

7.03 Problem Set 1

Due before 5 PM on Thursday, September 20, 2001

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

- 1.** You have isolated 10 new mutant yeast strains that are defective in synthesis of leucine, an amino acid. These Leu⁻ mutants do not grow on minimal medium, but they do grow on minimal medium supplemented with leucine. Your Leu⁻ mutants (numbered 1 through 10) were all isolated in a strain of mating type **a** (MAT **a**). As it turns out, your high school classmate, now at Harvard, has independently isolated 10 new Leu⁻ mutants (numbered 11 through 20), all in a strain of mating type α (MAT α). You and your high school classmate decide to combine your resources and determine how many different genes are represented by your 20 mutant strains. You cross each of the MAT **a** strains to each of the MAT α strains, and you include crosses to the appropriate wild-type strains. Your experimental observations are shown in the table below, where (-) indicates diploids that did not grow on minimal medium and (+) indicates diploids that did grow on minimal medium.

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

(a) What property do mutants 2 and 16 share?

(b) Which mutations do you know to be in the same gene?

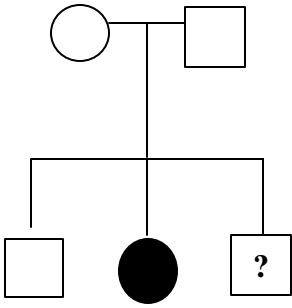
- (c) Could mutations 3 and 16 be in the same gene?
- (d) Based on these experiments, what is the minimum number of genes required for leucine synthesis?
- (e) What is the maximum number of genes that these 20 mutants could represent?

2. Being a well-rounded geneticist, you also maintain a colony of mice.

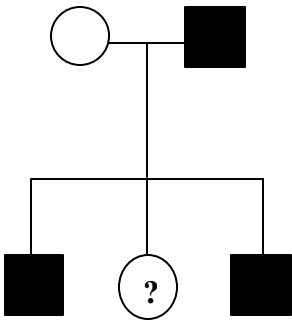
- (a) One day you spot a mouse in your colony with a novel and interesting phenotype: a kinked tail. You breed the kinked-tail mouse (a male) with several wild-type females and observe that about half of the offspring (both males and females) have kinked tails and half have normal tails. Is the kinked-tail mutation dominant or recessive to wild type?
- (b) When two of the kinked-tail offspring from part (a) are crossed, what fraction of the resulting mice would you expect to have kinked tails?
- (c) When you cross kinked-tail offspring from part (a), you find that one third of the resulting kinked-tail males produce no sperm and thus are sterile. The other two thirds of the resulting kinked-tail males (and all of the normal-tail males and all of the females) are fertile. Propose a model to account for these findings.
- (d) Not to be outdone, your high school classmate informs you that he has isolated a pure-breeding mouse strain in which males produce no sperm but have normal tails, and in which females are phenotypically normal (fertile; normal tails). You explain to your high school classmate that this is impossible. Why?
- (e) Eventually your classmate convinces you that the male sterile mutant that he has discovered displays autosomal recessive inheritance. You decide to test whether your male-sterile mutant and your classmate's male-sterile mutant are in the same gene. Diagram the series of crosses required to conduct an appropriate complementation experiment, including the expected ratios (in the final generation) of infertile to fertile males for two competing models: 1) two mutations in the same gene and 2) mutations in two different genes on different chromosomes.

- (f) Breeding experiments involving the two mutant strains are simplified when your classmate discovers that his mutant, with recessive male sterility, also displays a dominant phenotype in both sexes: short tail. Through extensive breeding, your classmate identifies a series of short-tail females that, when crossed with wild type, produce exclusively short-tail progeny. You cross these short-tail females with fertile, kinked-tail males and observe the following offspring: 35 short-kinked-tail females, 32 short-nonkinked-tail females, 28 short-kinked-tail males, 36 short-nonkinked-tail males (all offspring fertile). Propose a model to account for these ratios.
- (g) To test your model from part (f), you select, from among the offspring from part (f), short-kinked-tail females and short-kinked tail males, and you cross them. If your model is correct, what phenotypic classes do you expect to observe, and in what ratios?
- 3.** Each of the families below exhibits a different, extremely rare genetic disorder. Individuals expressing the trait (the disorder) are indicated by solid symbols. Assume that no new mutations have arisen in any of the individuals shown. Consider the following possible modes of inheritance: (i) X-linked recessive with complete penetrance, (ii) autosomal recessive with complete penetrance, (iii) autosomal recessive with 70% penetrance, (iv) autosomal dominant with complete penetrance, (v) autosomal dominant with 70% penetrance. For each pedigree state which, if any, of these five modes of inheritance are not possible. For the modes of inheritance that are possible, calculate the probability that the individual indicated by a “?” is affected.

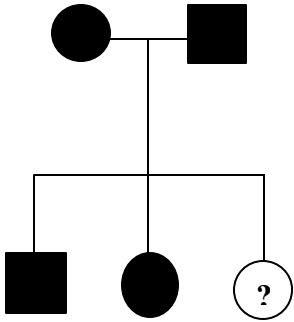
(a)



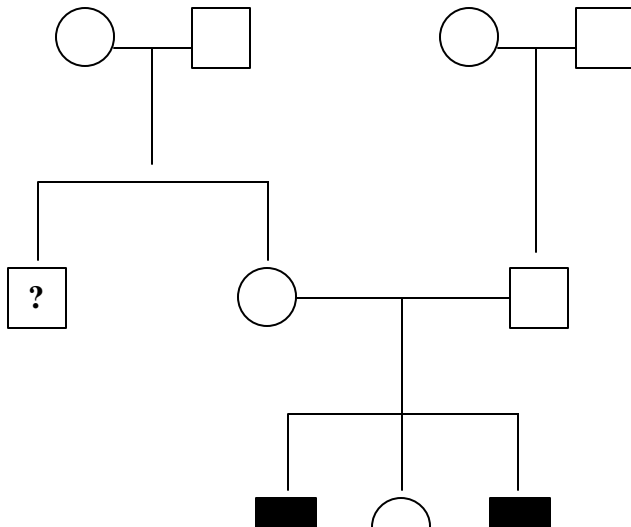
(b)



(c)



(d)



7.03 Problem Set 1 Answer Key
Fall 2001

- 1a. Mutants 2 and 16 have dominant Leu- alleles. Remember dominant or recessive alleles are tested by crossing a mutant to wildtype and observing the phenotype of the heterozygote.
- 1b. The following mutants are in the same genes: (1,19), (3,8,18), (4,13), (5,6,11,15), and (10,14). Mutant alleles in these strains fail to complement each other in the heterozygote. Nothing can be said definitively about strain 2 and 16 because complementation tests cannot be done with dominant alleles.
- 1c. Yes. Different alleles of the same gene can be both dominant and recessive. Again, complementation tests cannot be done with dominant alleles.
- 1d. There is a minimum of seven genes. Five groups defined by recessive alleles that noncomplement each other. One group minimally defined by mutants 7 and 9, and a second group defined by mutants 12, 17, and 20. 7 and 9 complement 11-20, excluding 16, and thus are not a part of any previously defined complementation group. They may however be in the same gene because we do not have data about complementation between mutants 7 and 9. Similar logic follows that 12,17, and 20 represent at least one new gene. These groups represent at least 2 new genes because they do not complement each other.
- 1e. There is a maximum of twelve genes. Five complementation groups in addition to mutants 7, 9, 12, 17, 20 which don't belong to any defined complementation groups and mutants 2 and 16 which contain alleles that have a dominant phenotype and may represent new genes.

- 2a. The kinked tail allele is dominant. Parental are (male) $t^K t^+$ x $t^+ t^+$ (female) where t^K = allele with dominant kinked tail phenotype, t^+ = allele with normal tail phenotype (wild type).

$$\begin{array}{ccc}
 \frac{t^K}{t^+} & \times & \frac{t^+}{t^+} \\
 & \Downarrow & \\
 1/2 \ t^K/t^+ & & 1/2 \ t^+/t^+
 \end{array}$$

- 2b. There will be $3/4$ kinked tail mice. You will get at 3:1 Mendelian ratio of kinked to non-kinked. All kinked tail offspring from first cross will be $t^K t^+$.

$$\begin{array}{ccc}
 \frac{t^K}{t^+} & \times & \frac{t^K}{t^+} \\
 & \Downarrow & \\
 1: \frac{t^K}{t^K} & 2: \frac{t^K}{t^+} & 1: \frac{t^+}{t^+}
 \end{array}
 \quad \text{Genotypic ratios}$$

- 2c. Homozygous t^K males are sterile. Thus $1/3$ of the kinked tail animals ($t^K t^K$) are sterile while $2/3$ ($t^K t^+$) of the kinked males are fertile. The t^K allele is dominant to t^+ for the kinked tail phenotype and recessive for the male specific sterility phenotype.
- 2d. Pure breeding males would be sterile! To have a pure breeding strain, you must be homozygous for all loci. In order for our high school friend to have a strain that is propagating, this would not be the case.
- 2e. Mate kinked tail males ($t^K t^+$) to normal appearing females homozygous for the male specific sterility gene (ss). S = wildtype allele, that is non-sterile, s = allele conferring a recessive male sterility phenotype. First, you must identify ss females. This requires a series of crosses with the inbred sterile strain from our friend.

