Problem Sets

Fall 1999

7.03 Problem Set 1

Due before 5 PM on Thursday, September 23

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

1. (a) You have isolated five new Arg⁻ mutants in yeast. After obtaining versions of each mutant in mating type **a** and in mating type α, you perform all of the possible pairwise matings shown in the table below. A "+" at the intersection of the two parental strains indicates that the diploid can grow without arginine added to the medium, whereas a "–" indicates that the diploid can't grow without arginine.

mutants of mating type **a** \ mutants of mating type α	1	2	3	4	5
1	–	+	–	+	–
2		–	–	+	+
3			–	–	–
4				–	+
5					–

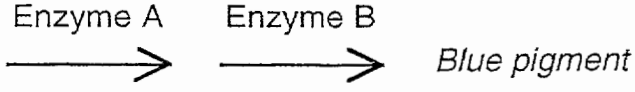
Give as much information as you can about your new Arg⁻ mutations. Indicate which mutations are dominant and which are recessive, also state how many genes are represented and which mutations lie the same gene. Assume each strain carries only a single Arg⁻ mutation.

(b) The amino acid permease that allows cells to take up arginine will also transport canavanine, which is a toxic analog of arginine. Mutants that interfere with transporter activity can block canavanine uptake and will therefore allow cells to grow in the presence of an amount of canavanine that would kill a wild-type yeast cell. You have isolated a set of five canavanine resistant mutants. As before, you obtain versions of each mutant in mating type **a** and mating type α, and you perform all of the possible pairwise matings shown in the table below. In this table a "+" at the intersection of the two parental strains indicates that the diploid can grow in the presence of high levels of canavanine, whereas a "–" indicates that the diploid is as sensitive to canavanine as wild-type.

mutants of mating type a	mutants of mating type α				
	1	2	3	4	5
1	+	+	−	+	−
2		+	−	+	−
3			+	+	−
4				+	+
5					+

Give as much information as you can about your new canavanine resistant mutations. Indicate which mutations are dominant and which are recessive also state how many genes are represented and which mutations lie the same gene. Again assume each strain carries only a single mutation.

2. A hypothetical insect has blue eyes, but mutant insects that can not form blue pigment have white eyes. Production of blue pigment involves the activities of enzyme A and enzyme B, both encoded by autosomal genes. First, suppose that the pathway for production of blue pigment involves enzyme A and enzyme B operating in series:

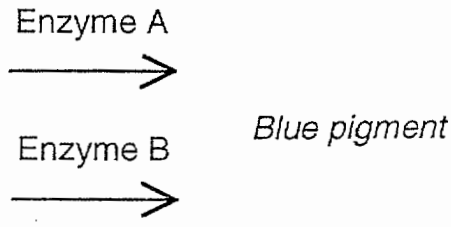


(a) A true breeding strain with a recessive mutation in the gene for enzyme A is crossed to a true breeding strain with a recessive mutation in the gene for enzyme B. Will the resulting F_1 progeny have blue or white eyes? When these F_1 insects are then crossed among themselves, what will the phenotypic ratio of blue to white eyed insects be among the F_2?

(b) A true breeding strain with a dominant mutation in the gene for enzyme A is crossed to a true breeding strain with a dominant mutation in the gene for enzyme B. Will the resulting F_1 progeny have blue or white eyes? When these F_1 insects are then crossed among themselves, what will the phenotypic ratio of blue to white eyed insects be among the F_2?

(c) A true breeding strain with a recessive mutation in the gene for enzyme A is crossed to a true breeding strain with a dominant mutation in the gene for enzyme B. Will the resulting F_1 progeny have blue or white eyes? When these F_1 insects are then crossed among themselves, what will the phenotypic ratio of blue to white eyed insects be among the F_2?

Now suppose that enzyme A and enzyme B act in parallel. That is, there are two different ways to make blue pigment and white-eyed insects only result when the steps carried out by both enzyme A and enzyme B are inactive.

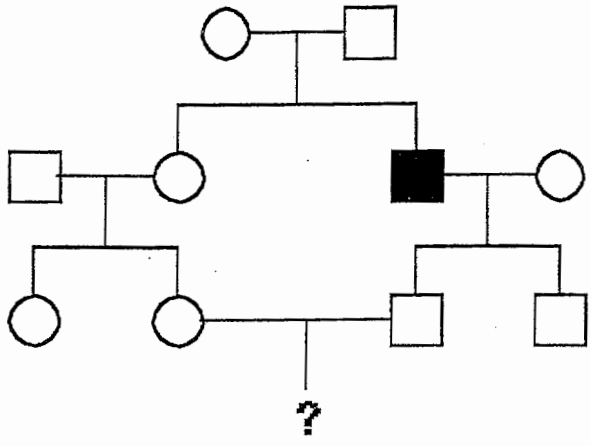


(d) A true breeding strain with a recessive mutation in the gene for enzyme A is crossed to a true breeding strain with a dominant mutation in the gene for enzyme B. Given the parallel pathway model, will the resulting F_1 progeny have blue or white eyes? When these F_1 insects are then crossed among themselves, what will the phenotypic ratio of blue to white eyed insects be among the F_2?

(e) Let's say that you are trying to distinguish the series model in part (c) from the parallel pathway model in part (d), but you decide to look at only eight flies from the F_2 generation. Apply the chi-square test to all nine possible phenotypic ratios for eight flies to determine which observed ratios are consistent with the expected ratios for each of the two models (Use the criteria that the hypothesis can be rejected only if the p value is < 0.05). How many of the possible outcomes will not definitively distinguish the two models?

(f) Finally, suppose that the gene for enzyme A is on the X-chromosome and that the gene for enzyme B is autosomal. Both males and females from a true-breeding strain with a recessive mutation in the gene for enzyme A are crossed to females and males from a true breeding strain with a dominant mutation in the gene for enzyme B. Given the parallel pathway model and the sex linkage of gene A, will the F_1 progeny have blue or white eyes? (Specify males or females). When the F_1 insects are then crossed among themselves, what will the phenotypic ratio of blue to white eyed insects be among males in the F_2? What will the phenotypic ratio be for females in the F_2?

3. You have just been hired as a genetic counselor for a royal family that still engages in a significant amount of inbreeding. As your first assignment you are presented with the following pedigree where the filled symbol represents a male in the royal family who has a rare recessive disease.



Your job is to calculate the probability that the child indicated by **?** will have the disease. To do this, assume that no new mutations arise within the pedigree and that no unrelated individual is a carrier (because this is a very rare disease).

(a) If the disease is caused by an autosomal recessive allele, what is the probability that the child indicated by **?** will have the disease?

(b) If the disease is caused by an X-linked recessive allele, what is the probability that a son will have the disease? What is the probability that a daughter will have the disease?

(c) If the disease is caused by an autosomal recessive allele, and the first child has the disease, what is the probability that the second child will have the disease?

7.03 Problem Set #1 Solutions
FA1999

1a.
Mutant **3** fails to complement with any other strain, but carries only a single mutation. It is likely that this strain carries a **dominant** mutation.
Mutants **1** and **5** are **recessive** mutations in the same complementation group. They are probably alleles of the **same gene**.
Mutant **2** is a **recessive** mutation in its own complementation group, as it complements the rest of the mutants. It is in a **separate gene** from the rest.
Mutant **4** is a **recessive** mutation in its own complementation group, in a **separate gene** from the rest.
Our set of strains represent mutations in **3 or 4 different genes**, with the ambiguity due to mutant 3 being dominant. By this test, we can not determine in which gene mutant 3 is in.

1b.
Mutant **4** fails to complement with any other strain, but carries only a single mutation. It is likely that this strain carries a **dominant** mutation.
Mutants **1** and **2** are **recessive** mutations in the same complementation group. They are probably alleles of the **same gene**.
Mutant **3** is a **recessive** mutation in its own complementation group, as it complements the rest of the mutants. It is in a **separate gene** from the rest.
Mutant **5** is a **recessive** mutation in its own complementation group, in a **separate gene** from the rest.
Our set of strains represent mutations in **3 or 4 different genes**, with the ambiguity due to mutant 4 being dominant. By this test, we can not determine in which gene mutant 4 is in.

2a.
All the F_1 progeny will have **blue** eyes. Both mutations are recessive to the wild-type allele, and thus, the two strains will complement each other in the F_1 generation. Using the punnet square gives a ratio of **9:7** blue to white eyed progeny for the F_2 generation

	AB	Ab	aB	ab
AB	AABB	AABb	AaBB	AaBb
Ab	AABb	AAbb *	AaBb	Aabb *
aB	AaBB	AaBb	aaBB *	aaBb *
ab	AaBb	Aabb *	aaBb *	aabb *

Only progeny with either 2 "a" alleles or 2 "b" alleles (or both) will produce white eyed flies.

2b.

All the F1 progeny will have **white** eyes. In the F_2 generation, blue eyed insects are only produced when no mutant allele is present in the progeny, giving a ratio of **15:1** ~~blue:white~~ W:B eyed progeny.

P(inherit A^{wt}/A^{wt} and inherit B^{wt}/B^{wt}) = P(A^{wt}/A^{wt}) * P(B^{wt}/B^{wt}) = 1/4*1/4 = 1/16

2c.

All the F1 progeny will have **white** eyes. Blue eyed insects in the F_2 generation must be (A/-, B/B), where "A" and "B" are the wild type alleles for enzymes A and B and "a" and "b" are the mutant alleles. ("B" is recessive to "b" for enzyme activity). Using the punnet square gives a ratio of **3:13** blue to white eyed insects in the F_2 generation.

	AB	Ab	aB	ab
AB	AABB *	AABb	AaBB *	AaBb
Ab	AABb	AAbb	AaBb	Aabb
aB	AaBB *	AaBb	aaBB	aaBb
ab	AaBb	Aabb	aaBb	aabb

2d.

All the F_1 progeny will have **blue** eyes. White eyed insects must be (a/a. b/-) (using the same conventions as in 2c). By punnet square, the phenotypic ratio is **13:3** blue to white eyed insects in the F_2 generation (not shown).

2e.

If we use the series model (c), then we expect (3/16 * 8) = 1.5 blue eyed insects and (13/16 * 8) = 6.5 white eyed insects out of 8 progeny.

If we use the parallel model (d), we expect (13/16 * 8) = 6.5 blue eyed insects and (3/16 * 8) = 1.5 white eyed insects out of 8 progeny.

Applying the χ^2 test to the different phenotypic ratios gives the following table (next page):

(degrees of freedom = 1)

If observed: Blue	White	Then	χ^2(series)	p(series)		χ^2(para)	p(para)
0	8		1.846	0.5-0.1		34.67	<0.005
1	7		0.205	0.9-0.5		24.82	<0.005
2	6		0.205	0.9-0.5		16.62	<0.005
3	5		1.846	0.5-0.1		10.05	<0.005
4	4		5.128	0.025-0.01		5.128	0.025-0.01
5	3		10.05	<0.005		1.846	0.5-0.1
6	2		16.62	<0.005		0.205	0.9-0.5
7	1		24.82	<0.005		0.205	0.9-0.5
8	0		34.67	<0.005		1.846	0.5-0.1

We can not distinguish between the two models if we observe **4 white and 4 blue** eyed progeny.

2f.

There are two different crosses possible for the F_1 generation:

♂ (X^a/Y; B/B) x (X^A/X^A; b/b) ♀ or ♂ (X^A/Y; b/b) x (X^a/X^a; B/B) ♀

|| ||

V V

♂ (X^A/Y; B/b) (blue) (X^a/Y; B/b) (white) ♂

♀ (X^A/X^a, B/b) (blue) (X^A/X^a; B/b) (blue) ♀

So, F_1 **males** will be both **blue** and **white** eyed while F_1 **females** will only be **blue** eyed.

