L^M$ + $L^M L^M$	
$L^M L^M$ + $L^M L^N$	
$L^M L^N$ + $L^N L^N$	
$L^M L^N$ + $L^M L^N$	
$L^M L^M$ + $L^N L^N$	
$L^N L^N$ + $L^N L^N$	

(f) Imagine a gene in the human population that has a mutation rate = 10^{-5} per generation, and the mutant form of the gene has a selective disadvantage, S = 0.1.

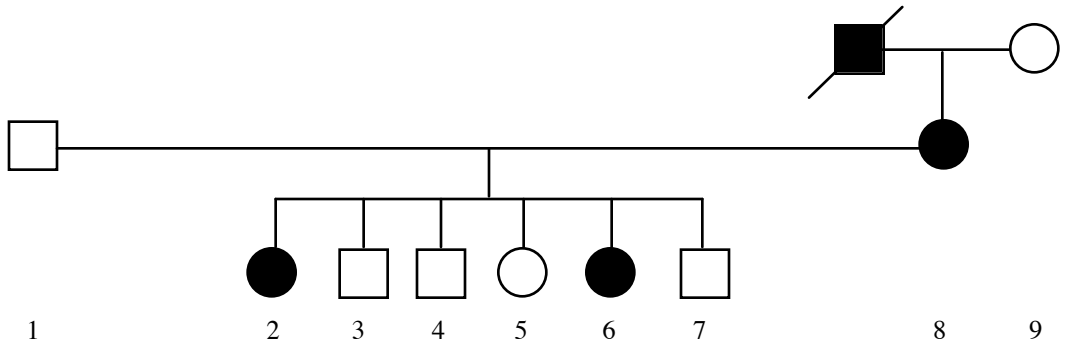
- If the mutation is autosomal recessive, calculate 1) the steady-state frequency of affected individuals in the population and 2) the fraction altered alleles that are due to new mutations.

- If the mutation is autosomal dominant, calculate 1) the steady-state frequency of affected individuals in the population and 2) the fraction altered alleles that are due to new mutations.

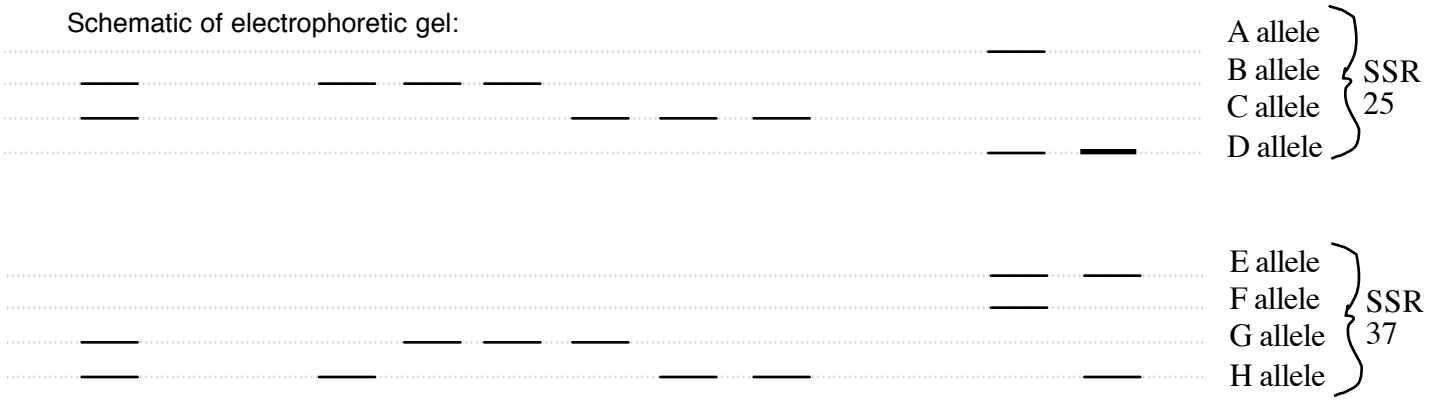
- If the mutation is X-linked recessive, calculate 1) the steady-state frequency of affected individuals in the population and 2) the fraction altered alleles that are due to new mutations.

3. The following is a pedigree for an autosomal dominant trait. This trait is caused by the Q allele. Thus individual 1 is "qq" and individual 8 is "Qq". You want to map the locus responsible for the trait with respect to two SSRs.

PLEASE NOTE: For individuals 2-7, only the alleles inherited from the father are shown. You will have to determine the alleles inherited from the mother later in the problem. Both the maternal and paternal alleles of SSR 25 and 37 are given for individuals 1, 8, and 9.



Schematic of electrophoretic gel:



Individual

	2	3	4	5	6	7
paternally inherited allele at the Q locus						
paternally inherited allele at SSR 25						
paternally inherited allele at SSR 37						

} to be filled in for part **(a)**

Individual

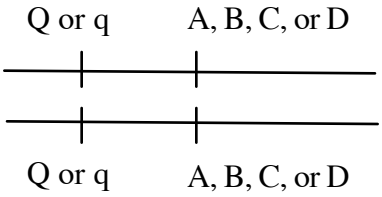
	2	3	4	5	6	7
maternally inherited allele at the Q locus						
maternally inherited allele at SSR 25						
maternally inherited allele at SSR 37						

} to be filled in for part **(c)**

→ to be filled in for part **(g)**

(a) Fill in the chart on page 4 with the paternally inherited alleles at all three loci.

(b) Draw all phases of individual 8 with respect to the Q locus and SSR 25 that are possible given everything you know about her. Make sure to draw phases using the proper notation:



(c) A human geneticist now tells you that the total probability of seeing the maternally inherited alleles of the Q and SSR 25 loci in individuals 2-7 equals $(0.5)^5 * (0.0)^0$, given that the Q and SSR 25 loci are linked at $\theta = 0$. Given this information, fill in chart above **and** draw in the maternally inherited bands in the above gel. (You have the SSR data for all the children, so fill all columns in.)

(d) Calculate the LOD score at $\theta = 0$ for SSR 25 and the Q locus. Show all calculations.

(e) Individuals 1 and 8 had six children and the human geneticist has taken the data for all six into account correctly in part **(b)**. Knowing this, give a brief explanation for why the human geneticist only included five meioses in the calculation of the LOD score shown above. Assume NO new mutations. Go back and make sure your data table and gel are consistent with your explanation.

(f) You now switch to mapping the two SSRs with respect to each other. Given everything you know about the mother (individual 8), draw all possible phases of the mother with respect to SSR 37 and SSR 25. Make sure to draw the phases using the proper notation.

(g) The human geneticist now tells you that the total probability of seeing the maternally inherited alleles of the SSR 25 and SSR 37 loci in individuals 2-7 equals $(0.45)^4$ * $(0.05)^1$, given that the two SSRs are linked at $\theta = 0.1$. Knowing this, draw in the maternally inherited bands of SSR 37 in the above gel. (You have the SSR data for all the children, so fill all columns in. Make sure you fill in the gel in a way that is consistent with the fact that the human geneticist only included 5 meioses in the calculation.)

(h) Given everything you know about the father (individual 1), draw all possible phases of the father with respect to SSR 37 and SSR 25. Make sure to draw the phases using the proper notation.

(i) Calculate the LOD score at $\theta = 0.1$ for the two SSRs. Make sure to incorporate the data from both the mother **and** the father, because both are informative parents regarding the inheritance of alleles at these two loci. Show all calculations.

Solutions to 7.03 Problem Set 6

1. You are studying an autosomal recessive trait in the fruit fly *Drosophila* called "vestigial wings," which is caused by the "vg" allele. A wild-type fly is "vg+" and has long wings. When answering the following parts, show all of your calculations.

A geneticist crosses wild-type males to vestigial-winged females (the P generation) to get the F1 generation. The geneticist then crosses the F1 flies together to get the F2 generation. The scientist removes the vestigial-winged flies from the F2 generation, and breeds the remaining flies together to get the F3 generation.

(a) What are the phenotypic frequencies of the two kinds of flies in the F3 generation?

First we will define our variables:

f(vg+) = p; f (vg) = q

F1 generation:

Since all individuals are heterozygous p = q = 0.5, and f (vg+/vg) = 1.

F2 generation

One generation of random mating will put the population in H-W equilibrium.
p = 0.5; q = 0.5
f (vg+/vg+) = p^2 = 0.25; f (vg/vg) = q^2 = 0.25 ; f (vg+/vg) = 2pq = 0.5

At this point vg/vg flies are removed and we recalculate genotypic frequencies
f (vg+/vg+) = (0.25/0.75) = 1/3; f (vg/vg) = 0 ; f (vg+/vg) = (0.5/0.75) = 2/3
p = f (vg+/vg+) + 0.5 f (vg+/vg) = 1/3 + .5x2/3 = 2/3
q = f (vg/vg) + 0.5 f (vg+/vg) = 0 + 0.5x2/3 = 1/3 (note that p + q =1)

F3 generation

Random mating of the remaining F2 flies will yield the following genotypic classes:
f (vg+/vg+) = p^2 = 4/9; f (vg+/vg) = 2pq = 4/9 ; f (vg/vg) = q^2 = 1/9

Thus wild type flies will make up 8/9 of the population, while flies with vestigial wings will make up 1/9 of the population

(b) The scientist then removes the vestigial-winged flies from the F3 generation, and breeds the remaining flies together to get the F4 generation. What are the allele frequencies of the two kinds of alleles in the F4 generation?