Take a female and a fertile male and mate them in isolation. Their genotypes are ambiguous so below are the possibilities:

$$S/S \text{ or } S/s \quad \times \quad S/S \text{ or } S/s \text{ or } s/s$$

In the next generation, you want to see sterile males meaning that the male and the female were of these possible genotypes:

$$S/s \quad \times \quad S/s \text{ or } s/s$$

Taking the parental male that is guaranteed to be S/s, mate to several of his offspring that are female until you find a female that gives birth to roughly 50% sterile and 50% fertile males. Below is your ideal cross:

$$\begin{array}{ccc}
 S/s & \times & ss \\
 & \Downarrow & \\
 1/2 S/s & \text{and} & 1/2 ss
 \end{array}$$

Then take that female, which is of the genotype ss, and use it in your cross.

$$\begin{array}{ccc}
 \text{Possibility 1: } t^K/t^+ & \times & s/s \\
 & \Downarrow & \\
 1/2 t^K/s & \text{and} & 1/2 t^+/s
 \end{array}$$

If t^K and s are alleles of the same gene the genotype of all kinked tailed males will be t^K/s . t^K and s would not complement and thus all kinked tailed males would be sterile.

$$\begin{array}{ccc}
 \text{Possibility 2: } t^K/t^+S/S & \times & t^+/t^+s/s \\
 & \Downarrow & \\
 1/2 t^K/t^+S/s & \text{and} & 1/2 t^+/t^+S/s
 \end{array}$$

If t^K and s are alleles of different genes, then the genotype of all kinked tailed males will be t^Kt^+Ss . For the recessive male sterility phenotype t^K and s should complement, thus all kinked tailed males would be fertile. Note that parental males are t^Kt^+ and could transmit either allele, but we know that the t^K allele has been inherited by following the dominant kinked tail phenotype.

- 2f. These two traits are caused by mutations in different genes. The two alleles complement each other. Parental genotypes are homozygous short tailed females ($t_1^S t_1^S$) and ($t^K t^+$) heterozygous kinked tailed males. Normal segregation of dominant alleles would produce equal frequencies of each class of progeny.

$$\begin{array}{ccc}
 t^K/t^+ t_1^+/t_1^+ & \times & t^+/t^+ t_1^S/t_1^S \\
 & \Downarrow & \\
 1/4 t^K/t^+ t_1^S/t_1^+ & & 1/4 t^K/t^+ t_1^S/t_1^+ \\
 1/4 t^+/t^+ t_1^S/t_1^+ & & 1/4 t^+/t^+ t_1^S/t_1^+
 \end{array}$$

If the t^S and t^K were alleles of the same gene then the heterozygote short-kinked-tail males would be sterile due to non-complementation. Since we know that all offspring are fertile, the two recessive sterile alleles must be in different genes.

- 2g. Taking a short-kinked-tail male and a short-kinked-tail female from the F_1 of part f. We have the cross:

$$t^K/t^+ t_1^S/t_1^+ \quad \times \quad t^K/t^+ t_1^S/t_1^+$$

$$\Downarrow$$

We would expect the following phenotypic ratios of progeny for the two dominant mutations:

9	kinked short tailed females	$(t^K_t_1^S_)$
3	kinked normal tailed females	$(t^K_t_1^+t_1^+)$
3	normal short tailed females	$(t^+t^+t_1^S_)$
1	normal tailed females	$(t^+t^+t_1^+t_1^+)$

The following are the genotypic ratios taking into account the sterility of males:

5	kinked short tailed males sterile	$(t^Kt^Kt_1^St_1^S, t^Kt^Kt_1^St_1^+, t^Kt^+t_1^St_1^S)$
4	kinked short tailed males fertile	$(t^Kt^+t_1^St_1^+)$
2	kinked normal tailed males fertile	$(t^Kt^+t_1^+t_1^+)$
1	kinked normal tailed males sterile	$(t^Kt^Kt_1^+t_1^+)$
2	normal short tailed males fertile	$(t^+t^+t_1^St_1^+)$
1	normal short tailed males sterile	$(t^+t^+t_1^St_1^S)$
1	normal tailed males fertile	$(t^+t^+t_1^+t_1^+)$

t^K = allele with dominant kinked tail and recessive male sterility phenotypes

t^+ = wild type kinked tail allele

t_1^S = allele with dominant short tail and recessive male sterility phenotypes

t_1^+ = wild type short tail allele

3. Note: In answering these problems, we make the primary assumption that the affected individual has the appropriate genotype under the conditions presented. Secondary assumptions are that since the mutation is extremely rare, affected individuals carrying dominant mutation are heterozygous. When dealing with a recessive disease gene, we assume that any individual mating into the pedigree from the general population is not a carrier. Secondary assumptions are held to be true UNLESS they violate the logic of the primary assumption. Remember that generations are numbered with roman numbers from top to bottom and that individuals are capital letters in alphabetical order from left to right.
- i. x-linked recessive
 - ii. autosomal recessive
 - iii. autosomal recessive with 70% penetrance
 - iv. autosomal dominant
 - v. autosomal dominant with 70% penetrance
- a.
 - i. Not possible because individual I-B is not affected.
 - ii. Possible, individual II-C has a $1/4$ chance of being affected. $1/2$ chance of inheriting the recessive allele from each parent. By the product rule, $1/2 \times 1/2 = 1/4$.
 - iii. Possible, individual II-C has a $7/40$ chance of being affected. We already know the probability that individual II-C will inherit the affected genotype. His chances of being affected is the probability of inheriting the affected genotype times the probability of being fully penetrant which is $7/10$. So, $1/4 \times 7/10 = 7/40$.
 - iv. Not possible, because neither parent is affected.
 - v. Possible, individual II-C has a $7/20$ chance of being affected. If we are assuming that this disorder has 70% penetrance and is rare, then one parent must be heterozygous for the dominant allele but not fully penetrant. Therefore the probability that the child will inherit the affected allele is $1/2$, but we have to account for the penetrance which is $7/10$. So, $1/2 \times 7/10 = 7/20$.
 - b.
 - i. Possible, individual II-B has a $1/2$ chance of being affected. To make an hypothesis of x-linked recessive fit the pedigree we have to assume that individual I-A is a carrier (heterozygous for the affected allele). This may be confusing with our assumption that this trait is rare in the population, but please note that because this pedigree has few individuals and few generations we don't know the background of individual I-A.
 - ii. Possible, individual II-B has a $1/2$ chance of being affected. The same reasons as i. apply here too.
 - iii. Possible, individual II-B has $7/20$ chance of being affected. The probability of inheriting the affected genotype is $1/2$ and the probability of being fully penetrant is $7/10$. So, $1/2 \times 7/10 = 7/20$.