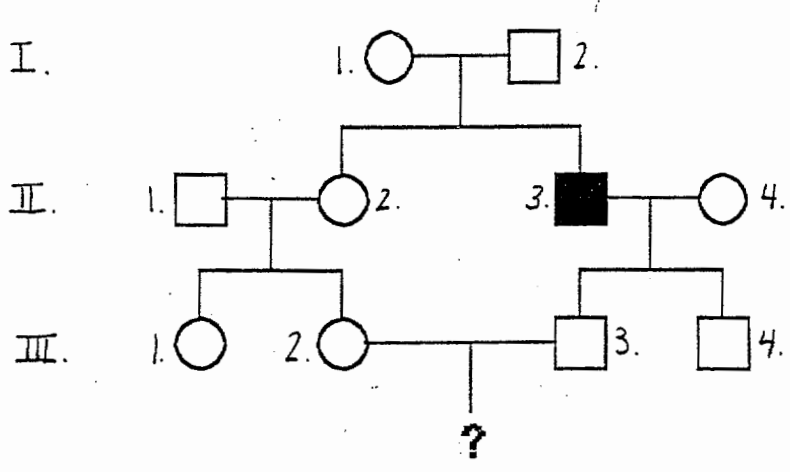
For the F_2 generation, if we cross a F_1 blue eyed male to a F_1 blue eyed female, all the **females** will inherit X^A and will be **blue** eyed. White eyed males must inherit X^a from the mother (1/2 probability) and at least 1 dominant mutant allele (depicted by "b"), which occurs at 3/4 probability. Multiplying these together gives a 3/8 chance of having white eyed males. This gives the **males** a ratio of **5:3** for blue to white eyed male insects from this cross.

If instead we cross a F_1 white eyed male to a F_1 blue eyed female, the phenotypic ratio is the same for the **male** progeny (**5:3** blue to white eyed males) -- but this time all females inherit X^a from their fathers and inherit X^a with 1/2 probability from their mothers. As in the F_2 males, then, if they inherit a "b" allele (3/4 probability) they will be white eyed. This gives the **females** a **5:3** blue to white ratio.

3a.
II-3 must be aa, as he has the disease. His parents, I-1 and I-2, must therefore be Aa, as they both passed him an a, but do not have the disease. III-3 must be Aa, because he must inhereit an a from II-3, but does not have the disease. The probability that heterozygous III-3 passes the a allele down to ? is 1/2. II-2 must be a carrier for ? to get the disease, and the probability of this is 2/3, as we know that she does not have the disease. The probability that III-2 will get the a from a heterozygous II-2 is 1/2, as is the probability of III-2 passing the allele to ?. Therefore the probability of ? getting the disease is 1/2*2/3*1/2*1/2=2/24=**1/12**.

3b.
If II-3 expresses the disease, he must be X^aY, with the X^a from I-1, as I-2 had to pass the Y. I-1 is X^A/X^a, as she does not have the disease. The chance that a heterozygous individual passes either allele to their offspring is 1/2, so p(I-1 to II-2)=p(II-2 to III-2)=p(III-2 to ?)=1/2, all of which must happen for ? to have the disease. If ? is male, the genotype of III-3 does not matter, as III-3 passes a Y. The probability that a **son** will have the disease is (1/2)^3=**1/8**. If ? is female, we must consider the genotype of III-3. He must be X^AY, as he does not express the disease and always passes X^A to his daughters. So, the probability that a **daughter** will have the disease is **0**.

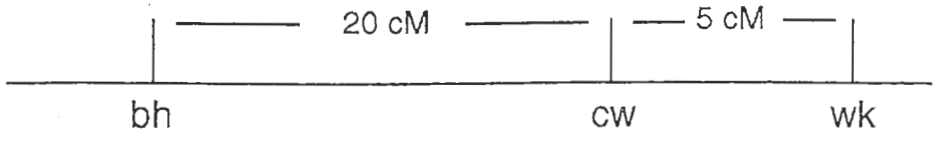
3c.
Once we know that a child of the fourth generation has the disease, we know that III-2 and III-3 must be carriers, heterozygous for the gene. Therefore the probability that a second child will have the disease is **1/4**.



7.03 Problem Set 2

Due before 5 PM on Thursday, September 30
Hand in answers in recitation section or in the box outside the class

1. The hypothetical *Drosophila* traits big head, curly wings, and weak knees are determined by the recessive alleles **bh**, **cw**, and **wk** respectively. The three genes are autosomal and are linked to one another. The gene order and two of the distances between genes are shown in the following genetic map.



A fly from a big-headed weak-kneed strain is crossed to a fly from a curly-winged strain and all of the F_1 progeny look normal. Females from the F_1 are collected and crossed to males from a big-headed, weak-kneed, curly-winged strain.

(a) List all eight of the possible phenotypes that could result from this cross. Indicate the two phenotypes that should be the most abundant and the two phenotypes that should be the least abundant.

(b) If 1,000 progeny flies were examined from this cross, how many flies of each of the eight phenotypic classes would you expect?

2. You have isolated three new His^- yeast mutants in mating type α. You cross each of the mutants to a wild-type (His^+) strain of mating type **a**, and then evaluate 50 of the resulting tetrads for the His^+ or His^- character. In principle, there are five different possible types of tetrads that could be obtained.

Type 1	Type 2	Type 3	Type 4	Type 5
4 His^+	3 His^+ : 1 His^-	2 His^+ : 2 His^-	1 His^+ : 3 His^-	4 His^-

(a) In the cross of mutant1 to wild-type, all of the tetrads are of Type 3. What does this tell you about the His^- mutation(s) in mutant 1?

(b) In the cross of mutant 2 to wild-type, 40 tetrads are Type 3; 8 tetrads are Type 4; and 2 tetrads are Type 5. Propose a genetic mechanism that would explain the behavior of mutant 2. Include in your answer a classification of each relevant tetrad type as PD, NPD, or T. Also give as much information as possible about the His^- mutation(s) in mutant 2.

(c) In the cross of mutant 3 to wild-type, 45 tetrads are Type 3 and 5 tetrads are Type 2. Propose a genetic mechanism that would explain the behavior of mutant 3. Include in your answer a classification of each relevant tetrad type as PD, NPD, or T. Also give as much information as possible about the His^- mutation(s) in mutant 3.

3. **(a)** You have two useful strains of phage λ with mutations in the **cI** gene. The **cI-1** mutation maps very close to the beginning of the **cI** gene coding sequence while the **cI-2** mutation maps close to the end of the coding sequence. Both mutations cause the phage to form clear plaques rather than the normal turbid plaques. When phage with **cI-1** are crossed to phage with **cI-2**, four plaques out of 1000 are turbid. What is the distance between **cI-1** and **cI-2** in map units?

(b) Given your answer for part **(a)** and that the repressor protein is 240 amino acids long estimate the recombination rate for phage λ in crossovers per kb (10^3 base pairs).

(c) You isolate a new mutation in the **cI** gene and in crosses between the new mutant and **cI-1** turbid plaques produced at twice the frequency as in crosses between the new mutant and **cI-2**. You discover that your new mutation introduces a stop codon into the coding sequence of the **cI** gene. Given that the average molecular weight of an amino acid is 110 Daltons, what is the expected molecular weight of the product of your mutant **cI** gene?

(d) A number of different mutagens cause what are known as transition mutations in which a T•A base pair is converted to C•G or a C•G base pair is converted to T•A. By examining the table for the genetic code, determine the sense codons (and the amino acids which they code for) that can be converted into a stop codon by a transition mutation.

AU

Problem Set 2 Answers

1. Your cross looks like this:

F1 $\frac{bh}{+}\ \frac{+}{cw}\ \frac{wk}{+}$ females $\times$ $\frac{bh}{bh}\ \frac{cw}{cw}\ \frac{wk}{wk}$ males

↓

	a. genotype	phenotype	b. expected #
1	$\frac{bh}{bh}\ \frac{+}{cw}\ \frac{wk}{wk}$	**big head, weak knees**	**380**
2	$\frac{+}{bh}\ \frac{cw}{cw}\ \frac{+}{wk}$	**curly wings**	**380**
3	$\frac{bh}{bh}\ \frac{+}{cw}\ \frac{+}{wk}$	**big head**	**20**
4	$\frac{+}{bh}\ \frac{cw}{cw}\ \frac{wk}{wk}$	**curly wings, weak knees**	**20**
5	$\frac{bh}{bh}\ \frac{cw}{cw}\ \frac{+}{wk}$	**big head, curly wings**	**95**
6	$\frac{+}{bh}\ \frac{+}{cw}\ \frac{wk}{wk}$	**weak knees**	**95**
7	$\frac{bh}{bh}\ \frac{cw}{cw}\ \frac{wk}{wk}$	**big head, curly wings, weak knees**	**5**
8	$\frac{+}{bh}\ \frac{+}{cw}\ \frac{+}{wk}$	**wildtype**	**5**

Classes 1 and 2 will be the most abundant because they are the parentals.
Classes 7 and 8 will be the least common because they represent the doubles.
b) The probability of a double crossover equals the probability of two single cross overs. p (DCO)= p(SCO between bh-cw) * p(SCO between cw-wk)=

20/100 * 5/100 =1/100

Thus, 0.5% will be class 7 and 0.5% will be class 8, or **5** of each class. Note: this is the expected value if two crossovers are equally likely to occur with those probabilities.

The distance between bh and cw is 20cM. Thus 10% of the total number will be class 5 and 10% will be class 6, or 100 in each case. However, you have to subtract out the number of DCO, leaving **95** for each class.

The distance between cw and wk is 5cM. 2.5% of the total number will be class 3 and 2.5% will be class 4, or 25. Again, subtracting out the double crossovers leaves **20**.

Now, the number of parentals can be determined by subtraction, **380** for each class.

2. a) All of the tetrads are Parental Ditypes. There is only one mutation in one of the genes in the histidine biosynthetic pathway.

b)

40	2His+ : 2His-	Parental Ditypes
8	1His+ : 3His-	Tetratypes
2	4His-	Nonparental Ditypes

There are two genes which are linked since PD >> NPD. A single mutation in one of these genes will give a mutant phenotype.

The distance between the genes is found using the formula

$$cm = \frac{T + 6NPD}{2\sum} \times 100 = \frac{8+12}{100} \times 100 = 20 \text{ cM}$$

c)

45	2His+ : 2 His-	Parental Ditypes
5	3His+ : 1 His-	Tetratypes

In this case, the Tetratypes give more His+ phenotypes. There are again two mutated genes involved, however, this time the two mutations must be together to get a mutant His- phenotype. Otherwise, with only a single mutation, they are His+. There are no NPDs, so the distance between the two genes is 5cM.

3. (**a**) The 4 turbid plaques we see must be of genotype **cl-1**+ / **cl-2**+ indicating that in 4 cases crossovers have taken place to restore the wildtype genes. There will also be 4 plaques that are **cl-1**-/**cl-2**- resulting from the same crossovers. We can not see these because they have the same phenotype (forming clear plaques) as the single mutants. However, all 8 are recombinants. Using the equation for distance:

Distance between mutations = (number of recombinants / total) x 100

=(8/1000) x 100

= .8 m.u.

(**b**) Since the repressor protein is 240 AA and 3 basepairs encode each amino acid, the coding sequence of the gene should be 720bp (or .720 kb). We know from part (a) that there were 4 crossovers in 1000 phage, so the calculation would be set up as:

$$\frac{\text{(4 crossovers)}}{\text{(1000 phage) (.720 kb)}} = .55 \text{ crossovers/kb for a single phage}$$

(**c**) Since the new mutant crossed with **cl-1** produces twice as many crossovers as the new mutant crossed with **cl-2**, we can assume that the mutation is twice as far from **cl-1** as from **cl-2**. Knowing that the gene is 720 bp, the distance from **cl-1** to the mutation must be (720bp)(2/3) = 480bp. Since this mutation is a stop codon, the protein will be truncated at this point. We can therefore calculate the size of the protein.