From part a) we know that before selection the genotypic ratios are the following:
f (vg+/vg+) = 4/9; f (vg+/vg) = 4/9 ; f (vg/vg) = 1/9
When we remove the 1/9 that are vg we must calculate new genotypic and new allele frequencies of generation F3:
f (vg+/vg+) = (4/9)/(8/9) = 1/2; f (vg+/vg) = (4/9)/(8/9) = 1/2 ; f (vg/vg) = 0
p = f (vg+/vg+) + 0.5 f (vg+/vg) = 0.5 + 0.5x0.5 = 0.75
q = f (vg/vg) + 0.5 f (vg+/vg) = 0 + 0.5X0.5 = 0.25
Random mating of the remaining F3 flies to produce the F4 flies will not change the allele frequencies across generations:
p = 0.75; q = 0.25

(c) The scientist removes the vestigial-winged flies from the F4 generation, and breeds the remaining flies together to get the F5 generation. What are the genotype frequencies of the three kinds of genotypes in the F5 generation?

From part b) we know that allele frequencies of the F3 generation after selection are:
p = 0.75; q = 0.25

Random mating of the F3 flies will give the following F4 genotypic classes:
f (vg+/vg+) = p^2 = 9/16; f (vg+/vg) = 2pq = 6/16; f (vg/vg) = q^2 = 1/16

Removal of the vestigial winged flies from the F4 will change the allele and genotypic frequencies:
f (vg+/vg+) = (9/16)/(15/16) = 9/15; f (vg+/vg) = (6/16)/(15/16) = 6/15;
f (vg/vg) = 0
p = f (vg+/vg+) + 0.5 f (vg+/vg) = 9/15 + 0.5x6/15 = 4/5
q = f (vg/vg) + 0.5 f (vg+/vg) = 0 + 0.5X6/15 = 1/5

Random mating of the remaining F4 flies will yield the following genotypes in the F5:
f **(vg+/vg+) = p^2 = 16/25;** f **(vg+/vg) = 2pq = 8/25;** f **(vg/vg) = q^2 = 1/25**

(d) Is the fly population described in this problem at Hardy-Weinberg equilibrium? Justify your answer in one sentence.

No, Hardy-Weinberg equilibrium assumes that there is no selection; however in these experiments vg/vg flies are being removed from the population before they mate. Thus selection against vg/vg flies is occurring with S=1.

2. Consider a small village in which 100 individuals are of blood group "M," 1300 individuals are of blood group "N," and 100 individuals are of blood group "MN." When answering the following parts, show all of your calculations.

(a) Is this small village at Hardy-Weinberg equilibrium? Justify your answer.

First calculate genotype frequencies:
f (m/m) =100/1500 = 0.067
f (m/n) = 100/1500 = 0.067
f (n/n) = 1300/1500 = 0.867

Next calculate allele frequencies:
f (n) = p = f (n/n) + 0.5 f (m/n) = 13/15 + 0.5x(1/15) = 0.9
f (m) = q = f (m/m) + 0.5 f (m/n) = 1/15 + 0.5x(1/15) = 0.1

Next calculate expected genotype frequencies:
$p^2 = 0.81$; $2pq = 0.18$; $q^2 = 0.01$

The expected genotype frequencies do not match the observed genotype frequencies, thus the population is not in H-W equilibrium.

(b) If the M and N blood groups were not co-dominant, would it be possible to determine whether the population is in Hardy-Weinberg equilibrium? Explain why or why not?

No, because you can't differentiate between homozygotes and heterozygotes.

(c) After one generation of random mating in this village, if the new generation has 2000 people in it, how many will be of each blood type ("M," "N," and "MN")?

Random mating in the population will put the population back in HW equilibrium:
The expected genotypic frequencies for a population with the calculated allele frequencies were calculated in 2a.
M = 0.01 x 2000 = 20; N = 0.81 x 2000 = 1620; MN = 0.18 x 2000 = 360

(d) In the generation of 2000 people from part **(c)**, how many copies of each allele ("L^M" and "L^N") will be present in the population?

There are 4000 alleles in the population (2xN):
f **(L^M) = 0.1x4000 = 400**
f **(L^N) = 0.9x4000 = 3600**

(e) Fill in the following chart to show the number of each type of couple in the generation from part (c), where 2000 people = 1000 couples. Assume that each person mates with one (and only one) other person, and mating choice is random with respect to blood group.

One parent + Other parent	Number of this sort of **couple** in the generation of 1000 couples from part **(c)**
$L^M L^M$ + $L^M L^M$	$1000 \times 0.01^2 = 0.1 \approx 0$
$L^M L^M$ + $L^M L^N$	$2 \times 1000 \times 0.01 \times 0.18 = 3.6 \approx 4$
$L^M L^N$ + $L^N L^N$	$2 \times 1000 \times 0.18 \times 0.81 = 291.6 \approx 292$
$L^M L^N$ + $L^M L^N$	$1000 \times 0.18^2 = 32.4 \approx 32$
$L^M L^M$ + $L^N L^N$	$2 \times 1000 \times 0.01 \times 0.81 = 16.2 \approx 16$
$L^N L^N$ + $L^N L^N$	$1000 \times 0.81^2 = 656.1 \approx 656$

There are 1000 total couples in the population.

For random mating the probability of such a pairing is simply equal to the product of the frequencies of the two individuals. We will assume that the allele frequencies are the same in the population of males and in the population of females.

For example the number of $L^ML^M + L^ML^M$ couples:

The frequency of L^ML^M individuals is:
$f(L^ML^M) = 0.01$ (from part a)

The probability of two such individuals mating is the product of the genotypic frequencies of the two individuals:
$f(L^ML^M \times L^ML^M) = f(L^ML^M) \times F(L^ML^M) = 0.01 \times 0.01 = .0001$

The expected number of couples is thus:
$Num = 1000 \times 0.0001 = 0.1$

For couples comprised of individuals of two different genotypes there are two possible combinations of individuals for each such couple. For instance, for $L^ML^M + L^ML^N$ couples, dad could be M and mom could be MN, or mom could be M and dad could be MN. Thus a factor of 2 is used. For example the number of $L^ML^M + L^ML^N$ couples:

The frequency of individuals is:
$2 \times \#couples \times f(L^ML^M) \times (L^ML^N)$
$2 \times 1000 \times 0.01 \times 0.18 = 3.6 \approx 4$

(f) Imagine a gene in the human population that has a mutation rate = 10^{-5} per generation, and the mutant form of the gene has a selective disadvantage, S = 0.1.

- If the mutation is autosomal recessive, calculate 1) the steady-state frequency of affected individuals in the population and 2) the fraction altered alleles that are due to new mutations.

    1) From lecture at steady state, for an autosomal recessive disorder
    $q^2 = m/S = 10^{-5}/0.1$
    $q = 10^{-2}$
    For an autosomal recessive disorder, the frequency of affected individuals = $f(aa) = q^2 = 10^{-4}$
    2) The frequency of alleles added each generation by mutation is 10^{-5}.
    Thus the fraction of alleles due to new mutations is:
    fraction new = (frequency new alleles)/(steady state frequency) = m/q
    fraction new = $10^{-5}/10^{-2} = 10^{-3}$

- If the mutation is autosomal dominant, calculate 1) the steady-state frequency of affected individuals in the population and 2) the fraction altered alleles that are due to new mutations.

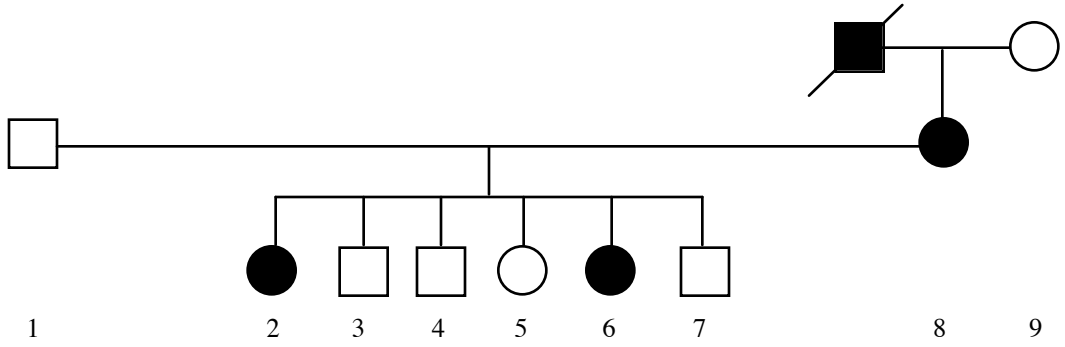
    1) From lecture at steady state, for an autosomal dominant disorder
    $q = m/S = 10^{-5}/0.1$
    $q = 10^{-4}$
    For an autosomal dominant disorder, the frequency of affected individuals = $f(Aa) + f(AA) = 2pq + q^2 = 2X10^{-4}$
    2) The frequency of alleles added each generation by mutation is 10^{-5}.
    fraction new = $10^{-5}/10^{-4} = 10^{-1}$

- If the mutation is X-linked recessive, calculate 1) the steady-state frequency of affected individuals in the population and 2) the fraction altered alleles that are due to new mutations.