- iv. Possible, individual II-B has a $1/2$ chance of being affected. Assuming that this is a rare mutation in the population, individual I-B would be a heterozygote for the affected dominant mutant allele. He has $1/2$ chance of transmitting the mutant allele to his daughter.
- v. Possible, individual II-B has a $7/20$ chance of being affected. She has a $1/2$ chance of inheriting the affected allele and a $7/10$ chance of being fully penetrant. So, $1/2 \times 7/10 = 7/20$.
- c. i. Possible, individual II-C has a 100% chance of being affected.
- ii. Possible, individual II-C has a 100% chance of being affected.
- iii. Possible, individual II-C has a $7/10$ chance of being affected. She will inherit both affected alleles from her homozygous affected parents, but she has a $7/10$ chance of being fully penetrant.
- iv. Possible, individual II-C has a $3/4$ chance of being affected. Again assuming a rare disease in the population, both parents will be heterozygous affected. Therefore individual II-C has a $3/4$ chance of being affected. She can be either homozygous for the dominant allele or heterozygous to be affected.
- v. Possible, individual II-C has a $21/40$ chance of being affected. She has a $3/4$ chance of inheriting an affected genotype and a $7/10$ chance of being fully penetrant. So, $3/4 \times 7/10 = 21/40$.
- d. i. Possible, individual II-A has a $1/2$ chance of being affected. Males in the generation III are affected, so we can assume that their mother was a carrier. For their mother (individual IIB) to be a carrier and not have an affected father (individual I-B), individual I-A must have been a carrier too. Therefore individual II-A has $1/2$ chance of inheriting an affected X from his mother.
- ii. Possible, but extremely unlikely.
- iii. Possible, but again extremely unlikely.
- iv. Not possible because there are no affected individuals in generations I or II.
- v. Possible, individual II-A has a $7/40$ chance of being affected. In order for individuals in generation III to be affected, there would need to have been individuals in generations I and II that were carriers of the dominant mutant allele but not fully penetrant. In addition, we do not know which side of the pedigree the dominant mutant allele has come from (either parents I-A or B had it or parents I-C or D). Therefore, there is a $1/2$ chance that the left side of the pedigree had the affected dominant allele (i.e. from parents I-A or B). If that side of the pedigree had the affected allele, then there is a $1/2$ chance that a

heterozygous parent passed the affected allele on to their child (individual II-A). In addition, we have the penetrance of 7/10. So, $1/2 \times 1/2 \times 7/10 = 7/40$.

7.03 Problem Set 2

Due before 5 PM on Thursday, September 27

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

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

<u>Phenotype</u>	<u>Number</u>	<u>Phenotype</u>	<u>Number</u>
a b c d	10	+ b c d	14
+ + + +	13	a + + +	11
a b c +	217	a b + +	25
+ + + d	198	+ + c d	29
a b + d	5	a + + d	203
+ + c +	2	+ b c +	214
a + c d	28	+ b + d	3
+ b + +	26	a + c +	2

- (a)** Determine the two-factor cross distance between markers **c** and **d** in cM.
- (b)** What can you conclude about the phenotypes of the two true-breeding parental lines. If it is not possible to specify their exact phenotype, note the nature of the ambiguities that remain.
- (c)** Draw a genetic map showing the relative positions and genetic distances between the genes for each of the four traits. Linked genes should be grouped together on the same chromosomal segment and any unlinked genes should be placed on a different chromosomal segment.
- (d)** If your best estimate for the distance between markers **c** and **d** is different than the two factor distance given in part (a), explain what additional data allowed a better map distance to be obtained and why.

2. In the late 1970s Gerald Fink (who is presently director of the Whitehead Institute) wanted to construct a yeast strain with two different mutations in the *Leu2* gene. To do this he crossed a strain that had a *Leu2-3* mutation to a strain with a *Leu2-112* mutation. (Both of these parental strains are Leu^- ; that is, they will not grow unless leucine is provided in the medium.) The tetrads from this cross were of two types:

<u>Type 1</u>	<u>Type 2</u>
4 Leu^-	3 Leu^- : 1 Leu^+

49 tetrads were of Type 1, while only 1 tetrad was of Type 2.

(a) Explain why this cross produced only two of the three possible tetrad types.

(b) The desired *Leu2-3 Leu2-112* double mutant was one of the three Leu^- spores in the Type 2 tetrad. Explain why complementation tests couldn't be used to find the double mutant.

(c) Describe a procedure that you would use to identify the desired double mutant. Be as specific as possible about the crosses that you would perform, how you would analyze the resulting tetrads, and how many tetrads you would analyze for each cross .

(d) Say that you have isolated a new Leu^- mutant that you call *LeuX*. In a cross between *LeuX* and the *Leu2-3 Leu2-112* double mutant the following tetrad types are obtained:

<u>Type 1</u>	<u>Type 2</u>	<u>Type 3</u>
4 Leu^-	3 Leu^- : 1 Leu^+	2 Leu^- : 2 Leu^+

From 50 tetrads you obtain 35 of Type 1, 13 of Type 2, and 2 of Type 3. What is the relationship between the *LeuX* mutation and the *Leu2* gene?

3. Imagine that you are studying a new type of phage. The chromosome of this phage has been measured to be 5×10^4 base pairs in length and by addition of many different map distances the total genetic length of the chromosome is estimated to be 100 m.u.. You have isolated two different mutants that appear to be in the same gene that cause the formation of small plaques. You call these mutants **sm-1** and **sm-2**.

(a) You cross an **sm-1** mutant with an **sm-2** mutant and among the progeny phage you find that 11 out of 1000 plaques have normal large plaques while the rest are small. What is the distance between the **sm-1** and **sm-2** mutations in m.u.?

(b) You have identified the protein product of the phage **sm** gene as a protein of 50 kDa (one kDa = 1000 daltons). On examination of the **sm** protein in lysates from mutant phage, you find that the **sm-1** mutant produces a protein of 45 kDa while the **sm-2** mutant produces a protein of 50 kDa. What type(s) of mutations would explain the behavior of the **sm-1** mutant? What type(s) of mutations would explain the behavior of the **sm-2** mutant? In thinking about your answer, you may find it useful to consider that the average molecular weight of an amino acid is approximately 110 Da.

(c) Based on everything that you know about the **sm-1** and **sm-2** mutants (effects on protein product, map distances, and the relationship between the genetic and physical lengths of the phage chromosome) draw an approximate map of the **sm** gene showing the coding sequence, the direction of transcription and the approximate positions of the **sm-1** and **sm-2** mutations.

(d) You have isolated a new phage mutant that gives clear plaques that you call **cl-1** (the wild type phage normally produces turbid plaques). You construct a **cl-1 sm-1** double mutant (which makes small clear plaques) and then you cross this double mutant to an **sm-2** single mutant. Of 1000 plaques produced from this cross, 10 are large. Of these large plaques, 9 are clear and 1 is turbid. Which side of the **sm** gene is the **cl** gene (relative to the direction of transcription of **sm**)? Approximately how far away from the **sm** gene is the **cl** gene in base pairs?

Problem 1

a. We only need to look at markers **c** and **d**.

c +	217 + 2 + 214 + 2 = 435	Parental
+ d	198 + 5 + 203 + 3 = 409	Parental
c d	10 + 28 + 14 + 29 = 81	Recombinant
++	13 + 26 + 11 + 25 = 75	Recombinant

$$\begin{aligned}\text{Distance} &= 100 \times \# \text{ recombinant} / \text{total} \\ &= 100 \times 156 / 1000 \\ &= 15.6 \text{ cM}\end{aligned}$$

b. The two true-breeding lines were crossed to give the F1 females that were used in this test cross to homozygous recessive males to give the data that we see. By analyzing the data to find what markers are linked and unlinked, we can determine the true-breeding parental phenotypes. To determine linkage we need to analyze the markers pairwise.

i. **a** and **b**

ab	10 + 217 + 5 + 25 = 257
++	13 + 198 + 2 + 29 = 242
a+	28 + 11 + 203 + 2 = 244
+b	26 + 14 + 214 + 3 = 257

All classes are represented equally, so conclude that **a** and **b** are unlinked.

ii. **a** and **c**

Two factor analysis shows that **a** and **c** are unlinked.

iii. **a** and **d**

Two factor analysis shows that **a** and **d** are unlinked.

iv. **b** and **c**

bc	10 + 217 + 14 + 214 = 455	Parental
++	13 + 198 + 11 + 203 = 425	Parental
b+	5 + 26 + 25 + 3 = 59	Recombinant
+c	2 + 28 + 29 + 2 = 61	Recombinant

b and **c** are linked \Rightarrow dist = $100 \times 120 / 1000 = 12 \text{ cM}$

v. **b** and **d**

bd	10 + 5 + 14 + 3 = 32	Recombinant
++	13 + 2 + 11 + 2 = 28	Recombinant
b+	217 + 26 + 25 + 214 = 482	Parental
+d	198 + 28 + 29 + 203 = 458	Parental

b and **d** are linked \Rightarrow dist = $100 \times 60 / 1000 = 6 \text{ cM}$

vi. **c** and **d**

From part a we know that **c** and **d** are linked \Rightarrow dist = 15.6 cM

From above analysis we conclude that **b**, **c**, and **d** are linked and that **a** is unlinked from these three. The parental classes are

$$\begin{array}{cc} \underline{a} & \underline{bc+} \\ + & ++d \end{array}$$

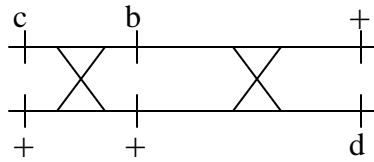
So one parent was $_bc+$ and the other parent was $_++d$. Since **a** is unlinked, we cannot determine which parent is **a**. Therefore there are two possible combinations:

- i. Parent 1: **abc** ii. Parent 1: **bc**
 Parent 2: **d** Parent 2: **ad**

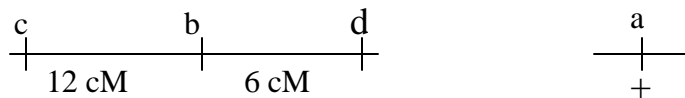
c. To determine the order of the three linked genes, we do a three factor cross.

bcd	10 + 14 = 24	
+++	13 + 11 = 24	
bc+	217 + 214 = 431	Parental, which confirms our answer in part b
++d	198 + 203 = 401	Parental
b+d	5 + 3 = 8	Double Cross-over
+c+	2 + 2 = 4	Double Cross-over
+cd	28 + 29 = 57	
b++	26 + 25 = 51	

To generate the double cross-over classes marker **b** would have to be in the middle.