(480bp) x (1 amino acid/ 3 bp) x (110 Daltons/ amino acid) = 17.6 kDa

(**d**) To answer this problem, you need to work backwards from the three possible stop codons:

UAG			Stop codon (sense mRNA)
ATC			DNA sequence that would translate to UAG (antisense strand)
GTC	ACC	ATT	DNA sequences that could result in ATC if a transition took place
CAG	UGG	UAA	sense mRNA sequences produced from above three DNA sequences
Gln	Trp	Stop	Amino Acids that above mRNAs encode

UAA			Stop codon (sense mRNA)
ATT			DNA sequence that would translate to UAA (antisense strand)
GTT	ACT	ATC	DNA sequences that could result in ATT if a transition took place
CAA	UGA	UAG	sense mRNA sequences produced from above three DNA sequences
Gln	Stop	Stop	Amino Acids that above mRNAs encode

UGA			Stop codon (sense mRNA)
ACT			DNA sequence that would translate to UGA (antisense strand)
GCT	ATT	ACC	DNA sequences that could result in ACT if a transition took place
CGA	UAA	UGG	sense mRNA sequences produced from above three DNA sequences
Arg	Stop	Trp	Amino Acids that above mRNAs encode

There are four possible mRNA sense codons where transitions could cause a premature stop codon and a protein truncation:

mRNA codon	Corresponding Amino Acid
CAG	Gln
CAA	Gln
UGG	Trp
CGA	Arg

7.03 Problem Set 3

due before 5 PM on Thursday, October 14

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

1. The mutagen hydroxyalmine preferentially causes transition mutations of the type C•G to T•A, because hydroxylamine converts cytosine residues into uracil residues by deamination. Determine the tRNA genes that could be converted (by a single base change) into an ochre (UAA) suppressor by mutagenesis with hydroxylamine.

2. Wild type *E. coli* has flagella that allow them to swim towards nutrient sources. Nonmotile mutants are easily detected because wild type *E. coli* makes colonies with diffuse edges when plated on soft agar, whereas nonmotile mutants make compact colonies. You have isolated two nonmotile mutants that you call **mot1**$^-$ and **mot2**$^-$.

(a) A this stage why is it not possible to carry out crosses between **mot1**$^-$ and **mot2**$^-$ by using **P1** transduction to determine whether the **mot**$^-$ mutants are linked?

In order to remedy the problem outlined in part **(a)** you set out to isolate a **Tn5** insertion linked to one of the **mot**$^-$ mutants. To do this you start with a collection of 1000 different random **Tn5** insertions in the otherwise wild type *E. coli* strain (these insertion strains are all kanamycin resistant (Kan^r) and motile (mot$^+$)). You grow **P1** phage on the entire mixed collection of **Tn5** insertion strains and then infect the **mot1**$^-$ mutant and select for Kan^r transductants. Most of the Kan^r transductants remain nonmotile, but one out of 1000 is motile.

(b) Explain the origin of this motile transductant.

Next you grow **P1** phage on the motile transductant isolated above and then infect the **mot1**$^-$ mutant with the resulting phage. You select for Kan^r transductants and then test these transductants for motility. You find that 60 out of 100 Kan^r transductants are motile.

(c) Give the distance between the **Tn5** insertion and **mot1**, expressed as a cotransduction frequency.

(d) By sequencing the junction between the **Tn5** insertion isolated above and the neighboring chromosomal sequences you find the **Tn5** insertion lies about 50 kbp from the **PhoS** gene. A headful of DNA for phage **P1** is 100 kbp. In transduction experiments you find that the **Tn5** insertion and the **PhoS** gene show 50% cotransduction, but **PhoS** and **mot1** are not linked by transduction. Explain.

Using the same procedure outlined in part **(c)** you find that the **Tn5** insertion shows about the same linkage to **mot2⁻** as to **mot1⁻**. You construct a strain that has both the **Tn5** insertion and **not1⁻** and another strain that has both the **Tn5** insertion and **mot2⁻**. Using these strains you perform two reciprocal crosses. In the first cross, **P1** is grown on the **Tn5 mot1⁻** strain and the resulting phage are used to infect a **mot2⁻** strain. In this transduction experiment, 80 out of 1000 Kan^r transductants are motile. In the reciprocal cross, **P1** is grown on the **Tn5 mot2⁻** strain and the resulting phage are used to infect a **mot1⁻** strain. In this experiment, 8 out of 1000 Kan^r transductants are motile.

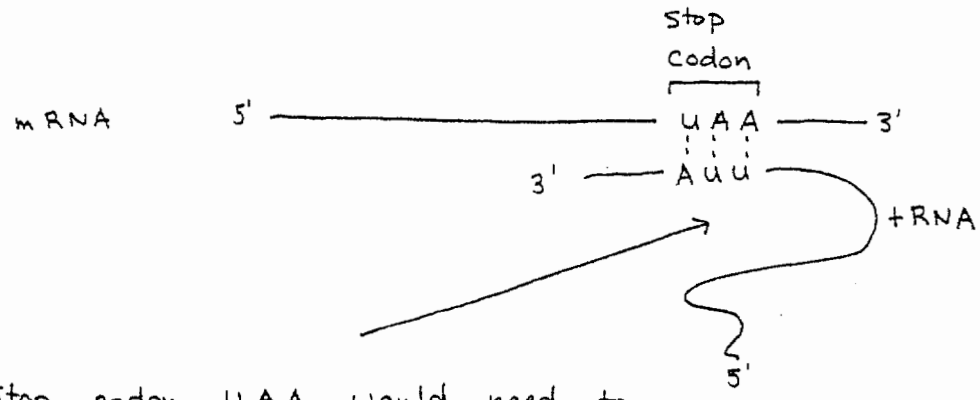
(e) Draw a map showing the relative order of the **Tn5** insertion, **mot1⁻** and **mot2⁻**.

3. Transposons are not only useful as portable genetic markers, they can also serve as portable regions of homology for recombination. In this problem we will see how a **Tn5** insertion can be used to construct an **F'** plasmid with desired characteristics. These methods rely on the use of a special **F** factor that carries an insertion of **Tn5** (this factor is designated **F::Tn5**).

(a) Starting with the **Tn5** insertion isolated in problem 2, you next introduce **F::Tn5** into this strain. An **Hfr** can then be isolated by selection for a strain that can transfer **PhoS** in a mating experiment. Given that the **Hfr** arose by homologous recombination between the **Tn5** on **F** and the **Tn5** on the chromosome draw a diagram showing the recombination event between **F** and the chromosome. For your answer include the location of **PhoS** and the **mot1** as well as the relative orientations of the two **Tn5** elements and the origin of transfer on **F** (more than one arrangement of orientations is possible).

(b) Having isolated an **Hfr** that can transfer **PhoS**, you are now in a position to isolate an **F'** that carries the **Mot** gene(s) by selecting for early transfer of Mot^+ into a **mot1⁻** mutant. Explain the rationale for this method for isolation of an **F'** and draw a diagram showing a possible recombination event that could produce the desired **F'**.

Problem Set #3 Solutions:



Stop codon UAA would need to be recognized by the anticodon UUA of a charged +RNA to be an ochre suppressor

therefore, what DNA sequences coding for +RNAs can be mutated to generate the +RNA sequence 5' —— UUA-3' by a C·G → T·A transition?

sired +RNA: 5' —— UUA — 3'

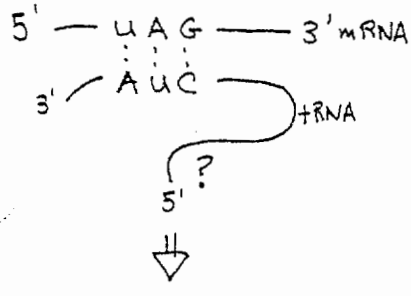
rresponding sense DNA strand: 5' —— TTA — 3'

C·G → T·A transition:

① 5' — CTA — 3' ② 5' — TCA — 3' ③ 5' — TTG — 3'

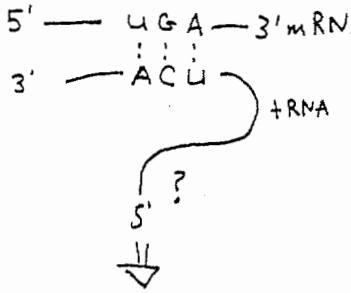
code for +RNA anticodons

RNA codons:



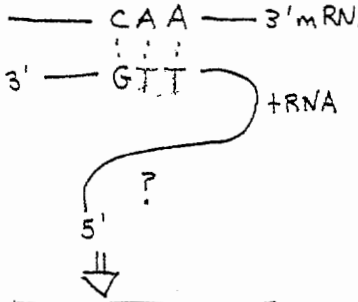
UAG = stop

② 5' — UGA — 3' mRNA



UGA = stop

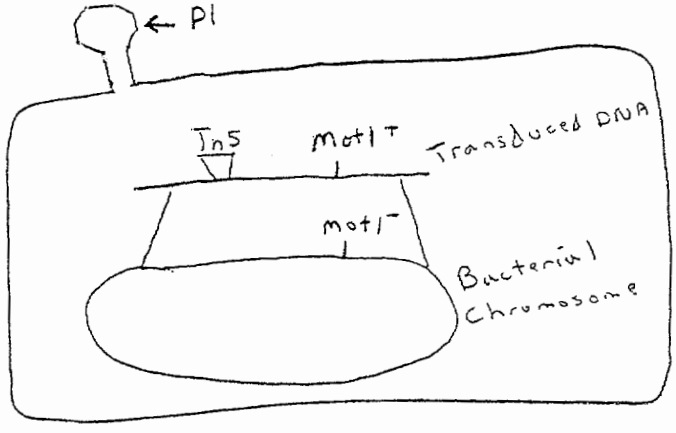
③ 5' — CAA — 3' mRN



OCHRE SUP. CAA = Gln

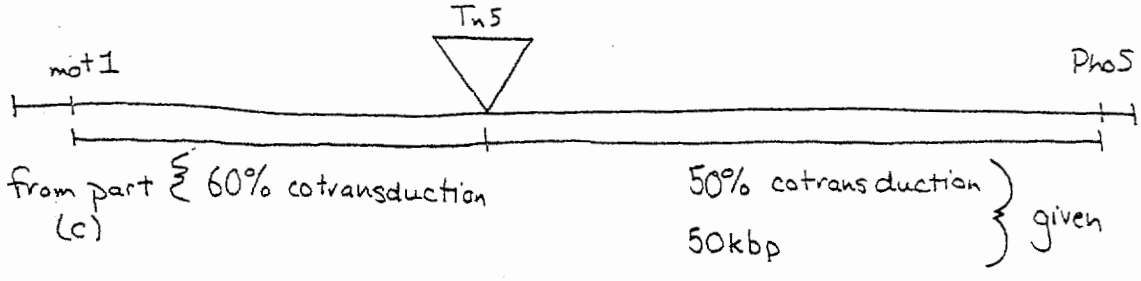
2. (a) It is not possible to carry out crosses between the **mot** mutants by **P1** transduction because a selectable marker near the mot locus is required. Without a selection it would be impossible to identify transductants. The antibiotic resistance gene encoded by **Tn5** fulfills this requirement.

2. (b) Presumably, some members of the **Tn5** insertion collection have transposons that are closely linked to the wild type **mot1** gene. When **P1** phage are grown on the insertion collection, some of the phage will be produced by these members of the collection. Occasionally one of the phage will accidentally package a portion of the bacterial chromosome that contains both the **Tn5** insertion and the wild type **mot1** gene. The motile transductant is the result of a phage like this infecting the nonmotile *mot1-* mutant followed by the pictured recombination event.

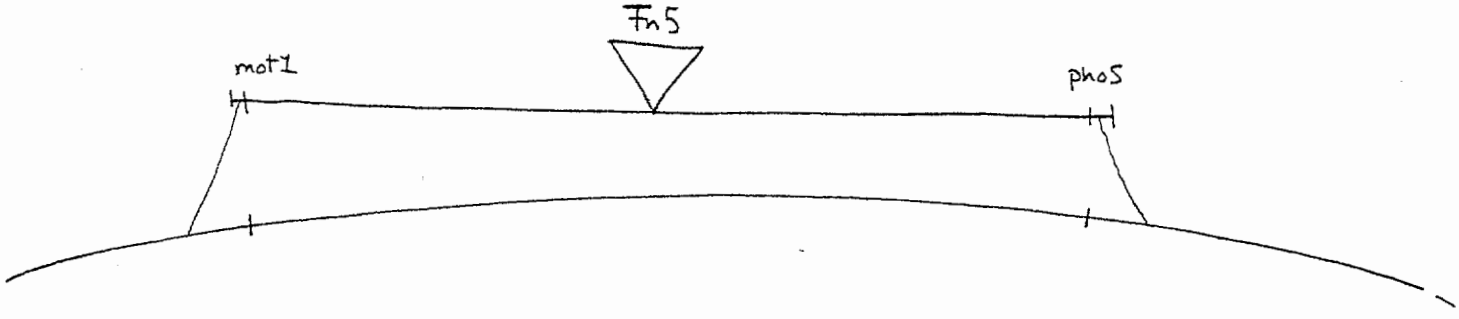


2. (c) In this case the cotransdution frequency between **Tn5** and **mot1** equals the number of Kan^r, motile transductants divided by the total number of Kan^r transductants multiplied by 100. There are 60 Kan^r, motile transductants and 100 total. Therefore the cotransduction frequency is 60%.