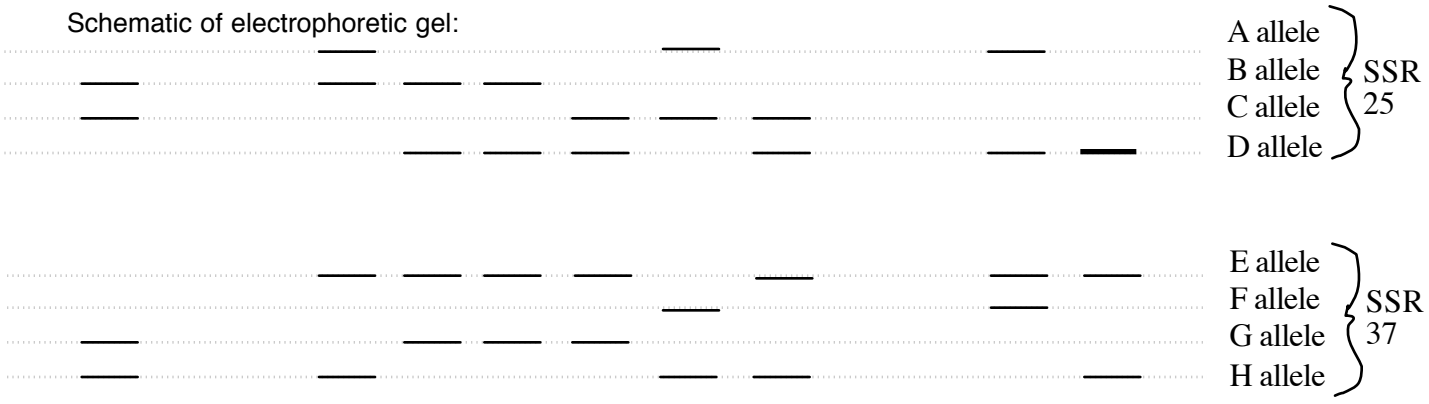
    1) From lecture at steady state, for an X-linked recessive disorder
    $q = 3m/S = 3x10^{-5}/0.1$
    $q = 3x10^{-4}$
    For an X-linked recessive disorder, the frequency of affected individuals = 1/2 f(affected males) + 1/2 f(affected females) = 1/2 $f(X^aX^a)$ + 1/2 $f(X^aY) = 1/2q^2 + 1/2q = 1.5X10^{-4}$
    2) The frequency of alleles added each generation by mutation is 10^{-5}.
    fraction new = $10^{-5}/3x10^{-4} = 10^{-1}/3$

3. The following is a pedigree for an autosomal dominant trait. This trait is caused by the Q allele. Thus individual 1 is "qq" and individual 8 is "Qq". You want to map the locus responsible for the trait with respect to two SSRs.

PLEASE NOTE: For individuals 2-7, only the alleles inherited from the father are shown. You will have to determine the alleles inherited from the mother later in the problem. Both the maternal and paternal alleles of SSR 25 and 37 are given for individuals 1, 8, and 9.



Schematic of electrophoretic gel:



Individual

	2	**3**	**4**	**5**	**6**	**7**
paternally inherited allele at the Q locus	**q**	**q**	**q**	**q**	**q**	**q**
paternally inherited allele at SSR 25	**B**	**B**	**B**	**C**	**C**	**C**
paternally inherited allele at SSR 37	**H**	**G**	**G**	**G**	**H**	**H**

} to be filled in for part **(a)**

Individual

	2	**3**	**4**	**5**	**6**	**7**
maternally inherited allele at the Q locus	**Q**	**q**	**q**	**q**	**Q**	**q**
maternally inherited allele at SSR 25	**A**	**D**	**D**	**D**	**A**	**D**
maternally inherited	**E**	**E**	**E**	**E**	**F**	**E**

} to be filled in for part **(c)**

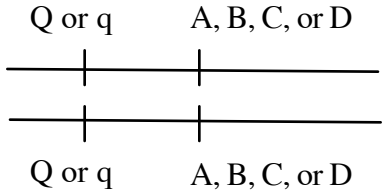
→ to be filled in for part **(g)**

allele at SSR 37						

(a) Fill in the chart on page 4 with the paternally inherited alleles at all three loci.

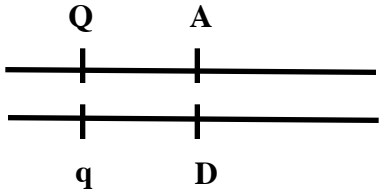
The father is qq, thus he must give each of his children a q allele. For SSR25 he is heterozygous with a B-C genotype; the SSR25 allele given to each child can be determined from the gel. For SSR37 he is heterozygous with a G-H genotype; the SSR25 allele given to each child can be determined from the gel.

(b) Draw all phases of individual 8 with respect to the Q locus and SSR 25 that are possible given everything you know about her. Make sure to draw phases using the proper notation:

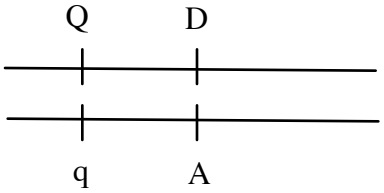


Based on the genotype of individual 8 (Q/q; A/D) there are two possible phases:

1



2



However, we know the genotype of individual 9 (q/q; D/D). Indivdual 8 must have inherited a q-D chromosome from individual 9 and a Q-A chromosome from her father. **Thus the correct phase is phase 1.**

(c) A human geneticist now tells you that the total probability of seeing the maternally inherited alleles of the Q and SSR 25 loci in individuals 2-7 equals $(0.5)^5 * (0.0)^0$, given that the Q and SSR 25 loci are linked at $\theta = 0$. Given this information, fill in chart above **and** draw in the maternally inherited bands in the above gel. (You have the SSR data for all the children, so fill all columns in.)

For the Q locus we know that affected individuals must have received a Q allele from their mother since they can only inherit a q allele from their father. Unaffected individuals must be qq, therefore they must have inherited a q from their mother.

If the Q locus and SSR25 are linked at $\theta = 0$ then no recombination between the two loci will take place and the maternal allele combinations (QA & qD) cannot be shuffled. Thus, all affected individuals will inherit allele A and all normal individuals will inherit allele D

(d) Calculate the LOD score at $\theta = 0$ for SSR 25 and the Q locus. Show all calculations.

$$LOD_{\theta = 0} = \log \frac{\text{p (data having arisen given that SSR25 and Q are linked at } \theta = 0)}{\text{p (data having arisen given that SSR25 and Q are not linked)}}$$

P data having arisen if SSR25 and Q are unlinked = 0.25^5

Probability of getting any combination of alleles by chance = 0.25, for five individuals this is 0.25^5. We're ignoring one because of the twins.

P data having arisen given linked at $0 = 0.5^5 x 0^0$

Probability of getting either QA or qD is 0.5 since the maternal allele combinations aren't shuffled. Probability of getting a recombinant is 0.

We have five parental types and no recombinants, therefore the probability of seeing these combinations is $0.5^5 x 0^0$.

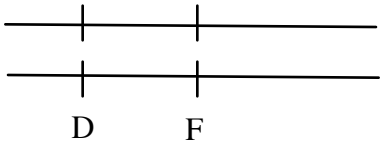
$LOD_{\theta = 0}$ = log ($0.5^5 x 0^0 / 0.25^5$) = 1.5

(e) Individuals 1 and 8 had six children and the human geneticist has taken the data for all six into account correctly in part **(b)**. Knowing this, give a brief explanation for why the human geneticist only included five meioses in the calculation of the LOD score shown above. Assume NO new mutations. Go back and make sure your data table and gel are consistent with your explanation.

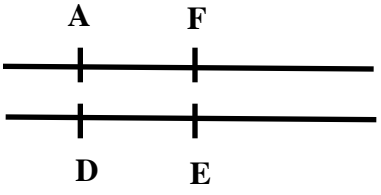
One possibility is that we are only looking at five meiotic events rather than six. This could happen if individuals 3 and 4 are identical twins, thus both were produced by a single meiosis.

(f) You now switch to mapping the two SSRs with respect to each other. Given everything you know about the mother (individual 8), draw all possible phases of the mother with respect to SSR 37 and SSR 25. Make sure to draw the phases using the proper notation.

1)



2)



Based on the genotype of individual 8 (A/D; E/F) there are two possible phases:
However, we know the genotype of individual 9 (D/D; E/H). Individual 8 must have inherited a D-E chromosome from individual 9 and an A-F chromosome from her father.

Thus the correct phase is phase 2

(g) The human geneticist now tells you that the total probability of seeing the maternally inherited alleles of the SSR 25 and SSR 37 loci in individuals 2-7 equals $(0.45)^4$ * $(0.05)^1$, given that the two SSRs are linked at $\theta = 0.1$. Knowing this, draw in the maternally inherited bands of SSR 37 in the above gel. (You have the SSR data for all the children, so fill all columns in. Make sure you fill in the gel in a way that is consistent with the fact that the human geneticist only included 5 meioses in the calculation.)

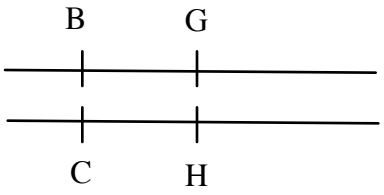
One of the meioses produced a recombinant gamete (i.e. A-E or D-F). We have no way of knowing which child is recombinant *a priori*, so we will just randomly assign one as a recombinant (the chart is filled out as if 2 is a recombinant). Thus there are five equally good solutions to this part.

	2	**3**	**4**	**5**	**6**	**7**
2 recombinant:	**E**	**E**	**E**	**E**	**F**	**E**
3 & 4 recombinant:	**F**	**F**	**F**	**E**	**F**	**E**
5 recombinant:	**F**	**E**	**E**	**F**	**F**	**E**
6 recombinant:	**F**	**E**	**E**	**E**	**E**	**E**
7 recombinant:	**F**	**E**	**E**	**E**	**F**	**F**

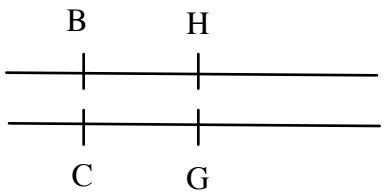
The bands are drawn on the gel consistent with the table (2 is a recombinant).