From the distances calculated in part b, we get the following map:



d. By looking at a three factor cross, we take into account double cross-overs that are not seen in a two factor cross. Without the knowledge of the number of double cross-overs, in part a we underestimate the map distance between the two markers that are the farthest apart.

Problem 2

a. From the cross we obtained:

49	4 Leu ⁻ : 0 Leu ⁻	Parental Ditype
1	3 Leu ⁻ : 1 Leu ⁻	Tetratype

We did not observe any non-parental ditypes, which indicates that there were no double cross-overs. Since the two different mutations are in the same gene the frequency of recombination between them is very small, which explains why we saw only one T (the result of a single cross-over) and no NPDs.

b. From the T tetrad, we observe that in the double mutant the two mutations, leu2-3 and leu2-112, do not complement each other. If we were to cross the three Leu⁻ spores from the T tetrad to a leu2-3 mutant, all three spores would be Leu⁻. Then if we were to cross these spores to a leu2-112 mutant, again all three would be Leu⁻. From this data we cannot determine which spore is the double mutant.

c. Take the three Leu⁻ spores from the T tetrad and cross each of them to wild type (Leu⁺). Then dissect tetrads of the three crosses.

If the spore is a single mutant it should give rise to the following tetrads:

All	2 Leu ⁻ : 2 Leu ⁺	Parental
-----	---	----------

If the spore is the double mutant it should give rise to the following tetrads:

49	2 Leu ⁻ : 2 Leu ⁺	Parental Ditype
1	3 Leu ⁻ : 1 Leu ⁺	Tetratype

The T resulted from a cross-over between leu2-3 and leu2-112. The recombination frequency should be the same as in part a, so we would expect 1 out of 50 tetrads to have a recombination event. Therefore we need to dissect at least 50 (but more is better) tetrads from each cross for a total of 150. The one spore that give rise to a Tetratype is the double mutant, the other two should always give rise to tetrads with a 2:2 pattern.

d. From LeuX x leu2-3leu2-112, we obtained:

35	4 Leu ⁻ : 0 Leu ⁺	PD
13	3 Leu ⁻ : 1 Leu ⁺	T
2	2 Leu ⁻ : 2 Leu ⁺	NPD

PD>>NPD, so the genes are linked. Distance is defined as:

$$D = 100 \times \frac{T + 6NPD}{2\Sigma}$$
$$D = 100 \times \frac{13 + 12}{100}$$

$$D = 25 \text{ cM}$$

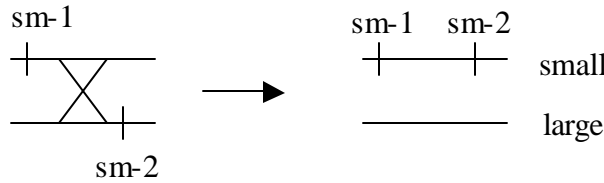
Problem 3

a. sm-1 x sm-2



989 small Parental
 11 large Recombinant

Since we only observe one product of a single cross-over, we need to multiply that number by 2 to get the total number of recombinant progeny.



$$D = 100 \times \frac{\# \text{ recombinant}}{\text{total}}$$

$$D = 100 \times \frac{22}{1000}$$

D = 2.2 m.u.

b. The wild type protein product from the sm gene is 50 kDa. The sm-1 mutation results in a truncated protein by 5 kDa. This could be the result of a mutation, either a base change or a frameshift, that introduced a new stop codon. Because the sm-2 mutation gives a protein product similar to wild type, this mutation is likely a base change. This base change results in a different amino acid, which then affects the function of the protein.

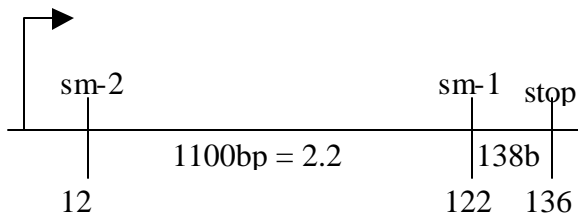
c. Using the information given we can calculate:

$$5 \times 10^4 \text{ bp} : 100 \text{ m.u.} \Rightarrow 500 \text{ bp} : 1 \text{ m.u.}$$

$$50 \text{ kDa} \approx 455 \text{ amino acids} \approx 1365 \text{ base pairs}$$

$$5 \text{ kDa} \approx 409 \text{ amino acids} \approx 1227 \text{ base pairs}$$

Since the sm-1 mutation gives a protein truncated by 5 kDa, we can conclude that the mutation is near the carboxy terminus of the protein. The sm-2 mutation is 2.2 m.u. (= 1100 base pairs) away and is therefore near the amino terminus and transcription start site. We get the following map:



d. cI-1sm-1 x sm-2



990 small

9 large, clear

1 large, turbid

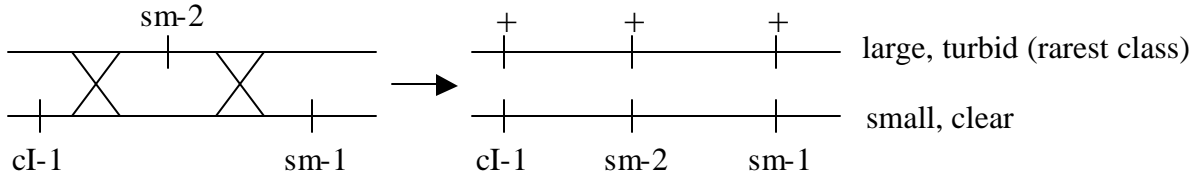
cI-1 sm+

cI+ sm+

Single Cross-over

Double Cross-over

The smallest class of recombinants is the result from a double cross-over. We can use this information to determine the order of the three mutations.

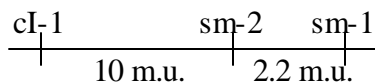


Therefore the cI-1 mutation is to the left of the sm-2 mutation and is upstream of the start of transcription. To calculate the distance between cI-1 and sm-2 we have to use different logic than before. A single cross-over in the interval between cI-1 and sm-2 would generate a small, clear plaque and a small plaque, which we cannot see because they are in the parental class. Therefore to calculate the distance, we need to examine the double cross-overs. A cross-over occurs 10 times in the interval between sm-2 and sm-1 and out of these 10 only one has a cross-over in the interval between cI-1 and sm-2. So the frequency of recombination in this interval is 10%.

Distance = frequency of recombination

$$D = 10 \text{ m.u.}$$

The map must be:



7.03 Problem Set 3

Due before 5 PM on Thursday, October 18

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

1. One way to isolate nonsense suppressor mutations in tRNA genes is to select for the simultaneous reversion of nonsense mutations in two different genes. This selection works because it is extremely unusual to get back mutations in two different genes at the same time. The *E. coli* **his1** and **his2** genes are required for histidine synthesis and strains harboring mutations in either gene will not grow unless histidine is provided in the growth medium.

(a) If you wanted to isolate nonsense suppressor mutations explain why it would be a bad idea to start with a strain that has an amber mutation (TAG) in **his1** and an ochre mutation (TAA) in **his2**.

(b) Say that you have a strain with amber mutations in both **his1** and **his2** and after treatment of this strain with a mutagen that causes G•C to A•T base changes you select for the ability to grow without histidine (**his⁺**). List all of the possible tRNA genes that could have been mutated to generate an appropriate suppressor mutation. For each possibility, Write out the RNA sequence of the anticodon segment of both the wild-type and suppressor tRNAs. Be sure to indicate the 5' and 3' ends. (Remember that in RNA, U [uracil] takes the place of T [thymine]). Also write out the DNA base pairs that encode the anticodon segment of the corresponding wild type and mutant tRNAs. Show both DNA strands with the 5' and 3' ends labeled.