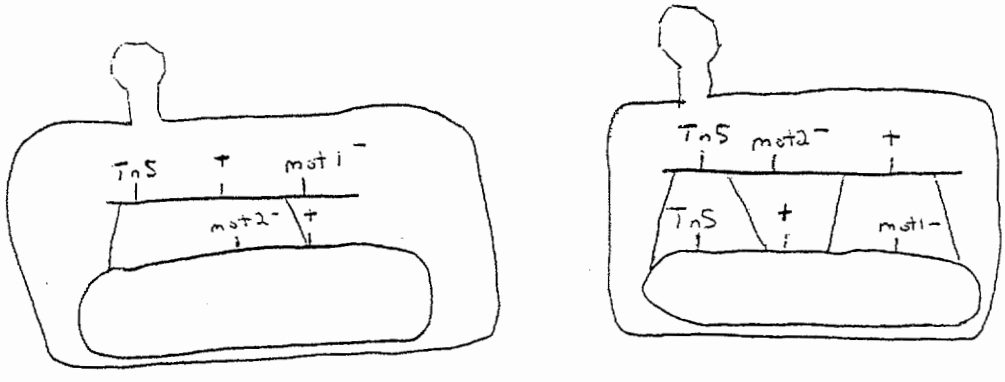
2. (d) First of all it means that **mot1** and **PhoS** are on opposite sides of the **Tn5** insertion. If they were not they would certainly be linked by transduction because they would easily be packaged together.



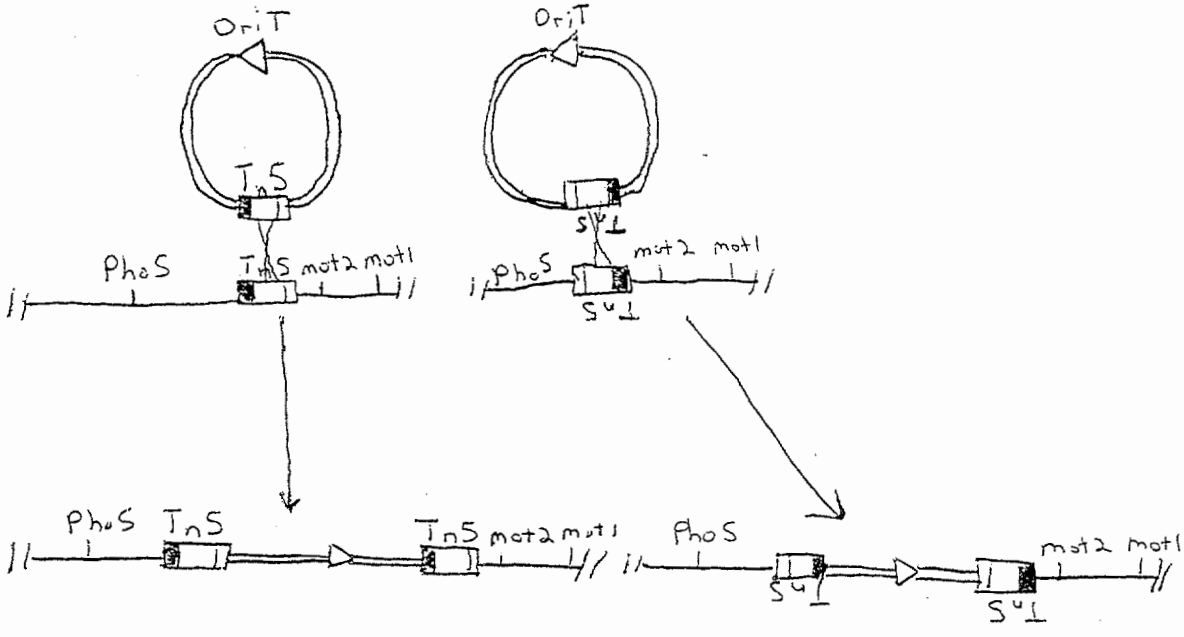
In this problem you are not sure of the exact distance between **mot1** and **PhoS** but it is approaching the 100 kbp size limit for a P1 phage head. The farther two mutations are apart the less likely they will be packaged together and cotransduced. The cotransduction frequency actually decreases by a power of 3 of the physical distance. Therefore it is unlikely that **mot1** and **PhoS** will be packaged together, although not impossible. Additionally to observe linkage by transduction in this case recombination would have to occur in the regions flanking **mot1** and **PhoS** which are both very small in this case. This also decreases the probability of seeing the linked transduction even more.



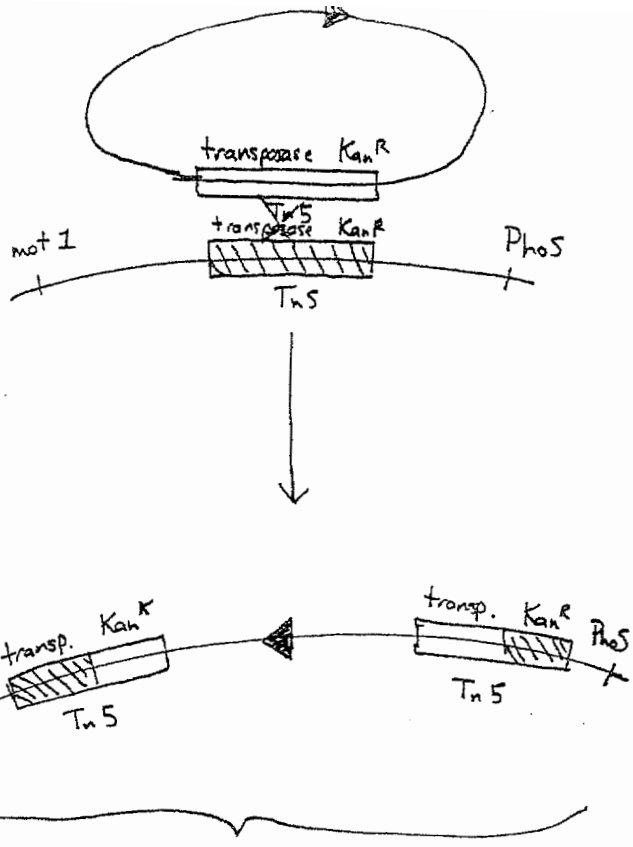
2. **(e)** The relative order between the three markers is **Tn5 - mot2 - mot1**. This order is derived from the deduction that the cross rarely giving motile transductants must be the result of four cross-overs between the phage and bacterial chromosome. The recombination event below on the left illustrates the transduction experiment yielding 80 motile transductants. The recombination event below on the right, illustrates the experiment yielding 8 motile transductants. This is the only gene order that could yield these results.



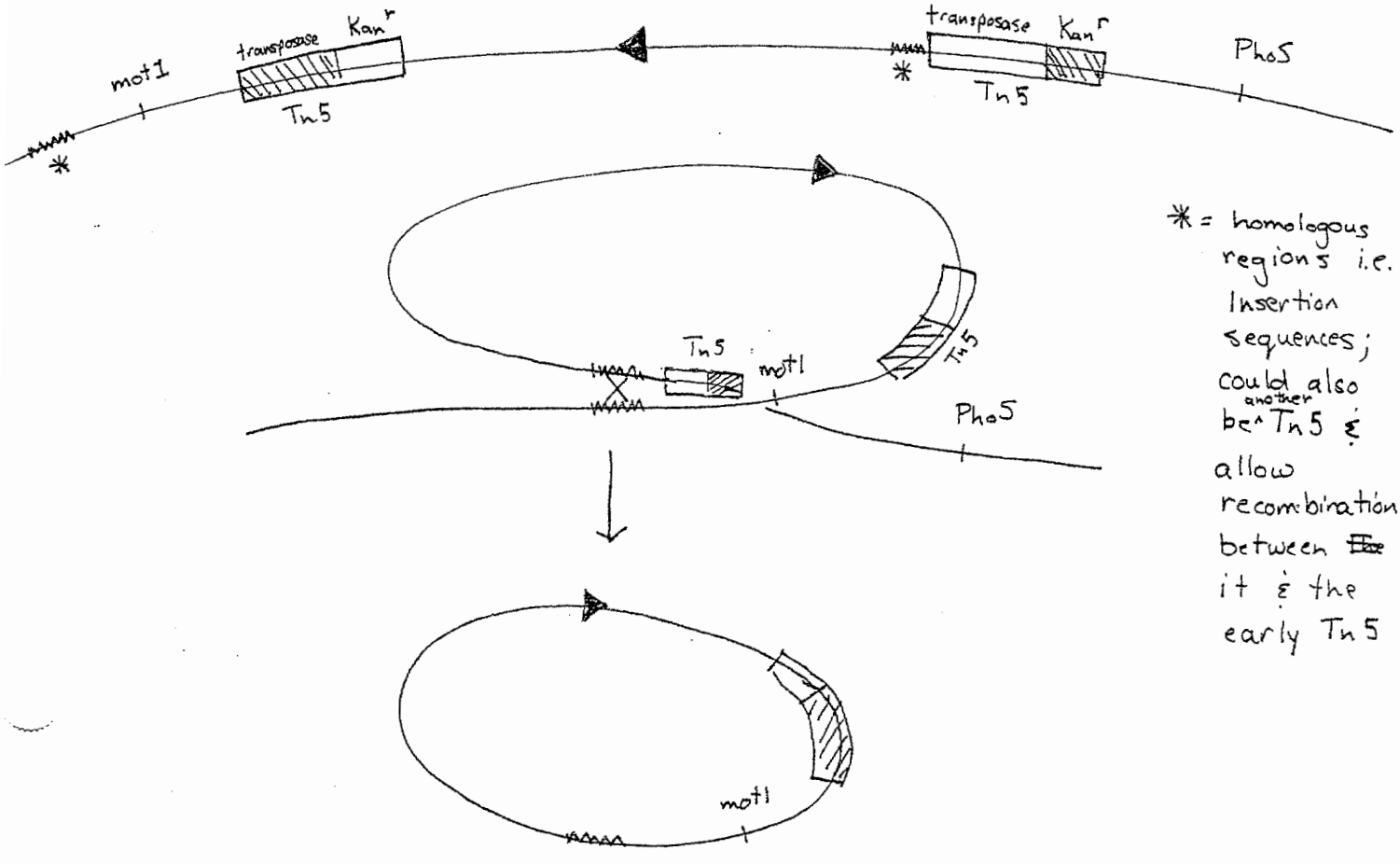
3. (a) The two potential Hfrs are determined by the orientation of **Tn5** relative to Ori T of the F plasmid. Note that because the Hfr must transfer the **PhoS** gene, the tail end of the Ori T must face that direction.



3. (b) (cont.) By looking for early transfer of **Mot**$^+$ into a **mot1**$^-$ mutant (using the **Hfr** from above) you ensure that you will only get out a **F'** which carries **Mot**$^+$, because in this **Hfr**, **Mot**$^+$ will never be transferred early. Only if **F'** is formed in which **Mot**$^+$ is incorporated will **Mot**$^+$ be transferred early.



use this Hfr which transfers PhoS frequently & early below



* = homologous regions i.e. Insertion sequences; could also be another Tn5 & allow recombination between ~~the~~ it & the early Tn5

7.03 Problem Set 4

due before 5 PM on Monday, October 25

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

1. You have isolated a **Tn5** insertion that is linked to the **Lac** operon. Since the **Tn5** insertion does not perturb **Lac** gene expression, the site of insertion can be assumed not to be within the **Lac** operon itself, but otherwise you don't know where the insertion site is.