(h) Given everything you know about the father (individual 1), draw all possible phases of the father with respect to SSR 37 and SSR 25. Make sure to draw the phases using the proper notation.

1)



2)



Based on the genotype of individual 1 (B/C; G/H) there are two possible phases.
Since we don't have any information about the configuration of alleles in the parents of individual 1 we cannot eliminate one of the phases.

(i) Calculate the LOD score at $\theta = 0.1$ for the two SSRs. Make sure to incorporate the data from both the mother **and** the father, because both are informative parents regarding the inheritance of alleles at these two loci. Show all calculations.

We are looking at meioses in two parents, the mother and the father, so we will calculate two separate LOD scores that will be additive to give us a final LOD score.

Maternal alleles:

P data having arisen because the two SSRs are unlinked = 0.25^5

Probability of getting any combination of alleles by chance = 0.25, for five individuals this is 0.25^5. We're ignoring one because of the twins.

P data having arisen given that the two SSRS are linked at $\theta = 0.1$

$= 0.45^4 \text{x} 0.05^1$

Given that the alleles are linked at $\theta = 0.1$, the probability of getting a parental type is 0.9 and the probability of getting a recombinant type is 0.1. Thus the probability of getting A-F is 0.45, and 0.45 for D-E. The probability of getting A-E is 0.05, and 0.0.5 for D-F. We see 1 recombinant and 4 parentals therefore the maternal probability is $(0.45)^4 \text{x} (0.05)^1$.

$LOD_{\theta = 0.1} = \log (0.45^4 \text{x} 0.05^1 / 0.25^5) = 0.322$

Paternal alleles:

P data having arisen because the two SSRs are unlinked = 0.25^5

P data having arisen given that the two SSRS are linked at $\theta = 0.1$

$= 1/2(0.45^3 \text{x} 0.05^2) + 1/2(0.45^2 \text{x} 0.05^3)$

We can't determine the phase for the father thus we must take both into account. *A priori* the probability of each phase is 1/2. For phase one there are 2 recombinants and 3 parentals, for phase two there are 3 recombinants and 2 parentals.

$LOD_{\theta = 0.1} = \log ([1/2(0.45^3 \text{x} 0.05^2) + 1/2(0.45^2 \text{x} 0.05^3)] / 0.25^5) = -0.887$

To calculate the total LOD score, sum the LOD score from the father and the LOD score from the mother.

$LOD_{\theta = 0.1}$ = $LOD_{\theta = 0.1}$ maternal + $LOD_{\theta = 0.1}$ paternal = -0.565

2004 7.03 Problem Set 7

NO DUE DATE This problem set is to provide practice on concepts from lectures 31-33.

1. For each child listed below, list **all** of the following non-disjunction events that could have led to the creation of a gamete that could have produced this child. Your choices are:

non-disjunction during gamete formation… in the mother during meiosis I
in the mother during meiosis II
in the father during meiosis I
in the father during meiosis II

(a) an XO female child

(b) an XYY male child

(c) an XXX female child

(d) an XXY male child

2. Predict the phenotypic concordance rates of monozygotic (MZ) and dizygotic (DZ) twins for the following rare traits on which environmental effects are negligible. Assume that each trait shows complete penetrance and does not affect fitness of those with the trait.

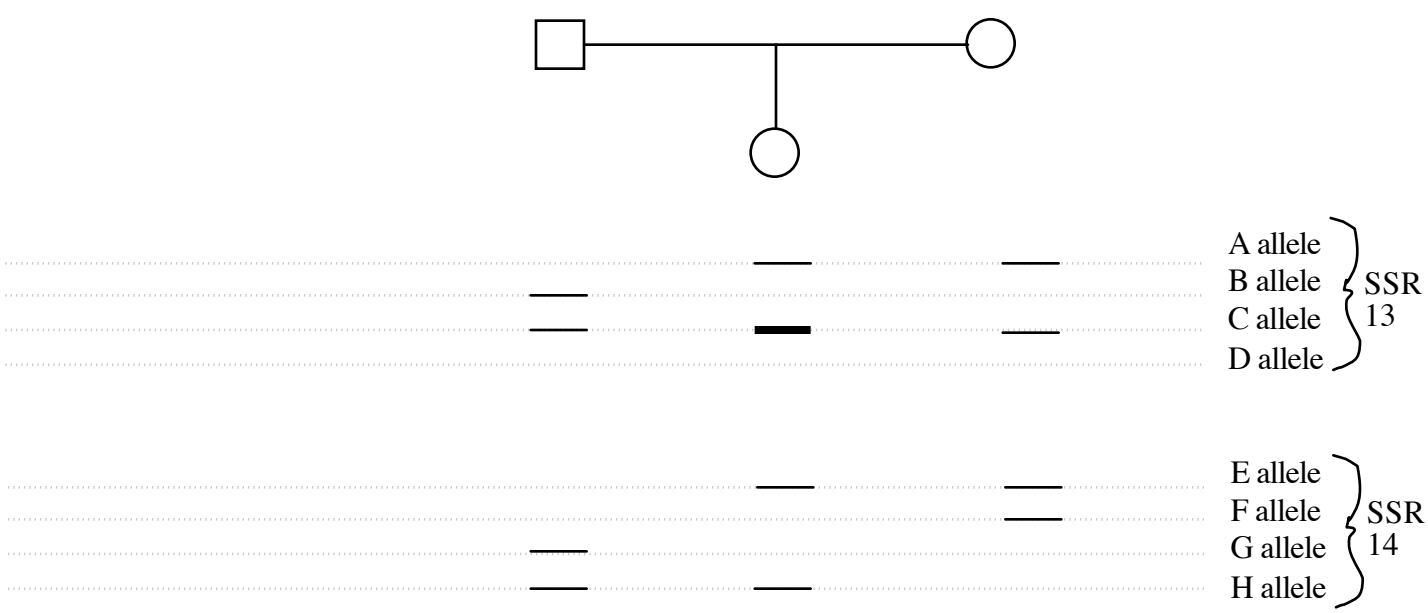
What are the concordance rates for MZ and DZ twins…

(a) … if at least one twin in each set is expressing a specific autosomal trait caused by the presence of dominant alleles at two unlinked loci (ie. only "A_B_" individuals show the phenotype)?

(b) … if at least one twin in each set is expressing a specific autosomal trait caused by the presence of recessive alleles at two unlinked loci (ie. only "aabb" individuals show the phenotype)?

(c) …if at least one twin in each set is a man expressing a specific X-linked recessive trait?

3. The following is a pedigree showing a couple that has a child with trisomy of chromosome 21. The schematic of a gel is shown below, which reveals the genotypes of each member of the family at two different SSRs found on chromosome 21.



For parts **(a)** and **(b)**, assume that the band in the gel corresponding to the H allele in the child is significantly more intense than the band corresponding to the E allele.

(a) Can you conclude in which parent the non-disjunction event occurred? If so, in which parent did the non-disjunction event occur? Explain your answer.

(b) Can you conclude whether non-disjunction occurred in meiosis I or meiosis II? If so, in which meiotic division did the non-disjunction event occur? Explain your answer.

For parts **(c)** through **(g)**, assume that the band in the gel corresponding to the E allele in the child is significantly more intense than the band corresponding to the H allele.

(c) Can you conclude in which parent the non-disjunction event occurred? If so, in which parent did the non-disjunction event occur? Explain your answer.

(d) Assuming that SSR13 is centromere-linked, in which meiotic division did the non-disjunction event occur? Explain your answer.

(e) Assuming that SSR13 is centromere-linked, draw the meiosis that created the gamete that led to the production of the child shown in the pedigree. Assume the parents' phases are as shown below.

In each cell, draw chromosome 21 AND another chromosome of a different size that undergoes meiotic chromosome segregation normally. Please label each SSR allele and the centromere on each homolog of chromosome 21. Show the initial cell and then that cell after having undergone each of the following steps of the meiotic cell cycle in the following order: DNA replication, chromosome alignment during metaphase I (please indicate where any recombination events occurred), and chromosome alignment during metaphase II. Then show the four final products of the meiosis. (Please indicate the gamete that led to the creation of the child with trisomy 21 with a star.)

…more space for part **(e)**…

(f) Now assume that SSR14 is centromere-linked (instead of SSR13). In which meiotic division did the non-disjunction event occur? Explain your answer.

(g) Assuming SSR14 is centromere-linked (instead of SSR13), draw the meiosis that created the gamete that led to the production of the child shown in the pedigree. Assume that the parents' phases are the same as shown in part **(e)**. Follow the instructions from part **(e)**.