(c) Now, say that you have a strain with ochre mutations in both **his1** and **his2**. In this case you use a mutagen that causes A•T to T•A base changes before selecting for the ability to grow without histidine (**his⁺**). List all of the possible tRNA genes that could have been mutated to generate an appropriate suppressor mutation. As in part **(b)**, give RNA and DNA sequences for both the wild-type and mutant (suppressor) allele of each tRNA.

2. Instead of starting with a double mutant as in problem 1, say that you have a strain with an amber mutation in just **his1**. After mutagenesis with EMS you select **his⁺** revertants by their ability to grow on medium without histidine. In this case, it is necessary to consider seriously the possibility of a back mutation in **his1** as well as extragenic suppressor mutations in tRNA genes.

(a) Explain why it would be very unlikely in this case to acquire an intragenic suppressor mutation in **his1**, at a site different from the mutant amber codon.

In order to distinguish back mutations and extragenic suppressors, you decide to isolate a **Tn5** insertion linked to the original **his1⁻** mutation. 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 **his⁺**). You grow **P1** phage on a mixture of the entire collection of **Tn5** insertion strains and then infect the **his1⁻** mutant and select for Kan^r transductants. Most of the Kan^r transductants are **his⁻**, but one out of 1000 is **his⁺**.

(b) Next you grow **P1** phage on this **his⁺** Kan^r transductant isolated above and infect the original **his1⁻** mutant with the resulting phage. After selecting for Kan^r transductants you test these transductants for their ability to grow in the absence of histidine. You find that among 100 Kan^r transductants 60 are **his⁻** and 40 are **his⁺**.

Give the distance between the **Tn5** insertion and **his1**, expressed as a cotransduction frequency.

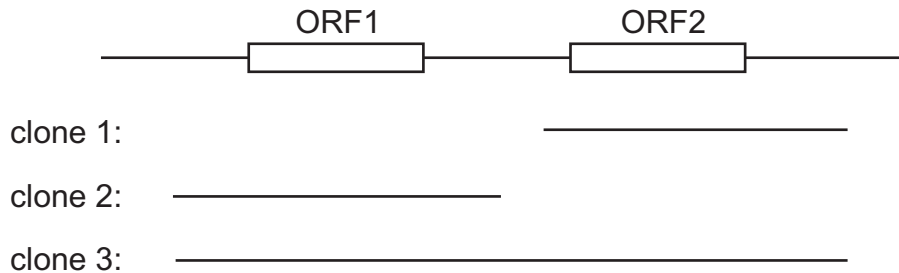
(c) You choose one of the **his⁺** revertants to test for the presence of a suppressor mutation. To do this you grow **P1** phage on one of the **his⁻** Kan^r transductant isolated in part **(b)** and use this phage to infect the **his⁺** revertant strain selecting for Kan^r transductants. Out of 100 Kan^r transductants, all are **his⁺**. What does this result tell you about the whether the **his⁺** revertant was a back mutation or an extragenic suppressor. Explain your logic.

(d) Next you choose a second **his⁺** revertant to test for the presence of a suppressor mutation. You use the same **P1** phage lysate generated in part **(c)** to infect the second **his⁺** revertant strain and select for Kan^r transductants. Out of 100 Kan^r transductants, 40 are **his⁻** and 60 are **his⁺**. What does this result tell you about the whether the **his⁺** revertant was a back mutation or an extragenic suppressor. Is this cotransduction frequency consistent with that obtained in part **(b)**? Explain.

(e) In this part we will use the linked **Tn5** insertion to determine the order of two different amber mutations in the **his1** gene (we will call these mutations **his1-1am** and **his1-2am**). To do this, you construct a strain that has both the **Tn5** insertion described above and **his1-1am**. **P1** is grown on this strain and used to infect the **his1-2am** strain. In this transduction experiment, 2 out of 1000 Kan^r transductants are **his⁺**. In the reciprocal cross, **P1** is grown on a **Tn5 his1-2am** strain and the resulting phage are used to infect a **his1-1am** strain. In this experiment, 12 out of 1000 Kan^r transductants are **his⁺**.

Draw a map showing the relative order of the **Tn5** insertion, **his1-1am** and **his1-2am**.

3. You are interested in introducing into *E. coli* the gene for a potential new restriction enzyme isolated from an obscure bacterial species. You have isolated a segment of a the bacterial genome that you suspect may contain the genes for a restriction enzyme and its corresponding modifying enzyme. Shown below is a map based on the DNA sequence of the cloned fragment showing two open reading frames labeled **ORF1** and **ORF2**.



(a) Shown on the map are three subclones that you have constructed by inserting a portion of the isolated fragment into a plasmid vector that contains an ampicillin resistance gene. Clone 1 contains **ORF2** only, clone 2 contains **ORF1** only, and clone 3 contains both **ORF1** and **ORF2**. You transform *E. coli* with each of the three clones selecting for ampicillin resistance. You obtain many amp^r transformants for clone 1, no amp^r transformants for clone 2, and some amp^r transformants for clone 3 (but not as many as for clone 1). Assuming that you are correct in that you are studying the genes of a restriction enzyme system, propose an explanation for the results of the subcloning experiments.

(b) Phage λ normally has a high efficiency of plating on *E. coli*, that is most of the phage particles will form plaques. However, unmodified phage λ will have a very low efficiency of plating on cells that express a restriction enzyme. For the following experiments predict whether a high or low efficiency of plating will be observed.

Phage λ grown on wild-type *E. coli* plated on *E. coli* carrying clone 1.

Phage λ grown on wild-type *E. coli* plated on *E. coli* carrying clone 3.

Phage λ grown on *E. coli* carrying clone 1 plated on *E. coli* carrying clone 1.

Phage λ grown on *E. coli* carrying clone 1 plated on *E. coli* carrying clone 3.

Phage λ grown on *E. coli* carrying clone 3 plated on *E. coli* carrying clone 1.

Phage λ grown on *E. coli* carrying clone 3 plated on *E. coli* carrying clone 3.

Solutions to 7.03 Problem Set 3

1.

a. In order to suppress TAG and TAA nonsense mutations, it requires two different tRNA's that have anti-codons mutated to CUA and UUA. Having two mutations in two different tRNA genes at the same time is as extremely unlikely as it is to get two back mutations at the same time.

b. The possible mutations are:

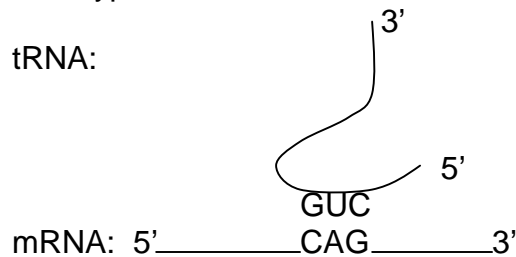
G	→	A
C	→	T

Amber mutation is TAG. tRNA has to have the anti-codon CUA. Gene that encodes this tRNA must have the sequence CTA. This sequence, given the possible mutations, could have mutated from:

CTG	→	CTA
CCA	→	CTA

Glutamine and Tryptophan tRNA's.

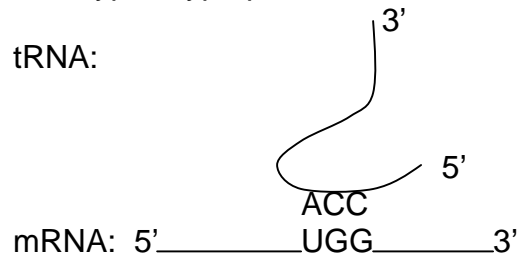
Wild-type Glutamine tRNA:



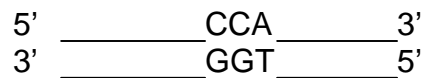
Coding sequence of DNA for the tRNA:

5'	_____	CTG	_____	3'
3'	_____	GAC	_____	5'

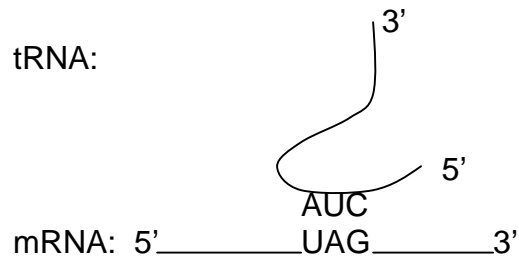
Wild-type Tryptophan:



Coding sequence of DNA for the tRNA:



Mutant:



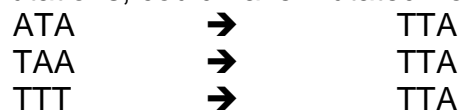
Coding sequence of DNA for the tRNA:



c. The possible mutations are:

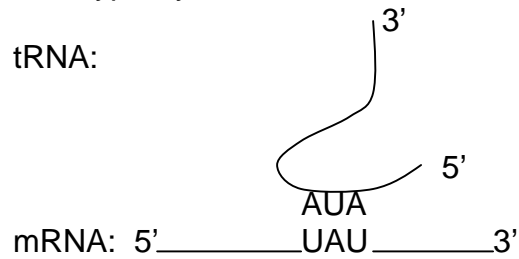


Ochre mutation is TAA. tRNA has to have the anti-codon UUA. Gene that encodes this tRNA must have the sequence TTA. This sequence, given the possible mutations, could have mutated from:

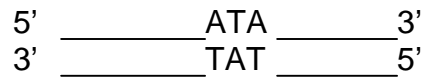


Tyrosine, Leucine, Lysine tRNA's.