(a) You grow phage **P1** on a strain that has the **Tn5** insertion and a **LacI**$^-$ mutation. This phage is used to infect a **LacZ**$^-$ mutant and transductants are selected by plating on kanamycin plates. Among the Kan^r transductants, 50% are regulated (by IPTG), 10% are constitutive, and 40% are uninducible. On the basis of this information draw a genetic map of the position of the **Tn5** insertion relative to **LacI** and **LacZ**, showing relative order and the relevant distances in cotransduction frequencies.

(b) You will now use the same **Tn5** insertion to determine the order of two different **LacI**$^-$ mutations. Phage **P1** is grown on a strain with the **Tn5** insertion and the **LacI-1**$^-$ mutation. This phage is then used to infect a **LacI-2**$^-$ strain and Kan^r transductants are selected. Among 500 Kan^r transductants, 5 are regulated by IPTG and the rest are constitutive. In a second transduction experiment, phage **P1** is grown on a strain with the **Tn5** insertion and the **LacI-2**$^-$ mutation. This phage is then used to infect a **LacI-1**$^-$ strain and Kan^r transductants are selected. Among 500 Kan^r transductants, all are constitutive. Draw a map showing the relative order of the **Tn5** insertion and the **LacI-1**$^-$ and **LacI-2**$^-$ mutations.

2. For the following merodiploid strains, determine the level ß-galactosidase expression in either the presence or absence of IPTG. Assume that when no repressor is bound to DNA, 100 units of ß-galactosidase activity is produced from each functional copy of the **LacZ** gene and when repressor is fully bound to DNA only 1 unit of enzyme is produced for each functional copy of **LacZ**. Finally, assume that the presence of **Lac I**$^{-d}$ protein will fully prevent other forms of the repressor in the same cell from binding to DNA.

lac I^{-d} Z^+ / F' lac I^+ Z^-

lac O^+ Z^- / F' lac O^c Z^+

lac I^+ Z^- Y^- / F' lac I^s Z^+ Y^+

lac I^+ O^c Z^+ / F' lac I^{-d} O^+ Z^+

lac I^+ O^c Z^+ / F' lac I^s O^+ Z^+

lac I^{-d} O^+ Z^+ / F' lac I^s O^+ Z^+

lac I^{-d} O^c Z^+ / F' lac I^s O^+ Z^+

lac I^{-d} O^c Z^- / F' lac I^s O^+ Z^+

3. Genes whose function is to metabolize compounds (for example the **Lac** and **Mal** genes) are often regulated in the sense that they are induced by the presence of substrate compounds. On the other hand, genes involved in biosynthesis are often repressed by the presence of the compound that is produced by the biosynthetic pathway. Consider a biosynthetic gene **X** that is regulated by the product (compound **Y**) in the sense that **X** is not transcribed when **Y** is present but is transcribed when **Y** is absent. You have identified two regulatory genes **A** and **B** that are both unlinked to gene **X**. Recessive mutations in gene **A** cause constitutive synthesis of gene **X** (even in the presence of **Y**), whereas recessive mutations in gene **B** cause gene **X** to be uninducible (even in the absence of **Y**).

(a) Draw two different regulatory models showing the interactions among genes **A**, **B**, and **X**. For your model, use the symbol $\rightarrow$ to designate an activating interaction and the symbol $\dashv$ to designate an inhibitory interaction. For your models show explicitly how compound **Y** would interact with the appropriate protein to give the observed regulatory behavior. Also, for each model describe the expected behavior of an $\mathbf{A^-}$ $\mathbf{B^-}$ double mutant.

(b) Further analysis of this regulatory system reveals that the products of both genes **A** and **B** bind to the DNA of the promoter region of gene **X**, but to no other sites on the genome. Diagram a new regulatory model to account for everything that you know about genes **A** and **B**. Again, show explicitly how compound **Y** would interact with the appropriate protein to give the observed regulatory behavior and describe the expected behavior of an $\mathbf{A^-}$ $\mathbf{B^-}$ double mutant.

Problem Set #4:

1 (a) P1 phage made on: Kan^R, I^-, Z^+

starting strain: Kan^S, I^+, Z^-

transductants: inducible = Kan^R, I^+, Z^+

constitutive = Kan^R, I^-, Z^+

uninducible = Kan^R, I^-, Z^-

or

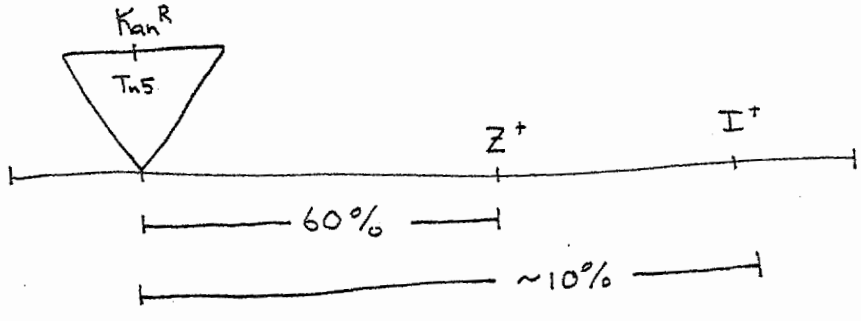
Kan^R, I^+, Z^-

therefore: · cotransduction frequency of Kan^R & Z^+ is

% Kan^R, I^+, Z^+ + % Kan^R, I^-, Z^+

= 50% + 10% = 60%

· the distance from Kan^R to I should be fairly close to 10% because the 4 crossover case occurs so rarely



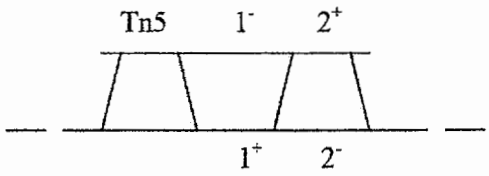
(b) The order is either: **Tn5 LacI-1 LacI-2** or **Tn5 LacI-2 LacI-1** because Tn5 cannot be inserted between LacI-1 and LacI-2 (if it were, you would have disrupted the LacI gene and you would not get back a wild-type inducible LacI-1^+ LacI-2^+ phenotype.) Now you want to compare the two different transductions for both orders and see how many crossovers it takes to get the rarest class (which is in both transductions: **inducible by IPTG**, or **LacI-1^+ LacI-2^+**):

Order #1:
Tn5 LacI-1 LacI-2

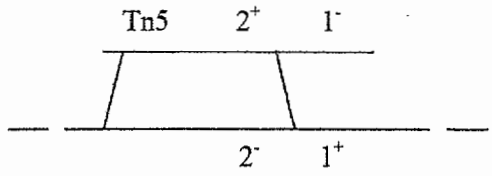
Order #2:
Tn5 LacI-2 LacI-1

Transduction #1:
Tn5 LacI-1^- infecting LacI-2^-

→ **5** are **LacI-1^+ LacI-2^+**



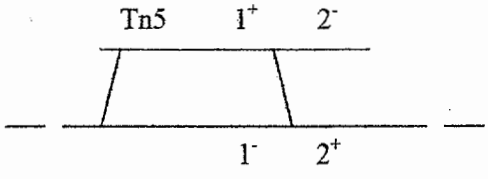
4 crossovers



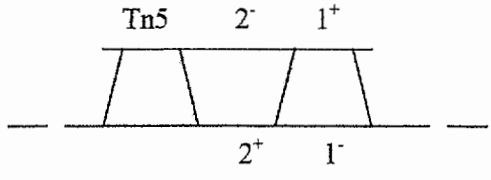
2 crossovers

Transduction #2:
Tn5 LacI-2^- infecting LacI-1^-

→ **0** are **LacI-1^+ LacI-2^+**



2 crossovers

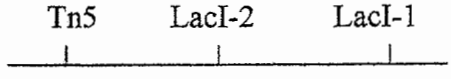


4 crossovers

Order #1 doesn't make sense because 4 crossovers should be rarer than 2 crossovers, yet we got **5** LacI-1^+ LacI-2^+ transductants in transduction #1 and **0** in transduction #2.

Order #2 makes sense because 4 crossovers should be rarer than 2 crossovers. Consequently, we saw less LacI-1^+ LacI-2^+ transductants in transduction #2 than in transduction #1.

∴The map order is:

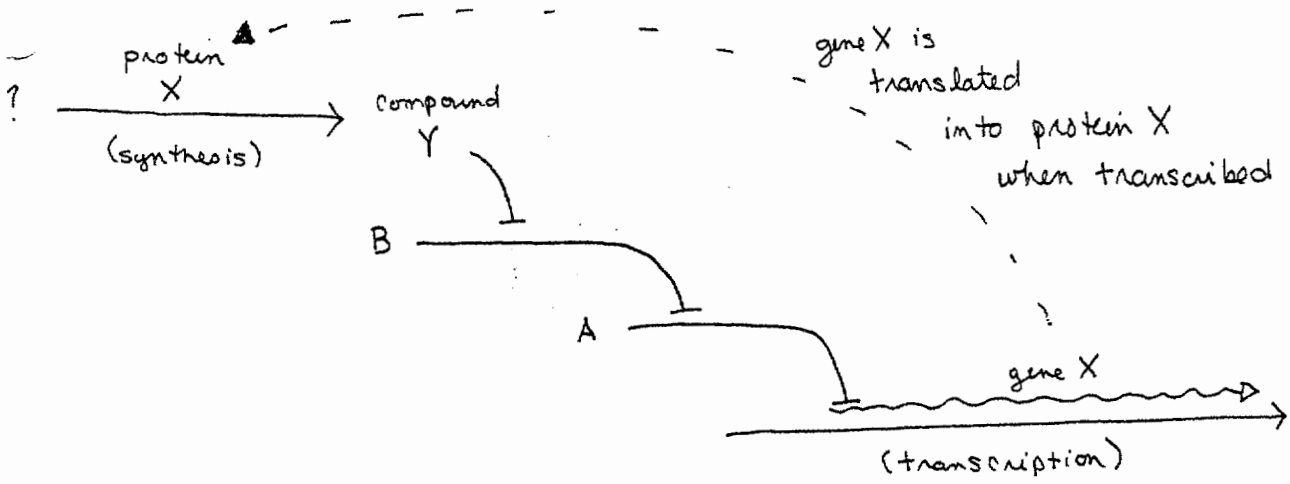


2.

Merodiploid Strain	- IPTG	+ IPTG	Explanation
lac I^{-d} Z^+ / F' lac I^+ Z^-	100	100	Lac I^{-d} protein prevents other forms of repressor from binding to DNA, so regulation is constitutive. There's only one wild-type copy of lacZ in the merodiploid so 100 U of B-gal activity is produced.
lac O^+ Z^- / F' lac O^c Z^+	100	100	O^c prevents all types of repressor from binding, so regulation is constitutive. There's only one wild-type copy of lacZ in the merodiploid so 100 U of B-gal activity is produced.
lac I^+ Z^- Y^- / F' lac I^s Z^+ Y^+	1	1	I^s binds to the operator and is not able to be induced by IPTG because of a mutation in the inducer binding pocket. I^s stays on bound to the operator permanently, preventing B-gal transcription. There's only one wild-type copy of lacZ in the merodiploid so 1 U of B-gal activity is produced.
lac I^+ O^c Z^+ / F' lac I^{-d} O^+ Z^+	200	200	Lac I^{-d} protein prevents other forms of repressor from binding to DNA, so regulation is constitutive. There are two wild-type copies of lacZ in the merodiploid so 200 U of B-gal activity is produced.
lac I^+ O^c Z^+ / F' lac I^s O^+ Z^+	101	101	O^c prevents all types of repressor from binding, so regulation is constitutive on the F^- strand. I^s binds to the operator (O^+) permanently in the F' strand. There are two wild-type copies of lacZ, one that is under the control of O^c (100 U B-gal) and one that is repressed by I^s (1 U of B-gal), so total B-gal produced is 101 U.
lac I^{-d} O^+ Z^+ / F' lac I^s O^+ Z^+	200	200	Lac I^{-d} protein prevents other forms of repressor from binding to DNA, so regulation is constitutive. There are two wild-type copies of lacZ in the merodiploid so 200 U of B-gal activity is produced.
lac I^{-d} O^c Z^+ / F' lac I^s O^+ Z^+	200	200	Lac I^{-d} protein prevents other forms of repressor from binding to DNA, so regulation is constitutive. There are two wild-type copies of lacZ in the merodiploid so 200 U of B-gal activity is produced.
lac I^{-d} O^c Z^- / F' lac I^s O^+ Z^+	100	100	Lac I^{-d} protein prevents other forms of repressor from binding to DNA, so regulation is constitutive. There's only one wild-type copy of lacZ in the merodiploid so 100 U of B-gal activity is produced.