Solution to 7.03 Problem Set 7

1. For each child listed below, list **all** of the following non-disjunction events that could have led to the creation of a gamete that could have produced this child. Your choices are:
non-disjunction during gamete formation… in the mother during meiosis I
in the mother during meiosis II
in the father during meiosis I
in the father during meiosis II

(a) an XO female child

in the mother during meiosis I
in the mother during meiosis II
in the father during meiosis I
in the father during meiosis II

(b) an XYY male child

in the father during meiosis II

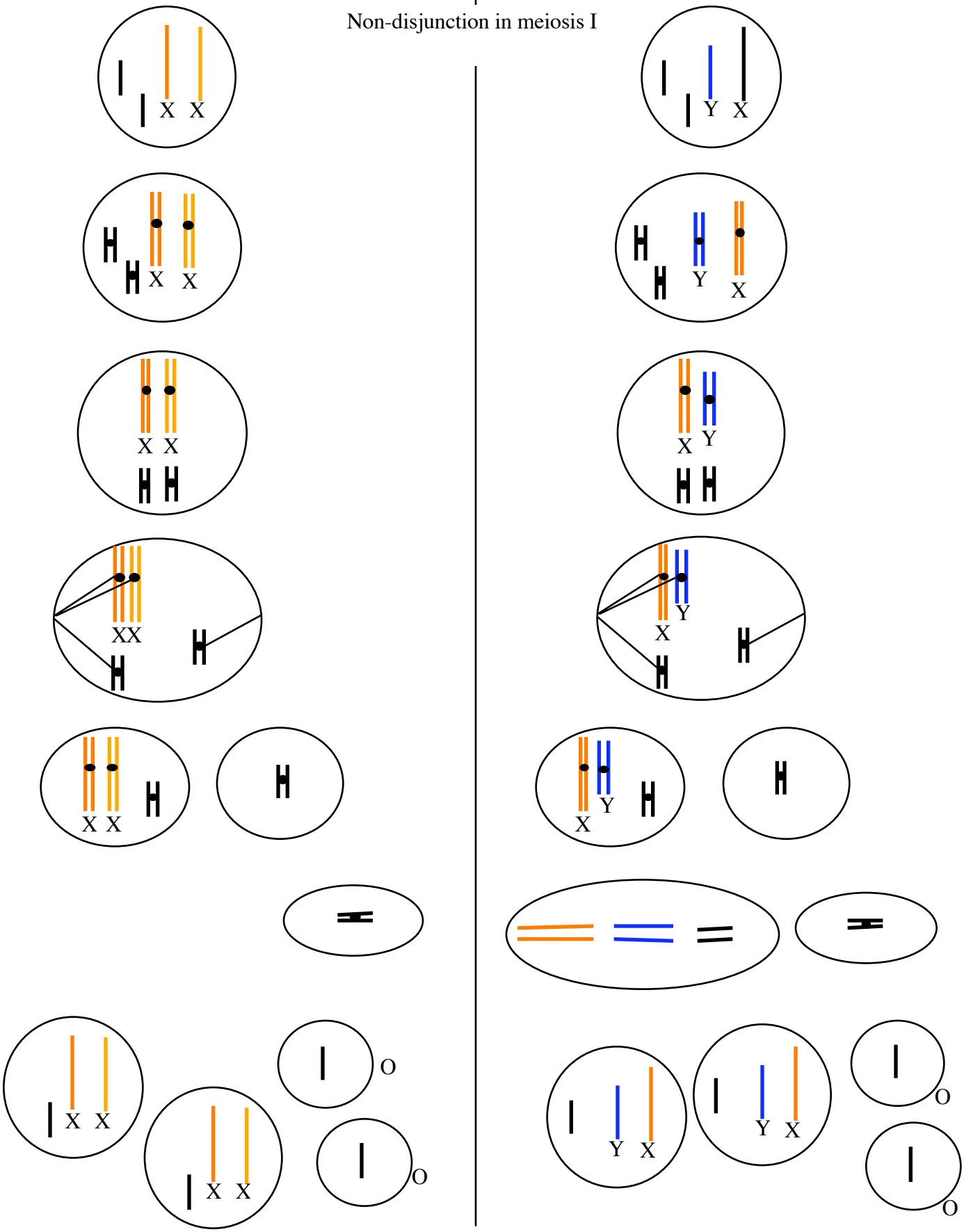
(c) an XXX female child

in the mother during meiosis I
in the mother during meiosis II
in the father during meiosis II

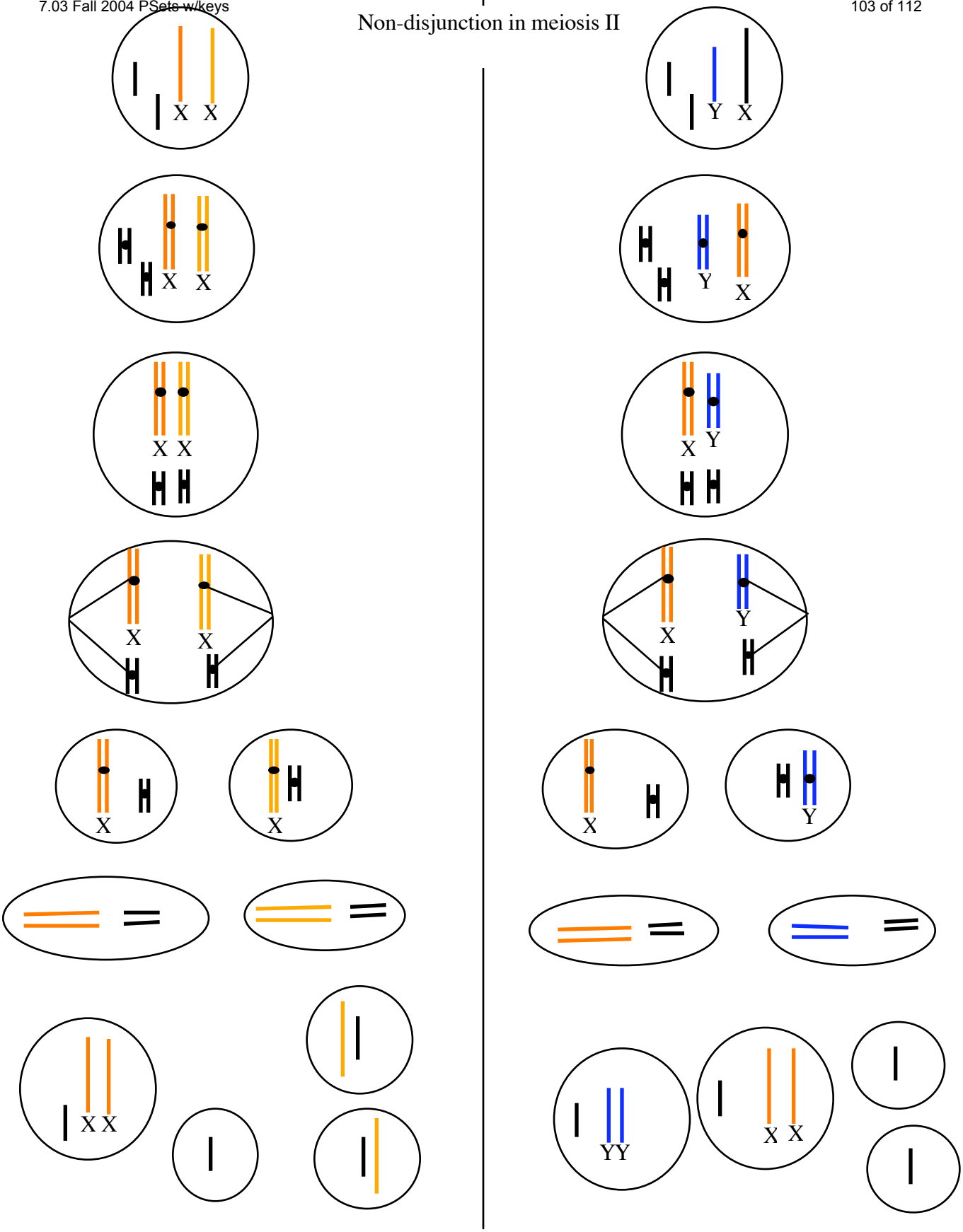
(d) an XXY male child

in the mother during meiosis I
in the mother during meiosis II
in the father during meiosis I

Non-disjunction in meiosis I



Non-disjunction in meiosis II



2. Predict the phenotypic concordance rates of monozygotic (MZ) and dizygotic (DZ) twins for the following rare traits on which environmental effects are negligible. Assume that each trait shows complete penetrance and does not affect fitness of those with the trait.

What are the concordance rates for MZ and DZ twins…

(a) … if at least one twin in each set is expressing a specific autosomal trait caused by the presence of dominant alleles at two unlinked loci (ie. only "A_B_" individuals show the phenotype)?

MZ – 1

DZ – 1/4. One twin is A_B_, and the alleles for this trait are rare, so the parents are most likely Aabb and aaBb. The DZ twin's sibling has 1/2 chance of inheriting A and 1/2 chance of inheriting B.

(b) … if at least one twin in each set is expressing a specific autosomal trait caused by the presence of recessive alleles at two unlinked loci (ie. only "aabb" individuals show the phenotype)?