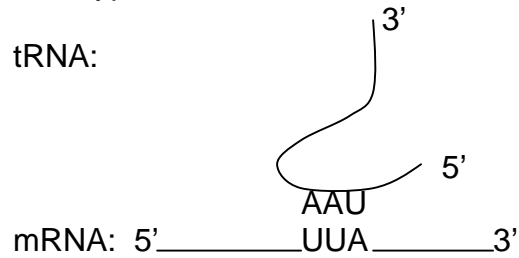
Wild-type Tyrosine:



Coding sequence of DNA for the tRNA:



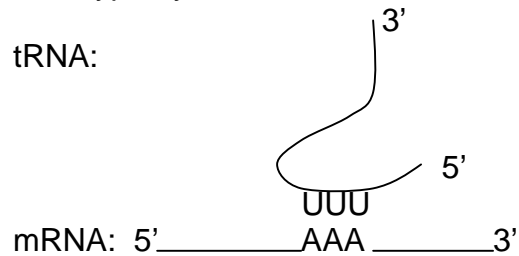
Wild-type Leucine:



Coding sequence of DNA for the tRNA:



Wild-type Lysine:

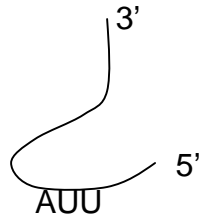


Coding sequence of DNA for the tRNA:



Mutant:

tRNA:



mRNA: 5' _____ UAA _____ 3'

Coding sequence of DNA for the tRNA:



2.

- a. In order to suppress the nonsense mutation with another mutation in the same gene but not in the nonsense codon, you would need a frameshift mutation before the nonsense codon and another frameshift mutation that puts the reading frame back in the original reading frame. This is a highly unlikely event, because EMS generally does not cause frameshift mutations.
- b. Because P1 was grown on **his**⁺ Kan^r strain, cotransductants selected for Kan^r are **his**⁺.

$$\text{cotransduction frequency} = \frac{\# \text{ of cotransductants}}{\text{total transductants}} = 40/100 = 40\%$$

- c. If the revertant were a back mutation, 40% of the transductants would have been **his**⁻, because **his**⁻ cotransduces with Kan^r 40% of the time. If the revertant were an extragenic suppressor, then all the transductants would be **his**⁺, because the extragenic suppressor would suppress the **his**⁻ mutation. Therefore, it must be an extragenic suppressor.
- d. The revertant must have been a back mutation. The cotransduction frequency is consistent. The only difference is that Kan^r is cotransducing with **his**⁻ allele instead of **his**⁺.

- e. Because his1-1am and his1-2am are both mutations in His1 gene, Tn5 can't be between these two mutations. Otherwise, His1 gene would be disrupted by Tn5.

	Experiment 1			Experiment 2			
P1 Grown on:	Tn5	1am	+	Tn5	+	2am	
P1 Infects:	+	+	2am	+	1am	+	
<u>Possible Transductant Genotypes</u>							<u>Count</u>
All cotransduce:	Tn5	1am	+	Tn5	+	2am	
Only Tn5:	Tn5	+	2am	Tn5	1am	+	
1am cotransduce:	Tn5	1am	2am	Tn5	+	+	12
2am cotransduce:	Tn5	+	+	Tn5	1am	2am	2

Since the locus of the mutations don't change between the experiments, each type of recombination should occur at the same frequency. In a quadruple crossover event, "All cotransduce" and "Only Tn5" classes are not possibilities. Therefore, Tn5 and 2am cotransductants must be the quadruple crossover class. We find that gene order is:

Tn5 his1-1am his1-2am

3.

- a. In cells that have restriction enzyme but no modifying enzyme, DNA is digested, therefore the cells would not survive.

In cells that have no restriction enzyme, DNA is not digested, therefore the cells can survive on Amp^r given that the plasmid transforms.

In cells that have both restriction enzyme and modifying enzyme, it's a race between these two enzymes recognizing the same sequence and either protecting or digesting it. Therefore, some of the cells with the plasmid will survive and some will not.

This follows that ORF1 encodes for the restriction enzyme and ORF2 encodes for the modifying enzyme.

- b. λ grown on wild-type *E. coli* plated on *E. coli* carrying clone 1.
DNA not modified, no restriction enzyme → high efficiency
- λ grown on wild-type *E. coli* plated on *E. coli* carrying clone 3.
DNA not modified, restriction enzyme → low efficiency
- λ grown on *E. coli* carrying clone 1 plated on *E. coli* carrying clone 1.
DNA modified, no restriction enzyme → high efficiency
- λ grown on *E. coli* carrying clone 1 plated on *E. coli* carrying clone 3.
DNA modified, restriction enzyme → high efficiency
- λ grown on *E. coli* carrying clone 3 plated on *E. coli* carrying clone 1.
DNA modified, no restriction enzyme → high efficiency
- λ grown on *E. coli* carrying clone 3 plated on *E. coli* carrying clone 3.
DNA modified, restriction enzyme → high efficiency

7.03 Problem Set 4

Due before 5 PM on Thursday, October 25

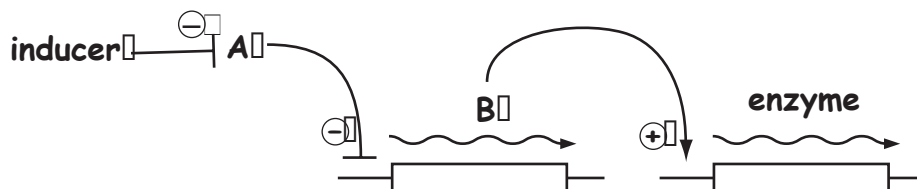
Hand in answers in recitation section or in the box outside the class

1. You are given an *E. coli* strain that you know contains an F' carrying the Lac genes, but you don't know precisely what Lac gene alleles are on the chromosome or on the F' .

(a) First you test the Lac phenotype of the strain and find that it expresses β -galactosidase constitutively. Next you set up a mating to an F^- strain that has a chromosomal deletion of the Lac operon and you find that the strains that receive the F' Lac express β -galactosidase in a normally regulated fashion. Given these observations, propose genotypes for both the Lac operon on the chromosome and the Lac operon carried on the F' in the strain that you were given.

(b) An F' carrying a segment of chromosomal DNA can occasionally recombine with the homologous chromosomal sequences to produce an Hfr. Starting with the F' strain that you were given, you isolate a number of derivatives that have become Hfrs. One way to test the genotype of these different Hfrs is to mate them to an F^- strain that has a $LacZ^-$ mutation on the chromosome and then to test for the properties of the Lac operon that are transferred at early times after mating. By performing this test you find that among the different Hfr derivatives that you have isolated some transfer a Lac operon that expresses β -galactosidase constitutively, some transfer a Lac operon that gives uninducible expression, while others transfer a Lac operon that shows normal regulation of β -galactosidase. On the basis of these results draw a map of the F' showing the direction of the origin of transfer relative to the arrangement of Lac genes. Also show where each recombination event must have taken place to produce each of the three different kinds of Hfrs. (Hint, to solve this tricky problem it will help to draw out both possible orientations and then determine the behavior of all possible recombination events between the F' and the chromosome.)

2. The following regulatory pathway, which depends on genes A and B, is responsible for turning on the expression of an enzyme in response to an inducer molecule.



Describe the minimum number of experiments and their outcomes that would have been needed to unambiguously deduce this model for the regulatory pathway.

3. You are studying the regulation of a new operon that is expressed in response to an inducer molecule. You have isolated a loss of function mutation in the gene for one of the enzyme products of the operon that you call E^- . In addition you isolate a regulatory mutation, which you call A^- that gives uninducible expression of enzyme E.