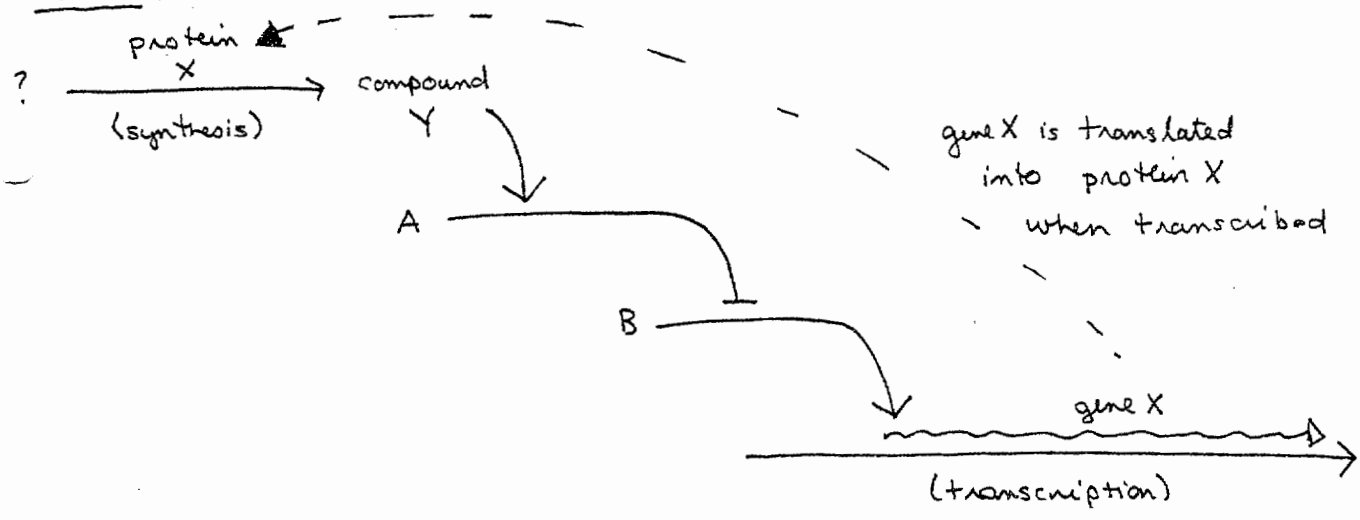
(a)

Model 1:



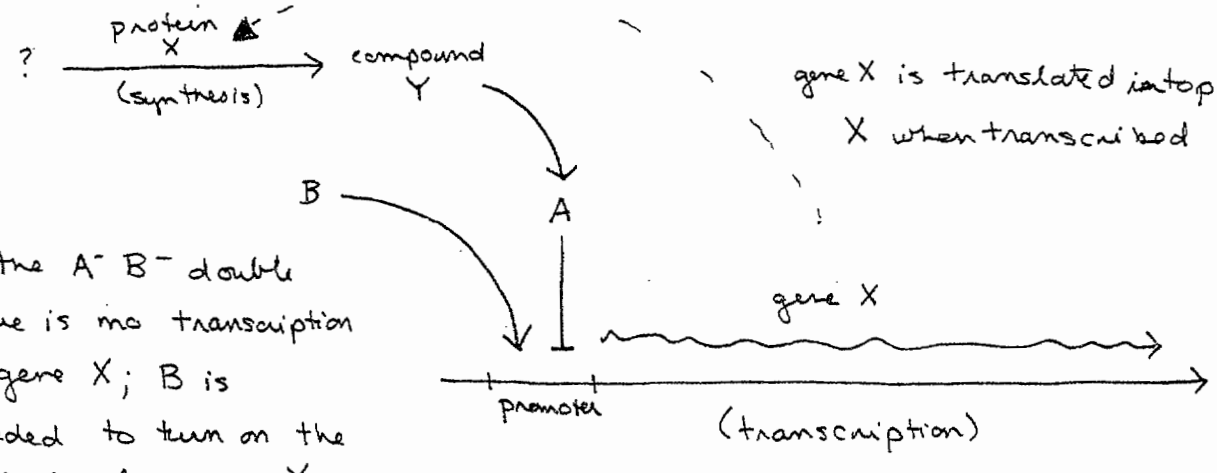
- in the $A^- B^-$ double there would be constitutive synthesis of gene X (constitutive)

Model 2:



- in the $A^- B^-$ double there would be no synthesis of gene X (uninducible)

(b)



- in the $A^- B^-$ double there is no transcription of gene X; B is needed to turn on the promoter for gene X

7.03 Problem Set 5

due before 5 PM on **Tuesday, November 9**

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

1. Consider a eukaryotic gene regulatory pathway where a small molecule X activates the expression of a reporter gene. You have isolated recessive mutations in two different genes, **A** and **B**, both of which give uninducible expression of the reporter. Genes **A** and **B** are not linked to each other and neither gene is linked to the reporter.

(a) Assuming that the regulatory factors encoded by **A** and **B** act in series, there are two possible orders in which these two regulatory factors can act. Draw out these two models showing the relationships between **A** and **B**, and the reporter. Also be sure to indicate where and how the inducer X acts.

(b) In order to distinguish between the two models from part **(a)**, an epistasis test would be useful. Because recessive mutants in **A** and **B** have the same phenotype (uninducible) it is not possible to perform an epistasis test using these alleles. Fortunately, you are able to isolate an allele of gene **A** that gives constitutive expression of the reporter. This allele, called $\mathbf{A^S}$, is dominant. Describe in molecular terms how the allele $\mathbf{A^S}$ might work for each of your models from part **(a)**.

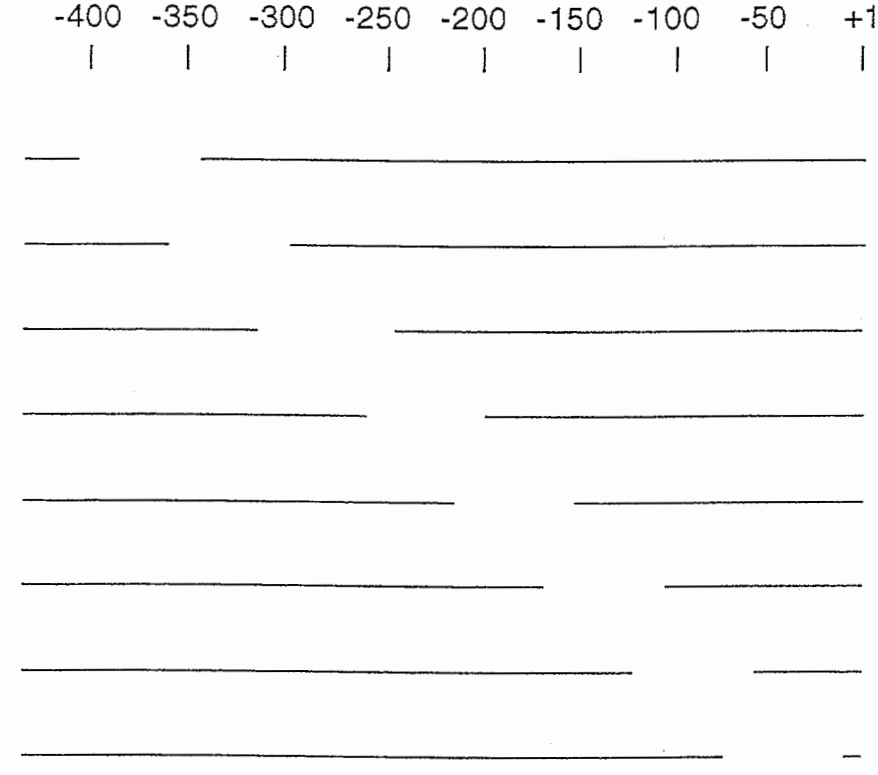
(c) Assume that you are studying this regulatory pathway in yeast and you wish to perform an epistasis test by constructing the $\mathbf{A^S\ B^-}$ double mutant. To do this you cross a MAT**a** $\mathbf{A^S\ B^+}$ strain to a MATα $\mathbf{A^+\ B^-}$ strain and examine the resulting tetrads. For each of the two models from part **(a)**, give the types of tetrads that you would expect and their relative frequencies. The tetrad types should be described by the phenotypes (constitutive, uninducible, or regulated) of the four spores.

(d) Now assume that you are studying this regulatory pathway in *Drosophila*. To perform the epistasis test you cross a $\mathbf{A^S/A^S\ B^+/B^+}$ male to a $\mathbf{A^+/A^+\ B^-/B^-}$ female. For each of the two models from part **(a)**, give the expected phenotype of the F_1 flies from this cross. Now you cross the F_1 flies among themselves to produce F_2 flies. For each of the two models, give the expected ratio of constitutive, uninducible, or regulated phenotypes among the F_2 flies.

2. The yeast **Gap1** gene is regulated by nitrogen source. Under most growth conditions **Gap1** is not expressed, but when yeast are grown on either glutamate or urea as a nitrogen source **Gap1** is expressed at high level. In cells with recessive mutations in **Gln3**, **Gap1** is induced on urea but not on glutamate. On the other hand, in cells with recessive mutations in **Nil1**, **Gap1** is induced on glutamate but not on urea.

(a) Propose a model for **Gap1** regulation that takes into account the action of the **Gln3** and **Nil1** genes as well as the inducers glutamate and urea.

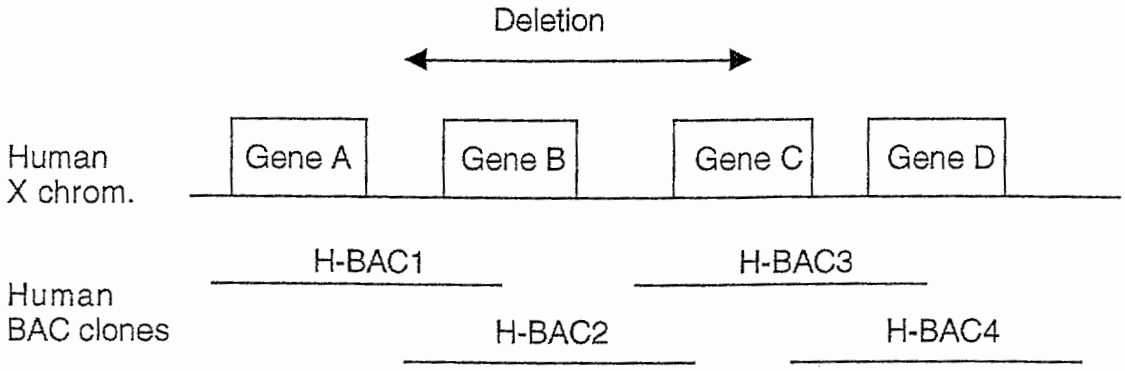
Regulation of **Gap1** has also been studied by examination of the behavior of cis-acting mutations within the **Gap1** gene. Below is a diagram showing a series of deletions that remove different 50 base pair segments from the **Gap1** promoter region and the ability of these deletions to express **Gap1** on either glutamate or urea. Base pairs are numbered relative to the position of the start codon for the Gap1p protein.



(b) Why do you think that deletion 8 does not express **Gap1** on either glutamate or urea?