MZ – 1

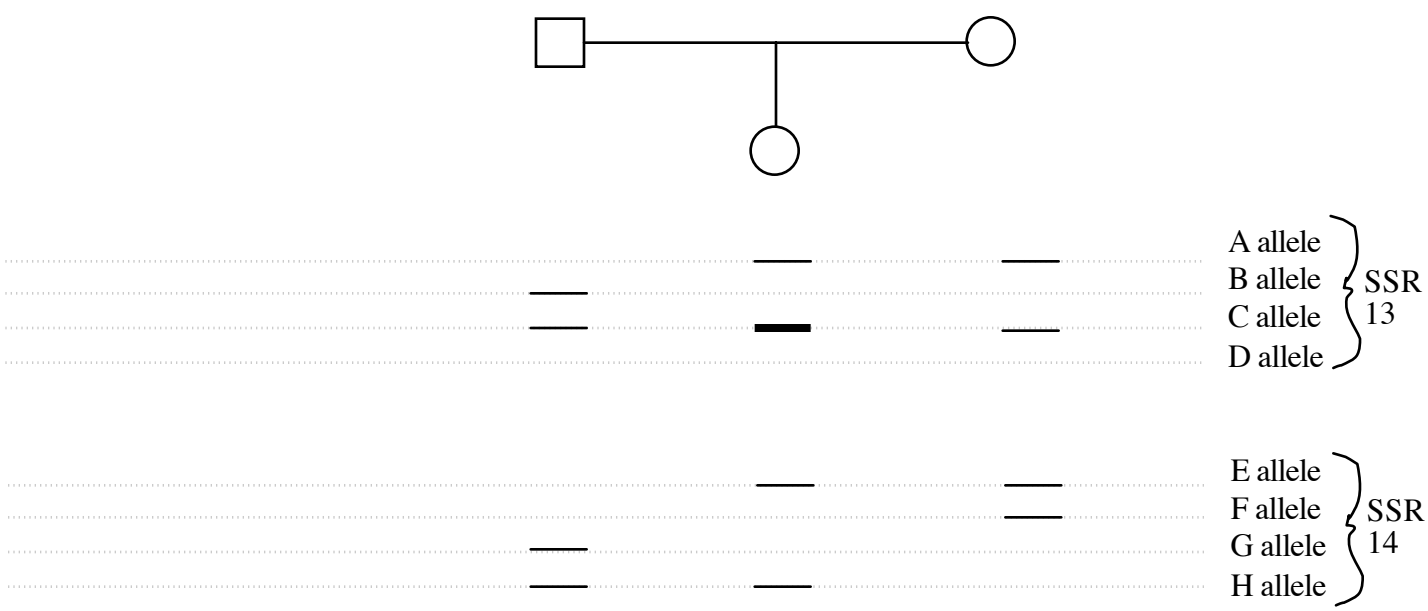
DZ – 1/16. **One twin is aabb and the alleles for this trait are rare, so the parents are most likely AaBb and AaBb. There is 1/4 chance that mom produces an egg with genotype ab and 1/4 chance that dad produces a sperm with genotype ab.**

(c) …if at least one twin in each set is a man expressing a specific X-linked recessive trait?

MZ – 1

DZ – 1/4. **One twin is X^aY and the alleles for this trait are rare, so the parents are most likely X^AY and X^AX^a. There is 1/2 chance that mom gives X^a to her son, so 1/2 of the boys will show the trait. However, 1/2 of the time the other twin is a girl (X^AX^A or X^AX^a). None of the girls will show the trait.**

3. The following is a pedigree showing a couple that has a child with trisomy of chromosome 21. The schematic of a gel is shown below, which reveals the genotypes of each member of the family at two different SSRs found on chromosome 21.



For parts **(a)** and **(b)**, assume that the band in the gel corresponding to the H allele in the child is significantly more intense than the band corresponding to the E allele.

(a) Can you conclude in which parent the non-disjunction event occurred? If so, in which parent did the non-disjunction event occur? Explain your answer.
The dark band corresponds to two copies of the H allele, so the child must have two copies of Dad's chromosome #21.

(b) Can you conclude whether non-disjunction occurred in meiosis I or meiosis II? If so, in which meiotic division did the non-disjunction event occur? Explain your answer.
The child has two copies of both C at SSR13 and two copies of H at SSR 14. Thus you expect non-disjunction occurred in meiosis II. To get this chromosomal arrangement due to failure of meiosis I, you would have to have recombination at both loci, which is unlikely.

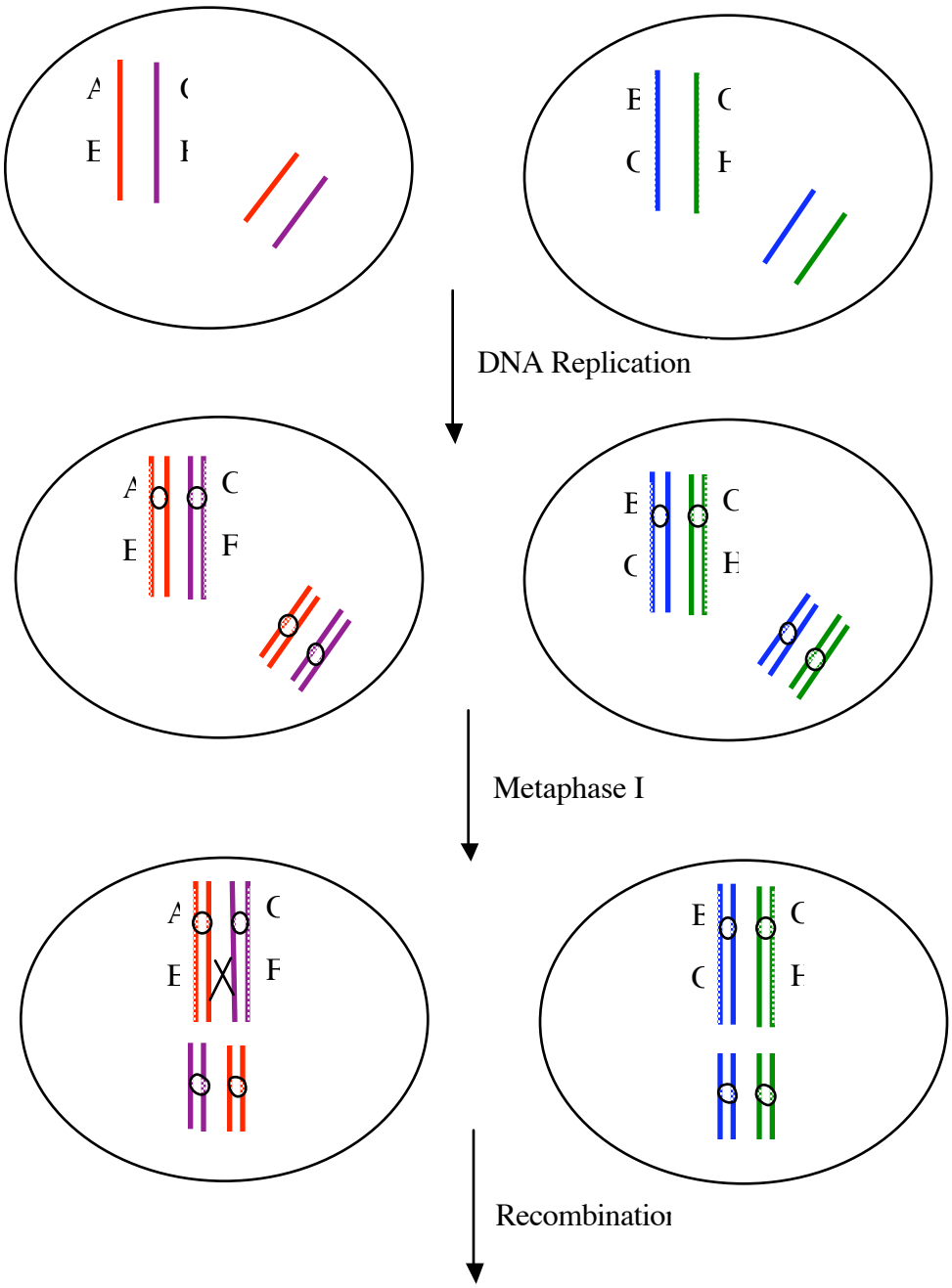
For parts **(c)** through **(g)**, assume that the band in the gel corresponding to the E allele in the child is significantly more intense than the band corresponding to the H allele.

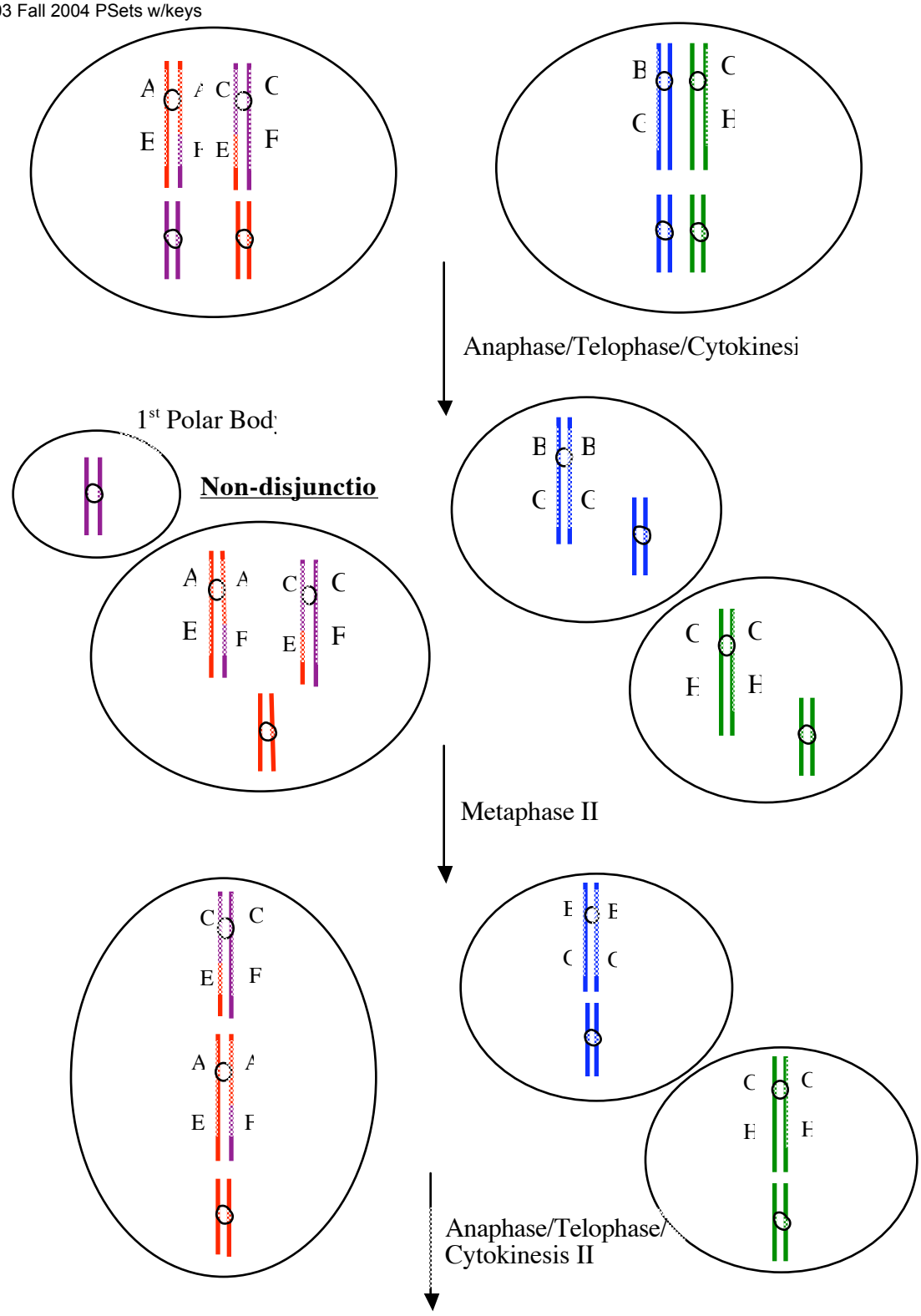
(c) Can you conclude in which parent the non-disjunction event occurred? If so, in which parent did the non-disjunction event occur? Explain your answer.
The dark band corresponds to two copies of the E allele, only Mom could give two E alleles of SSR13.