You construct a number of strains to determine the nature of the A^- mutation. Their genotypes and properties are given below:

<u>Genotype</u>	<u>Phenotype</u>
A^-E^+	uninducible
$A^-E^+ / F' A^+E^+$	uninducible
$A^-E^+ / F' A^+E^-$	uninducible
$A^-E^- / F' A^+E^+$	uninducible

(a) On the basis of these results, propose a function for the A gene product in the regulation of enzyme E.

Next, you isolate a second regulatory mutation which you call B^- which causes constitutive of enzyme E. The genotypes and phenotypes of a number of a number of strains carrying the B^- mutation are as follows:

<u>Genotype</u>	<u>Phenotype</u>
B^-E^+	constitutive
$B^-E^+ / F' B^+E^+$	regulated
$B^-E^+ / F' B^+E^-$	regulated
$B^-E^- / F' B^+E^+$	regulated

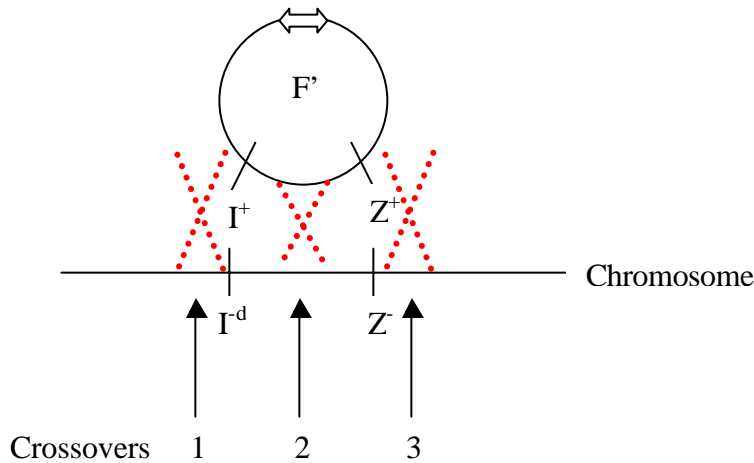
(b) On the basis of these results, propose a function for the B gene product in the regulation of enzyme E.

(c) Finally, you construct an $A^- B^-$ double mutant, which gives constitutive of enzyme E. Give two possible models for regulatory pathways that can explain these results. For each model, be as specific as you can about the normal function of the A and B gene products as well as the nature of the A^- and B^- mutations. Be sure to include in your models a role for the inducer molecule.

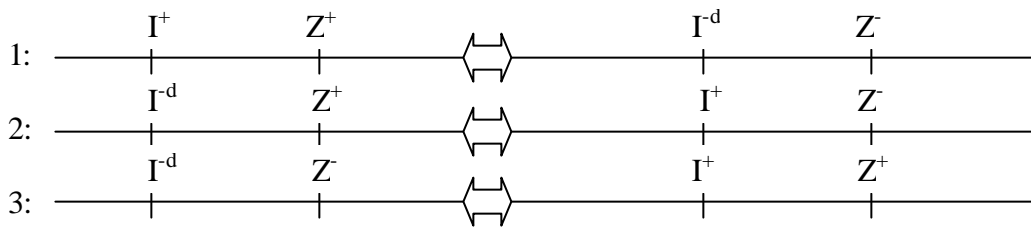
1.a) F' must be I⁺ P⁺ O⁺ Z⁺ (could possibly be I⁻ if transduced into lacZ⁻ strain since that strain may be lacI⁺ and would mask an I⁻ on the F') and the chromosome must be either I⁻ or P⁺ O^c Z⁺ (we do not know if either P, O or Z are mutated in the I⁻ nor do we know if the I is mutated in the case of an O^c mutation).

When the F' is transferred into a LacZ⁻ (or Lac deleted) strain we get normal regulation of the Lac Z. This suggests that the F' is LacI⁺ LacP⁺ LacO⁺ and LacZ⁺ (or in other words that it carries a complete normal lac operon). When this F' is in our strain the result is a constitutively active lac operon. This suggests that the chromosome must have either an I⁻ or an O^c since both of these would result in constitutive activity (when combined with the F' in the case of the I⁻). In the case of an I⁻ mutation we do not know if the chromosome is Z⁺ or Z⁻ but we know that for an O^c mutation to be constitutively active the chromosome would have to be Z⁺.

b) Here we can rule out the possibility that the chromosome has an O^c mutation because that will not allow us to get the three different Hfrs that we are seeing. We also now know that the chromosome must also have either a P⁻, O⁻ or Z mutation (this is required to get the three Hfrs). Since the F' is P⁺ O⁺ Z⁺ it does not matter which is mutated on the chromosome so for simplicity I will chose Z. Now we know that the chromosome was I⁻ Z⁻ and the F' was I⁺ and Z⁺ so we can now test the possible orders.

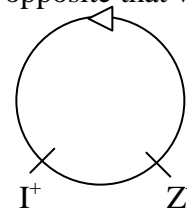


Now we can determine what the Hfr resulting from each of these crossovers would look like.



Now we can look at which alleles will be transferred early and determine the direction of the origin of replication on the F'.

From this we can see that if transfer begins left of the origin (↙ facing origin in Hfr) we would get transfer of normally induced operon from Hfr 1, a constitutively active operon from Hfr 2 and an uninducible operon from Hfr 3. If transfer goes to the right of the origin (↘ facing origin in Hfr) we would not get an operon being transferred early that is constitutively active. Therefore we know that the map of the F' must be (note that the orientation of the origin in the F' is opposite that when it is in the Hfr) and the recombination events are described above:



2. The minimum number of experiments needed to deduce this model is 6.

Exp 1: Look at the phenotype of the A^- mutant: Will be constitutively active.

Exp 2: Look at the phenotype of the B^- mutant: Will be uninducible.

Exp 3: Look at A^-B^- double mutant: Will be uninducible, since the double mutant has the same phenotype as the B^- we know that B is epistatic to A (or in other words the B gene acts after the A gene in the pathway).

Exp 4: Put an $A^+E^+ F'$ into an A^-E^+ cell: Will have normal regulation of the operon, this tells us that A^- is recessive and we can deduce that it is trans-acting (from chart in notes for lecture 18) and is a repressor.

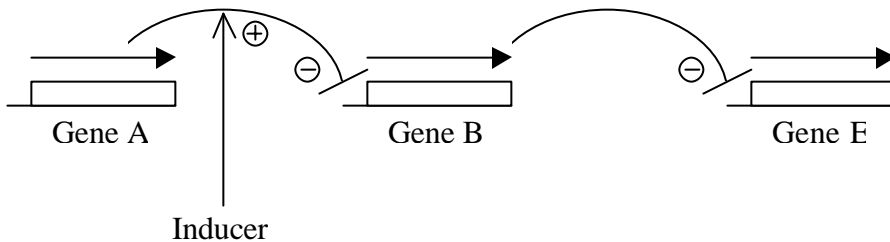
Exp 5: Put a $B^+E^+ F'$ into a B^-E^+ cell: Will have normal regulation. This tells us that B is recessive. Now if we look at the chart from lecture 18 we see that just knowing that it is uninducible and recessive does not narrow it down to one type of mutation so we need one more test.

Exp 6: Put $B^+E^+ F'$ into B^-E^+ cell: Will have normal regulation. Then put $B^-E^+ F'$ into B^+E^+ cell: Will also have normal regulation. This tells us that B is also trans-acting and thus according to the chart it is an activator.

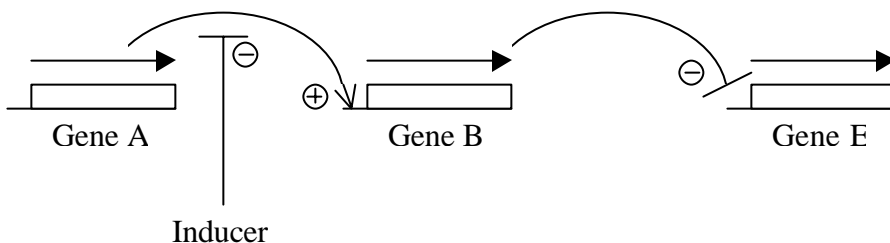
3.a) From the data given we can determine that A^- is dominant and trans-acting and we are told that it is uninducible. Therefore, we can look at the chart from lecture 18 and determine that it is either a repressor^s or activator^{-d}. Thus A is either a repressor or an activator.

b) From the data given we can determine that B- is constitutive, recessive and trans-acting. Therefore, according to the chart it is a repressor^r. Thus B is a repressor.

c) Model 1: A is activator^{-d} (prevents binding of Gene B operator) and so constitutively activates Gene B.