(c) Given that **Gln3** and **Nil1** are both DNA-binding proteins, identify the regions in the **Gap1** promoter where **Gln3** or **Nil1** are likely to bind.

3. A mammalian geneticist interested in identifying molecular causes of infertility discovers a deletion on the X chromosome in a man who produces no sperm but is otherwise healthy. A series of overlapping BAC clones spanning the relevant region of the X chromosome are available. All four BAC clones have been sequenced.



The geneticist suspects that the absence of either Gene B or Gene C is the cause of the man's infertility, but further studies in humans do not validate or refute this hypothesis. The geneticist wonders whether transgenic mouse studies could 1) provide support for this hypothesis and 2) pinpoint whether Gene B or Gene C is the culprit. The geneticist discovers a nearly identical arrangement of Genes A, B, C, and D on the mouse X chromosome, as well as an equivalent set of overlapping mouse BACs (M-BAC1, 2, 3, 4), all of which have been sequenced.

(a) Design an experiment, utilizing mouse transgenic technology, to test the hypothesis that Gene B is required in sperm-producing cells. (Warning: Your experiment should employ transgenes AND gene targeting if it is to have a good chance of success. Think carefully about the order in which these two technologies should be employed in this experiment.)

(b) What is the predicted outcome of your experiment if Gene B is required in sperm-producing cells?

(c) What is the predicted outcome of your experiment if Gene B is not required for spermatogeneis?

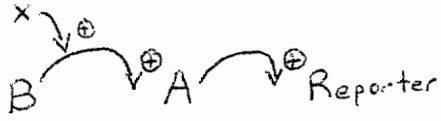
You carry out your proposed experiments on Gene B, as well as parallel experiments to test the hypothesis that Gene C is required in sperm-producing cells. Unfortunately, the results suggest that neither loss of Gene B alone nor loss of Gene C alone impairs spermatogenesis. You then consider the hypothesis that simultaneous loss of Gene B and of Gene C is required to disrupt spermatogenesis. You decide to generate male mice that are mutant in both Gene B and Gene C.

(d) Your first thought is to generate double-mutant mice by crossing the single-mutant strains that you generated in your earlier experiments. Why is this approach likely to fail?

(e) You decide that the best approach to test your new hypothesis will require that you isolate ES (embryonic stem) cells from genetically altered mice. In fact, you realize that you should isolate ES cells from mice carrying a combination of genetic alterations that you engineered in your earlier experiments. What should be the genotype of your newly isolated ES cells, and what additional modification will you make to them to test the hypothesis that simultaneous loss of Gene B and Gene C is required to disrupt spermatogenesis?

7.03 Problem Set 5 Answers
FA '99

1.
(a)
Model 1:



Model 2:

(b)
In Model 1, A^S does not need to be activated by B (it's always turned on) and will give constitutive reporter gene expression even in the absence of the inducer molecule X.

In Model 2, A^S always activates B and does not need the inducer molecule X in order to be turned on. Activation of B will lead to constitutive reporter gene expression.

(c)

Model 1:

PD	T	NPD
2 Constitutive	2 Constitutive	2 Constitutive
2 Uninducible	1 Uninducible	2 Regulated
	1 Regulated	

(A^SB^- double mutants will be Constitutive)

Model 2:

PD	T	NPD
2 Constitutive	1 Constitutive	2 Uninducible
2 Uninducible	2 Uninducible	2 Regulated
	1 Regulated	

(A^SB^- double mutants will be Uninducible)

(d)
Model 1: 12 Constitutive: 3 Regulated: 1 Uninducible
(A^SA^S, B^-B^- and A^SA^+, B^-B^- flies will be Constitutive for gene expression)

Model 2: 9 Constitutive: 4 Uninducible: 3 Regulated
(A^SA^S, B^-B^- and A^SA^+, B^-B^- flies will be Uninducible for gene expression)

2(a) In the presence of urea, the **Nil1** gene product activates expression of **Gap1**. Deletion of the **Nil1** gene will give uninducible expression of **Gap1** with respect to the inducer urea. However this has no effect on **Gln3**'s ability to activate expression of **Gap1** by the inducer glutamate.
In the presence of glutamate, the **Gln3** gene product activates expression of **Gap1**. Deletion of the **Gln3** gene will give uninducible expression of **Gap1** with respect to the inducer glutamate. However this has no effect on **Nil1**'s ability to activate expression of **Gap1** by the inducer urea.

(b) The deletion of this region most likely changes the location of the TATA box and will thus cause uninducible expression of the **Gap1** gene even in the presence of urea or glutamate.

(c) **Nil1** binds somewhere in between the -350 to -250 region of the **Gap1** promoter
Gln3 binds somewhere in between the -150 to -100 region of the **Gap1** promoter.

3(a) First, you want to make a transgenic mouse that carries an extra copy of gene B. You can use M-BAC2 as a transgene. (This makes more sense than using H-BAC2, because you don't need a human protein to be made, you just need a functional copy of Gene B.) Once you have your line of transgenic mice with a randomly inserted transgene of M-BAC2 and the original copy of Gene B on the X chromosome, you can do a gene targeting experiment to knockout the endogenous X chromosome copy of Gene B. If you had not made the transgenic first and gene B was essential for spermatogenesis, then your resulting chimera would have been infertile. You can not efficiently test chimeras because they are tetraparental, so you need a fertile chimera to actually make a strain missing gene B.

For the knockout, you would target a neo cassette with flanking regions of Gene B to knockout Gene B in a male ES cell of a mouse with black fur. Since it is a male ES cell, there is only one copy of Gene B anyway. These ES cells would be put back into a 4 day embryo of a foster mother with white fur. Some of the resulting pups would be chimeric. Chimeric males will be bred to mice with white fur to get females who are carriers of the gene B knockout. Finally breed for males with the knockout but that have no transgene and test to see if they are sterile or not.

(b) If Gene B is required in spermatogenesis, all males carrying the gene B knockout and no transgenic copy of Gene B will be infertile and will not produce sperm.

(c) If Gene B is not required, spermatogenesis will still take place and the mice will be fertile.

(d) Gene B and Gene C are right next to each other on the mouse X chromosome. Since they are so tightly linked, it will be extremely difficult to find a mouse that had a crossover event between the two genes which would produce a double knockout.

(e) To produce the double knockout mouse you would need to isolate ES cells from your mouse carrying a neo replacement of Gene B on the X chromosome but a transgenic copy of Gene B elsewhere in the genome. (You made this mouse in part a.) You could then use the male ES cells isolated to knockout Gene C. The ES cell now has Genes B and C deleted on the X chromosome, but an extra transgenic copy of B restoring fertility. From here, you can make a chimera and breed to find animals that have lost the transgene in the same manner as part a. (Note: this could also have been done the opposite way, with a mouse who has endogenous gene C knocked out, but a transgenic copy of C restoring fertility.)

7.03 Problem Set 6

Due before 5 pm on ~~**Wednesday November 17**~~ Thursday, Nov 18
Hand in answers in recitation section or in the box outside the class

1. Consider an autosomal recessive disease that affects one in every 3,000 children in the U.S. Assume that the population is in Hardy-Weinberg equilibrium.

(a) What is the frequency (in the U.S. population) of the allele that is associated with the disease?

(b) What is the frequency of heterozygotes in the U.S. population?

What is the probability that a child born in the U.S. will have the disease if:

(c) Neither the mother nor the father has the disease, and there is no history of the disease in either family.

(d) The mother has the disease. The father does not have the disease and has no family history of the disease.

(e) The mother does not have the disease but has an affected brother. The father does not have the disease and has no family history of the disease.

(f) The mother has the disease. The father does not have the disease and has no family history of the disease, but he is an immigrant from a population (itself in Hardy-Weinberg equilibrium) where the disease affects one in every 11,000 children.

2. In a large but isolated human population, a striking phenotype is observed in 1 in every 1000 births: affected babies experience irregular heart rhythms during the first day after birth, within only 20% of affected babies surviving. Those babies that do survive this dangerous first day are subsequently unaffected and have normal lives thereafter. This heart rhythm defect shows autosomal recessive inheritance, and the population is in Hardy-Weinberg equilibrium for this gene.

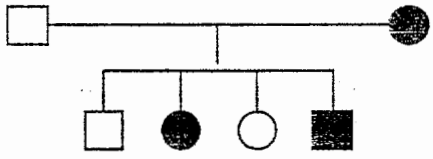
(a) Calculate the mutation rate for this gene in this population.

(b) Suppose that a therapy is devised and implemented so that 60% of affected babies survive the heart rhythm problems experienced during the first day after birth. Calculate the new frequency (at steady-state, after many generations) for the mutant allele in this population.

(c) Suppose that no such therapy is devised but that the mutation rate for this gene doubles in this population. Calculate the new frequency (at steady-state, after many generations) for the mutant allele in this population.

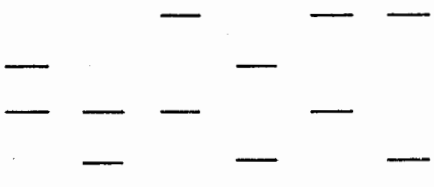
3. Your colleague, a human geneticist, is conducting genetic linkage studies of an autosomal dominant disease whose chromosomal location has not been firmly established. Your colleague is presently focused on two SSR markers that may be linked to each other and to the disease. Here are two families in which some individuals are affected:

Family 1

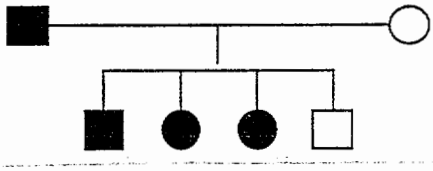


SSR 1

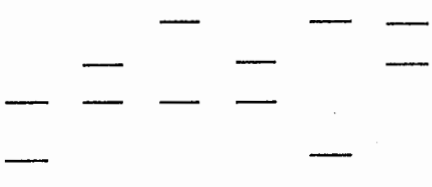
SSR 2



Family 2



SSR 1



SSR 2

Problem Set #6 Solutions

1. a) $f(a/a) = q^2 = 1/3000 = 3.3 \times 10^{-4} \Rightarrow$ **$q = 0.018$**

b) f(heterozygotes) = f(A/a) = 2pq = 2q (because in rare diseases $p \approx 1$)

= **$2q = 0.036$**

c) p(child affected) = p(Mom is carrier)*p(Mom passes on allele)* p(Dad is carrier)*p(Dad passes on allele)

= 2q * 1/2 * 2q * 1/2

= **$q^2 = .00033$**

d) p(child affected) = p(Mom passes on allele) * p(Dad is carrier) * p(Dad passes on allele)

= 1 * 2q * 1/2

= **$q = 0.018$**

e) p(child affected) = p(Mom is carrier)*p(Mom passes on allele)*p(Dad is carrier)*p(Dad passes on allele)

= 2/3 * 1/2 * 2q * 1/2

= **$1/3q = 0.006$**

f) p(child affected) = p(Mom passes on allele) * p(Dad is carrier) * p(Dad passes on allele)

= 1 * 2q * 1/2

= **q**

But **q** has changed because the father is from a different population:

$f(a/a) = q^2 = 1/11{,}000 = 9.1 \times 10^{-5} \Rightarrow$ **$q = 0.0095$**

2. a) $f(a/a) = q^2 = 1/1000 \Rightarrow q = 0.032$

at steady-state: $q = \sqrt{\frac{\mu}{S}} \Rightarrow \mu = Sq^2$

- since 20% survive, the selective disadvantage (S) is 80%, or 0.8

$\mu = Sq^2 = 0.8 * (0.032)^2 \Rightarrow$ **$\mu = 0.0008$**

b) Now that the new survival rate is 60%, the selective disadvantage is 40%, or 0.4

$$q = \sqrt{\frac{\mu}{S}} = \sqrt{\frac{0.0008}{0.4}} = 0.045$$

c) Without the therapy in part (b), the selective disadvantage remains at 80%, or 0.8

$$q = \sqrt{\frac{2\mu}{S}} = \sqrt{\frac{2*0.0008}{0.8}} = 0.045$$

3) For each of these families, the phase is unknown. This means that we do not know which length allele of the SSR is linked to the disease. We can not assume the phase, So to determine the LOD, we must take both into consideration.

$$LOD = \log_{10}\left[\frac{\text{probability data having arisen if linked at theta = 0}}{\text{probability data having arisen if totally unlinked}}\right]$$

Since the phase is unknown:

$$LOD = \log_{10}\left[\frac{0.5\,(\text{prob phase I}) + 0.5\,(\text{prob phase 2})}{\text{probability data having arisen if totally unlinked}}\right]$$

The factor of 0.5 represents that each phase is just as likely.

a) Let's name the alleles A, B, C, D. Where allele A is the highest and D is the lowest. Not knowing which phase is correct, let's declare that phase I represents allele D segregating with the disease and allele C segregating with the wildtype version of the disease. Phase 2 will represents allele D segregating with the wildtype allele of the disease and allele C segregating with the Disease. Note: we do not need to consider the mothers alleles because she is not affected.