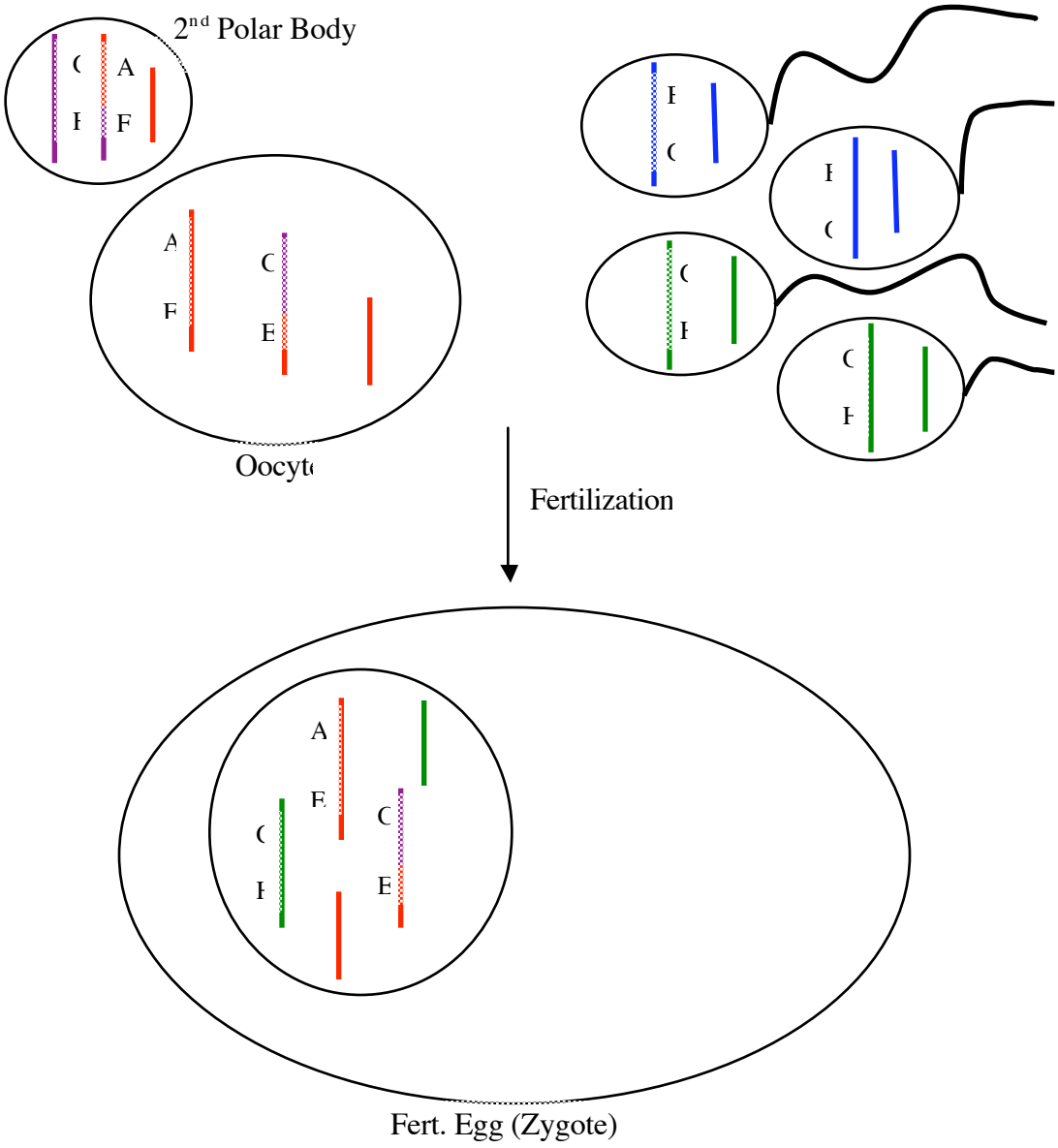
(d) Assuming that SSR13 is centromere-linked, in which meiotic division did the non-disjunction event occur? Explain your answer.
Meiosis I. By assuming that SSR13 is centromere-linked, we conclude that no recombination occurs at SSR13. Only failure of homologous chromosome to separate at meiosis I can give two different centromere-linked alleles.

(e) Assuming that SSR13 is centromere-linked, draw the meiosis that created the gamete that led to the production of the child shown in the pedigree. Assume the parents' phases are as shown below.

In each cell, draw chromosome 21 AND another chromosome of a different size that undergoes meiotic chromosome segregation normally. Please label each SSR allele and the centromere on each homolog of chromosome 21. Show the initial cell and then that cell after having undergone each of the following steps of the meiotic cell cycle in the following order: DNA replication, chromosome alignment during metaphase I (please indicate where any recombination events occurred), and chromosome alignment during metaphase II. Then show the four final products of the meiosis. (Please indicate the gamete that led to the creation of the child with trisomy 21 with a star.)


DNA Replication
Metaphase I
Recombinatio

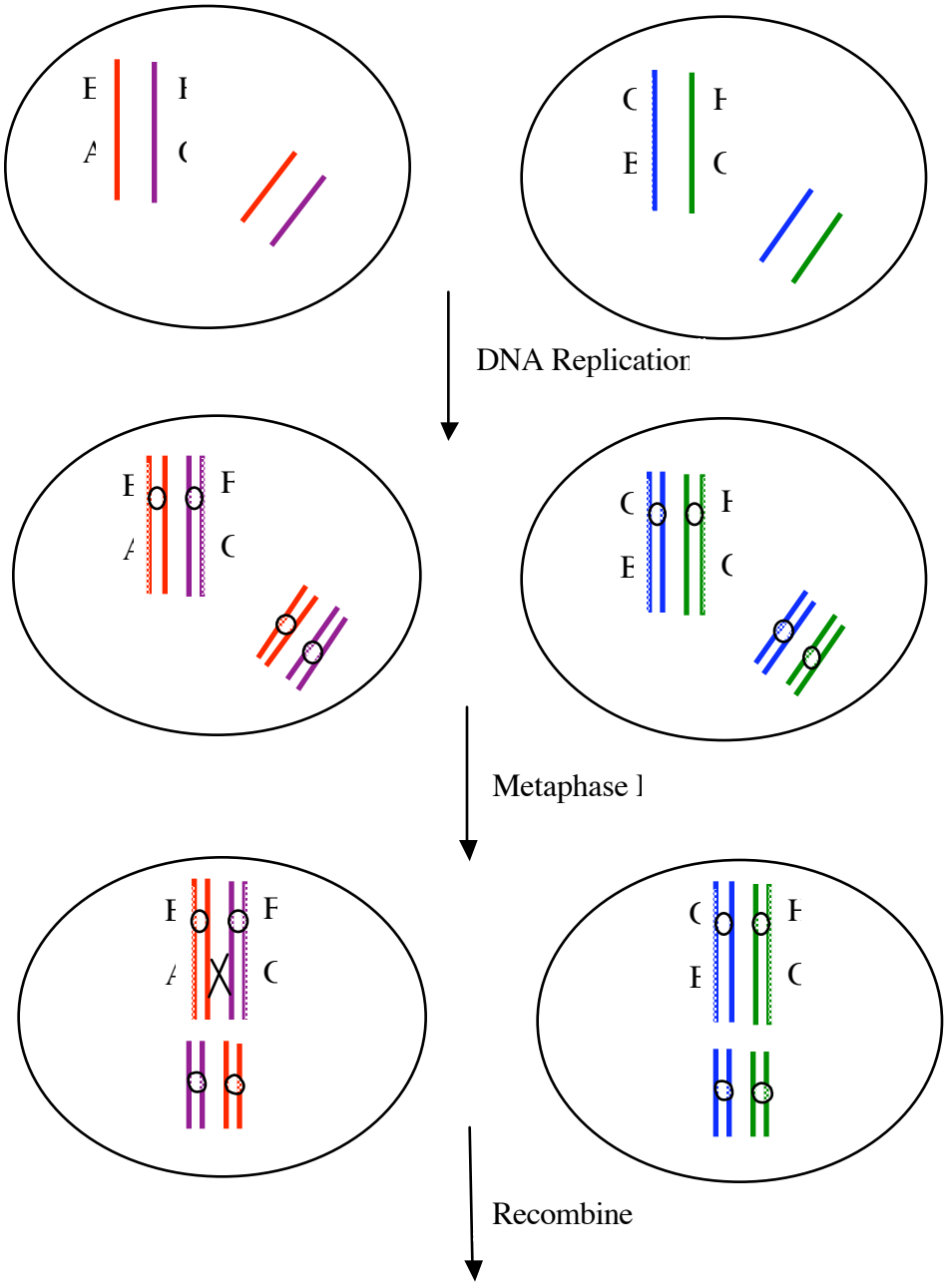

Anaphase/Telophase/Cytokinesi
1st Polar Bod
Non-disjunctio
Metaphase II
Anaphase/Telophase/
Cytokinesis II