The probability of phase 1 when theta =0.is 0.5, with no recombination there is a one half chance that the child will inherit C and the wildtype version and a 0.5 chance the child will inherit D and the disease. There are a total of four children each with either allele Dand the disease or allele C without the disease. The probability in phase 2 of getting D and the disease or C and no disease with no recombination is 0.

$$LOD = \log_{10}\left[\frac{0.5\,(0.5)4 + 0.5\,(0)4}{[(0.5)(0.5)]4}\right] = 0.903$$

b) $$LOD = \log_{10}\left[\frac{0.5\,(0.5)4 + 0.5\,(0)4}{[(0.5)(0.5)]4}\right] = 0.903$$

c) $$LOD = \log_{10}\left[\frac{0.5\,(0.5)4 + 0.5\,(0)4}{[(0.5)(0.5)]4}\right] = 0.903$$

d) $$LOD = \log_{10}\left[\frac{0.5\,(0.5)4 + 0.5\,(0)4}{[(0.5)(0.5)]4}\right] = 0.903$$

e) $$LOD = \log_{10}\left[\frac{0.5\,(0.5)4 + 0.5\,(0)4}{[(0.5)(0.5)]4}\right] = 0.903$$

Because both mom and dad give informational meiosis products, multiply 0.903 by 2 = 0.186

f) $$LOD = \log_{10}\left[\frac{0.5\,(0.5)4 + 0.5\,(0)4}{[(0.5)(0.5)]4}\right] = 0.903$$

"ditto." ↑

g) You should add these two parts together because you are looking at the same marker and the same gene.

h) You can not add these two parts together because they are different markers and may not be the same distance apart.

i) You can add these two together because they are the same markers.

j) Right now all three seem to be absolutely linked, but you can not yet publish a conclusion based on the linkage between the disease and either SSR. You must keep looking for more data to get a LOD score of greater than three.

(Note: if you add parts e and f together, LOD > 3, which means that SSR1 and SSR2 are linked, but that doesn't really tell us much.)

7.03 Problem Set 7

Due before 5 pm on **Friday December 3**
Hand in answers in recitation section or in the box outside the class

1. Johnny Lunchbucket and Betty Juicebox were both raised by single mothers. They met in 7.03 and fell in love calculating LOD scores together. They have been married for several years and want to start a family. Recently, however, they discovered that they share the same deadbeat father. Thus, Johnny and Betty are half-brother and half-sister.

a) If Johnny and Betty decide to have a child, what would the inbreeding coefficient be?

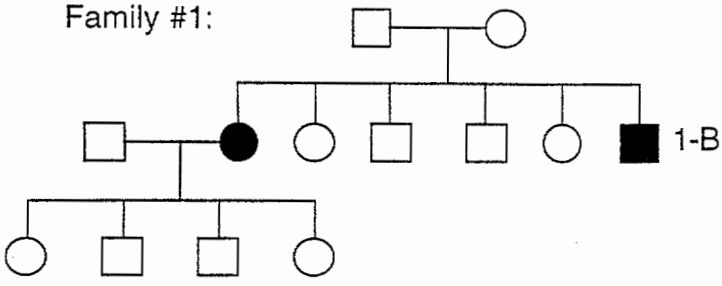
b) They seek the advice of a genetic counselor, who describes the increased risk of autosomal recessive traits among inbred children. To illustrate the risks, the counselor tells them about a rare autosomal recessive disease known as *dormus in lecturitis*. The counselor tells Johnny and Betty that the probability of their child being afflicted with *dormus in lecturitis* is 2×10^{-4}. What is the frequency of the mutant allele in the population?

c) The genetic counselor also tells them that 200 times as many children with *dormus in lecturitis* come from random matings as from half-sibling marriages. What is the frequency of half-sibling marriages in this population?

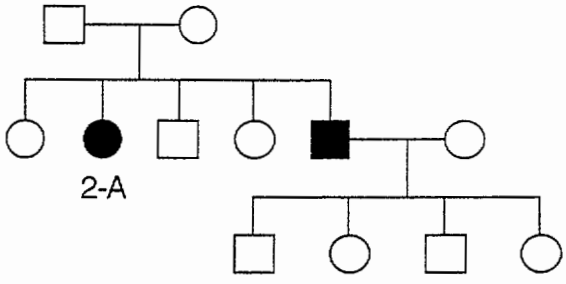
d) Betty and Johnny are particularly concerned about genetic load and the risk of a stillbirth or neonatal death. The genetic counselor informs them that their father was likely to carry 2 recessive lethal alleles, and that the frequency of stillbirth or neonatal death in children of unrelated parents is 0.04. What is the likelihood that their child would be stillborn or die neonatally?

2. Childhood deafness is often hereditary. Consider two pedigrees in which some individuals were deaf from birth:

Family #1:



Family #2:



a) What is the likely mode of inheritance of this deafness from birth?

b) The affected male from Family #1 (individual 1-B) and the affected female from Family #2 (individual 2-A) attend the same school for deaf children, and they ultimately marry and have two children. Both children have normal hearing. Provide a likely genetic explanation for their children having normal hearing.

3. Congenital dislocation of the hip is a common birth defect, with an incidence in the general population of 0.002. Scarlet fever is caused by a bacterial infection.

One of the following two sets of twin data is for congenital dislocation of the hip. The other set of twin data is for scarlet fever. Your task is to figure out which set of data goes with which disease.

	Concordance in MZ twins	DZ twins
Disease #1	23%	2.3%
Disease #2	55%	47%

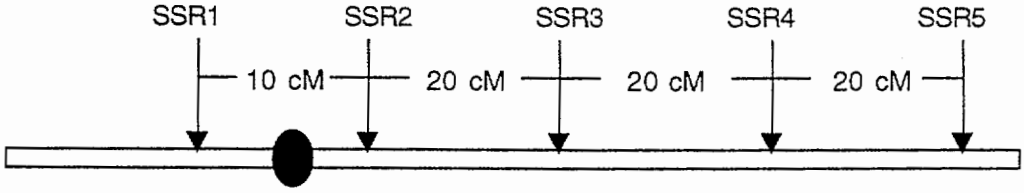
a) Is Disease #1 more likely to be congenital dislocation of the hip, or scarlet fever? Justify your answer.

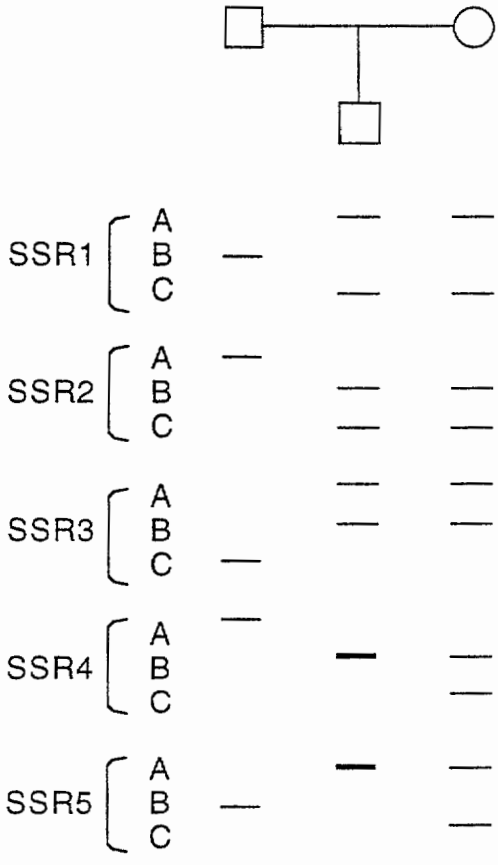
b) Is Disease #2 more likely to be congenital dislocation of the hip, or scarlet fever? Justify your answer.

c) Is congenital dislocation of the hip a purely genetic disorder, or is there some "environmental component"? (In considering a possible "environmental component, realize that the defect is present at birth.) Justify your answer.

d) Is the data consistent with one gene predisposing to congenital dislocation of the hip? Justify your answer.

4. After taking 7.03, you're so interested in human genetics that you sign up for a UROP to study meiotic nondisjunction. You prepare a DNA sample from a young boy with Klinefelter Syndrome, which means that the boy has two X-chromosomes and a Y chromosome. You also prepare DNA samples from the boy's parents. You then type the boy and his parents for a series of SSRs distributed along the X chromosome:





a) In which parent did nondisjunction occur?

b) In which division of meiosis did nondisjunction occur?

c) Sketch the meiotic event in which nondisjunction occurred. Your drawing should include the SSRs present along the X chromosome.

d) Later you have the opportunity to study a boy with one X chromosome and two Y chromosomes. You realize that you do not even need to use SSRs or other genetic markers to figure out the meiotic division in which nondisjunction occurred. In which parent and at which meiotic division did nondisjunction occur?

Solutions to 7.03 Problem Set 7

1. A)F= $(1/2)^4 * 2 = 1/8$
   B)Frequency of half sib with disease = $2x10^{-4}$
   $F * q = f$
   $1/8\ (q) = 2x10^{-4}$
   $q = 1.6 \times 10^{-3}$
   C)frequency of random marriages * q^2 = 200 * frequency of half-sib marriages * F q
   Assume frequency of random marriages = 1
   Assume number of children from normal marriage is equal to number from half-sib marriage
   q^2 = 200 * frequency of half-sib marriages * 1/8 * q
   Frequency of half-sib marriages = (8/200)q
   Frequency of half-sib marriages = $4x10^{-4}$
   D)p(homozygous for an allele carried by father) = 1/16
   Recessive lethal alleles carried by father = 2
   P (homozygous for a recessive lethal allele carried by father = (1/16)*2 =.125
   With random mating p (still birth) = 0.04
   For these half-sibs p(stillborn child) = 0.165
2. A) The most likely mode of inheritance is autosomal recessive.
   B) One likely explanation is that these individuals have deafness caused by two different mutations. Therefore we see complementation in the offspring and no deafness.
3. A)Disease #1 is congenital dislocation of the hip. This congenital birth defect is likely to have a genetic component. In disease #1 the concordance rate is 10 times higher in monozygotic twins then in dizygotic twins. This difference in concordance rates is consistent with a disease like hip dislocation.
   B)Disease #2 is likely to be Scarlet fever as there is little difference between the concordance rates. This result is consistent with a disease that is the result of environmental factors like infection.
   C)Clearly congenital hip dislocation is not a purely genetic disorder. If it were, we would expect 100% concordance in monozygotic twins. However, we only see 23% concordance suggesting there is some environmental component to the disease.
   D)This data is not consistent with one gene predisposing to dislocation. If only one gene was responsible, we would not expect to see a ten fold difference in concordance between monozygotic and dizygotic twins.

4. A) Nondisjunction occurred in the mother. This is evident from the fact the affected individual only carries markers also carried by his mother. He does not carry any of his fathers SSR markers, indicating both of his X chromosomes came from his mother.
B)Nondisjunction occurred at meiosis one. If nondisjunction had occurred at meiosis two, the boy would have been homozygous for the majority of his SSR markers.
C)See below.
D)As the boy has two Y chromosomes, nondisjunction must have occurred in the father. In order for two Y chromosomes to be passed on to his son, nondisjunction must have occurred in meiosis two. If nondisjunction had occurred in meiosis one, the father would have passed on an X and a Y chromosome.