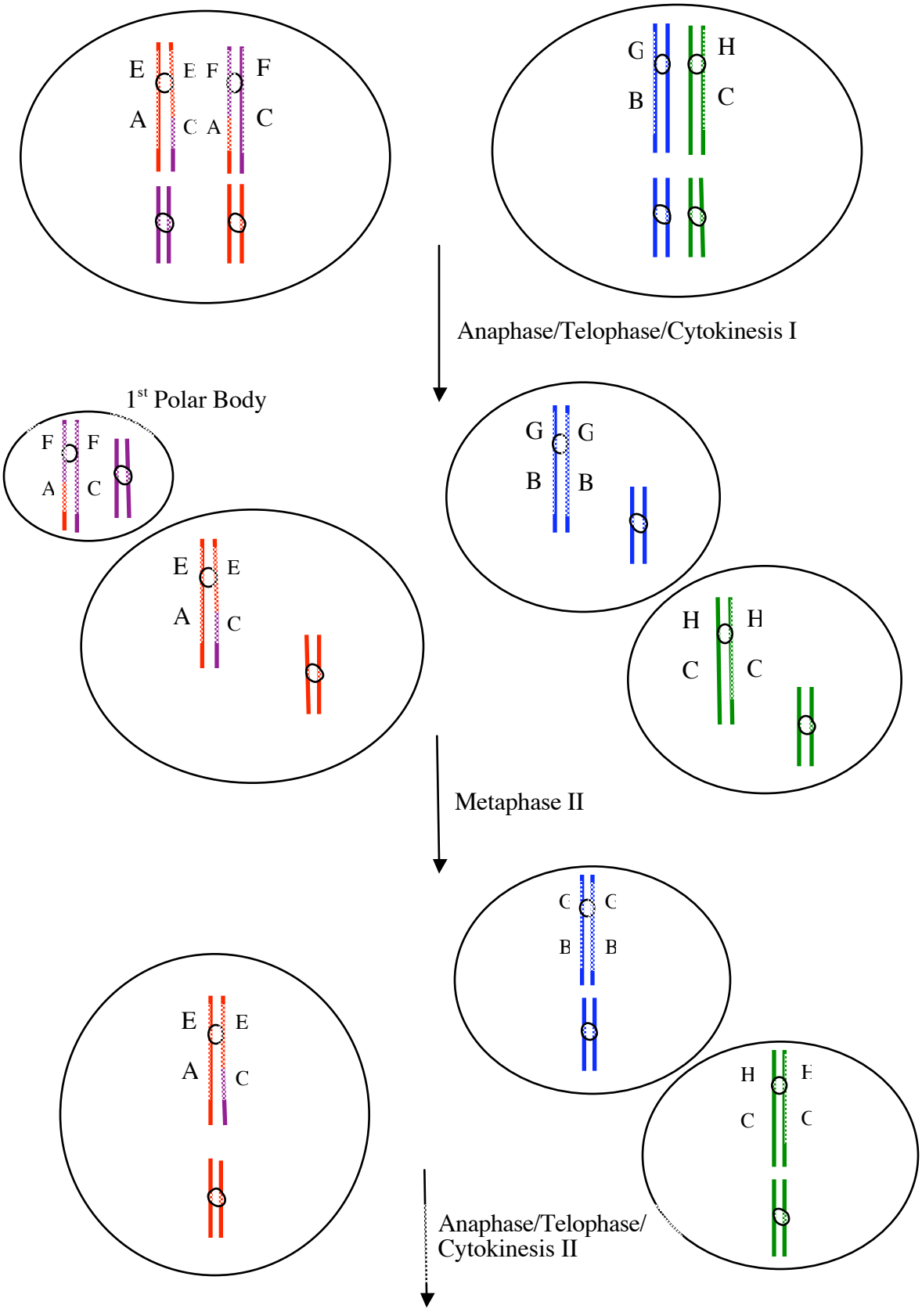

2nd Polar Body
Oocyte
Fertilization
Fert. Egg (Zygote)

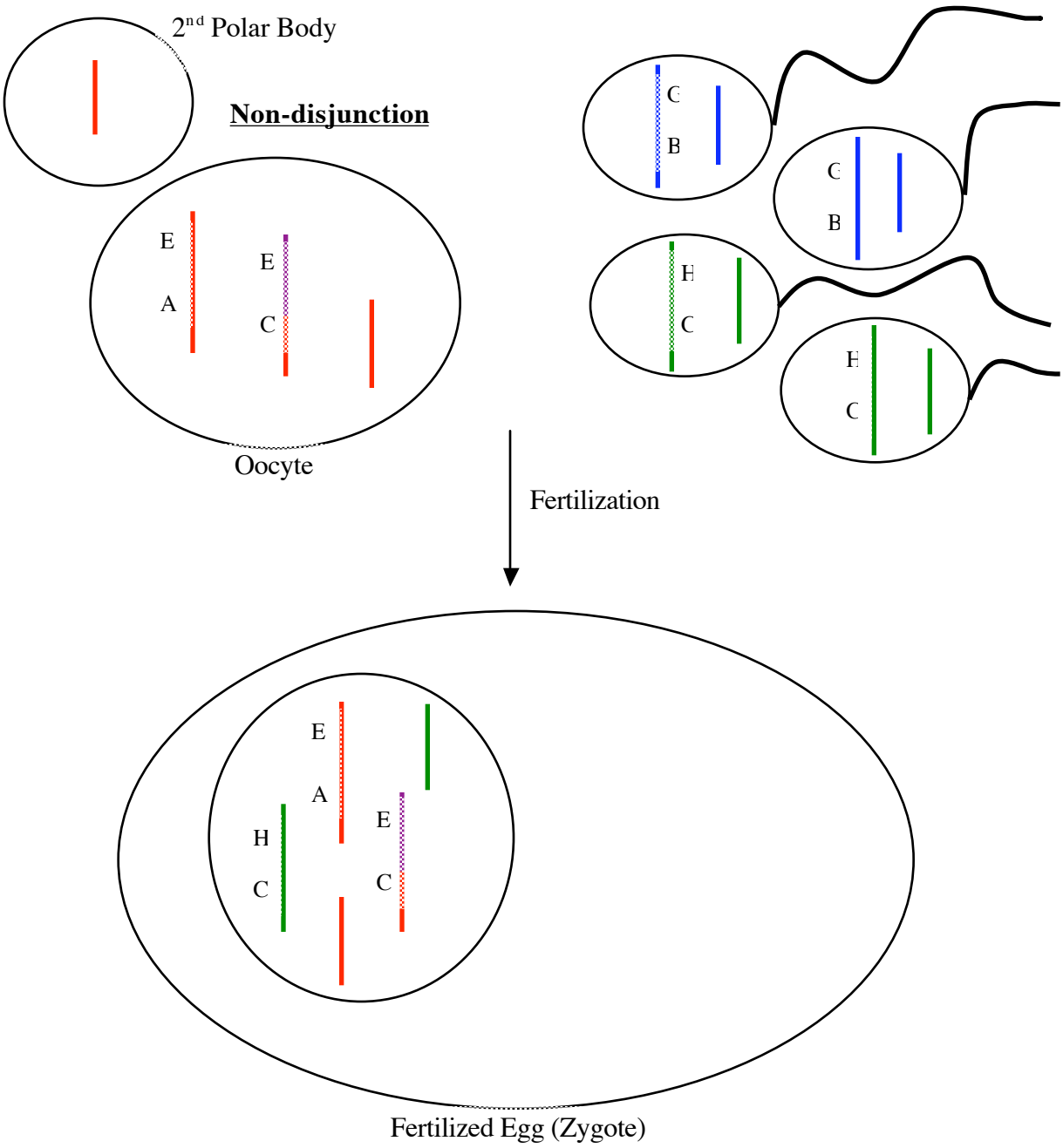
(f) Now assume that SSR14 is centromere-linked (instead of SSR13). In which meiotic division did the non-disjunction event occur? Explain your answer.

MII. Recombination did <u>not</u> occur at SSR14, and chills has two copies of the E allele at SSR14. So sister chromatids failed to separate at meiosis II.

(g) Assuming SSR14 is centromere-linked (instead of SSR13), draw the meiosis that created the gamete that led to the production of the child shown in the pedigree. Assume that the parents' phases are the same as shown in part **(e)**. Follow the instructions from part **(e)**.




E E F F
A C A C
G H
B C
Anaphase/Telophase/Cytokinesis I
1st Polar Body
F F
A C
G G
B B
E E
A C
H H
C C
Metaphase II
C C
B B
E E
A C
H F
C C
Anaphase/Telophase/
Cytokinesis II


2nd Polar Body
Non-disjunction
E
A
E
C
Oocyte
C
B
C
B
H
C
H
C
Fertilization
E
A
H
C
E
C
Fertilized Egg (Zygote)