Model 2: A is repressor^s and so it cannot bind the inducer and is always activating B.



In both models B is a repressor and B^- results in constitutive activation of Gene E.

7.03 Problem Set 5

Due before 5 PM on Thursday, November 15, 2001

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

1. You are studying the regulation of the yeast gene for glutamine synthesis (Gln1), which is *not* expressed when there is glutamine in the growth medium and is expressed when glutamine is absent. To begin your analysis of regulation you fuse the promoter region of the enzyme to the LacZ coding sequence and then place this hybrid gene on a yeast plasmid. You find that yeast cells carrying the hybrid gene only expresses β -galactosidase activity when glutamine is absent, meaning that the promoter region you have selected contains all of the necessary cis-acting sequences for normal regulation. You next identify two different mutants that show abnormal regulation of your reporter. Gln2⁻ gives constitutive expression whereas Gln3⁻ shows uninducible expression.

(a) Is Gln2⁻ cis-acting or trans-acting? Is Gln3⁻ cis-acting or trans-acting? Explain.

(b) You cross a Gln2⁻ to wild type and the resulting diploid shows normal regulation of the Gln1 promoter fused to LacZ. Is Gln2⁻ dominant or recessive? Explain.

(c) You cross a Gln3⁻ to wild type and the resulting diploid shows normal regulation of the Gln1 promoter fused to LacZ. Is Gln3⁻ dominant or recessive? Explain.

Next, you cross a Gln2⁻ to a Gln3⁻ mutant. Three different tetrad types are obtained:

<u>Type 1</u>	<u>Type 2</u>	<u>Type3</u>
regulated (wt)	regulated(wt)	constitutive
regulated (wt)	constitutive	constitutive
uninducible	uninducible	uninducible
uninducible	uninducible	uninducible

Tetrad Type 2 is the most abundant and Type 1 and Type 3 occur at roughly equal frequencies.

(d) Are the Gln2⁻ and Gln3⁻ mutations linked? What is the phenotype of a Gln2⁻ Gln3⁻ double mutant?

(e) Produce a model to explain the regulation of the enzyme that is consistent with all of the data you have.

Next you evaluate the promoter sequences necessary for expression of the gene. The figure below shows the effect of different 50 bp deletions in the promoter region on the amount of β -galactosidase activity expressed by the reporter gene fused to the Gln1 promoter.

	-300	-250	-200	-150	-100	-50	+1		<u>β-galactosidase</u>	
									+glutamine	-glutamine
wt	—————LacZ								10 units	100 units
1	———		—————LacZ						10 units	100 units
2	—————			—————LacZ					10 units	100 units
3	—————				—————LacZ				10 units	10 units
4	—————					—————LacZ			10 units	100 units
5	—————						—————LacZ		10 units	100 units
6	—————						—————LacZ		0 units	0 units

(f) Deletions 3 and 6 behave similarly in that they both prevent induction of the reporter to high levels. Based on their positions it appears that deletion 6 may remove the TATA sequence, whereas deletion 3 probably removes a UAS for a transcriptional activator. Describe an experiment or series of experiments that would enable you to test this hypothesis and distinguish the cis-acting sequences removed by deletion 3 and 6.

2. Mammalian genes, their mRNAs, and the proteins that they encode vary greatly in size, but an understanding of typical or averages sizes is useful. Knowing what you do about the estimated percentage of the human genome that codes for protein, and the estimated number of genes in the human genome, calculate:

- The total amount of coding sequence in the human genome.
- The average length (in base pairs) of the coding sequence in a human gene.
- The average length (in amino acid residues) of a human protein.

You set out to study the molecular basis of hearing (and its absence: deafness) in mice and humans. Through genetic linkage analysis in mice, you map an autosomal recessive deafness mutation called *Ard* to a 0.5 Mb region of mouse chromosome 14. The responsible gene has not yet been defined at a molecular level. The DNA

sequence of this 0.5-Mb region has not yet been determined, but you have available 7 BACs (order unknown) and 11 STSs (order unknown) that derive from the region. By PCR, you test each of the 7 BACs for the presence (+) or absence (-) of each of the 11 STSs, and you obtain the following results:

	BACs						
	1	2	3	4	5	6	7
STS1	-	+	-	+	-	-	-
STS2	-	-	-	-	-	+	+
STS3	-	-	-	+	-	-	-
STS4	+	-	+	-	-	+	-
STS5	-	+	-	-	-	-	-
STS6	-	-	+	-	+	-	-
STS7	+	-	-	-	-	+	+
STS8	-	-	-	+	+	-	-
STS9	-	-	-	-	-	-	+
STS10	-	-	-	-	-	-	+
STS11	+	-	+	-	-	-	-

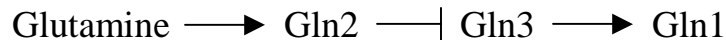
- (d) Using this STS content data, construct a physical map of the 0.5-Mb region in which you indicate the order of the BACs and the overlaps among them, and the order of the STSs. Indicate any pairs of STSs that cannot be ordered based on this information.
- (e) Like STSs, mammalian exons are typically about 100 to 200 base-pairs in length. You discover that STS6 contains the translation initiation codon (ATG) of a gene, optimistically called *Pindrop*, that is expressed in the ear and that is likely to contain multiple exons scattered across approximately 50 kb. You have available the complete sequence of the *Pindrop* mRNA. Describe an experiment (without sequencing any of the BACs) that would allow you to determine the approximate location of *Pindrop*'s stop codon on your physical map of overlapping BACs and ordered STSs.
- (f) For the experiment that you proposed in (e), list a set of experimental results that are consistent with the presence of *Pindrop*'s stop codon in three BACs, including BAC3. Based on your new data, redraw your physical map, including the location

of *Pindrop*'s translation initiation and stop codons and the direction of *Pindrop* transcription (5'→3').

- 3.** You hypothesize that loss of *Pindrop* function is the cause of deafness in the *Ard* strain of mice. You decide to test this hypothesis using a variety of transgenic methods. You have available the set of mouse BACs described in Problem 2, as well as a human genomic BAC that contains the entirety of the human *Pindrop* gene. (You find that the human genome contains a *Pindrop* gene whose DNA sequence is very similar but not identical to that of the mouse *Pindrop* gene.)
- (a) What modification to the mouse genome would allow you to generate a mouse with only one copy of the *Pindrop* gene? Draw the DNA construct that would allow you to modify the mouse genome, and explain how your construct would integrate into the mouse genome.
- (b) What if any additional steps would be required to test the hypothesis that *Pindrop* function is required for hearing in mice (regardless of whether *Pindrop* and *Ard* are the same gene)?
- (c) What if any additional experiments would be required to test the hypothesis that *Pindrop* (a molecularly defined gene) and *Ard* (a gene defined by phenotype) are the same gene?
- (d) What additional experiments would allow you decide to test the hypothesis that the human and mouse *Pindrop* genes are functionally interchangeable? Draw any needed DNA constructs and explain how they would integrate into the mouse genome.

Problem 1

- a) $Gln2^-$ and $Gln3^-$ are both trans-acting because all of the cis-acting elements are contained on the plasmid.
- b) $Gln2^-$ is recessive because the $Gln2^-/Gln2^+$ heterozygote diploid shows the wild type phenotype, normal regulation.
- c) $Gln3^-$ is recessive because the $Gln3^-/Gln3^+$ heterozygote diploid shows the wild type phenotype, normal regulation.
- d) Type 1 is NPD, Type 2 is T, and Type 3 is PD. $Gln2^-$ and $Gln3^-$ mutations are unlinked because $PD \approx NPD$ and $T \gg PD$, NPD. By looking at the NPD, we determine that the double mutant phenotype is uninducible.
- e) The double mutant tells us that $Gln3^-$ is epistatic to $Gln2^-$. So a possible model is:



f) One possible set of experiments would be to replace the deleted sequence in deletion 3 and in deletion 6 by a different activating sequence, i.e. the LexA operator sequence. Then test if there is transcription in the presence of the activator, LexA. If there is transcription, then the original deletion was of the UAS. If there is no transcription, then the original deletion was of the TATA sequence. The following shows what you would expect for β -galactosidase activity if deletion 3 was the UAS and deletion 6 was the TATA sequence.

		β -galactosidase	
		-LexA	+LexA
3	————— LexA OP ————— LacZ	10	100
6	————— LexA OP ————— LacZ	0	0

Problem 2

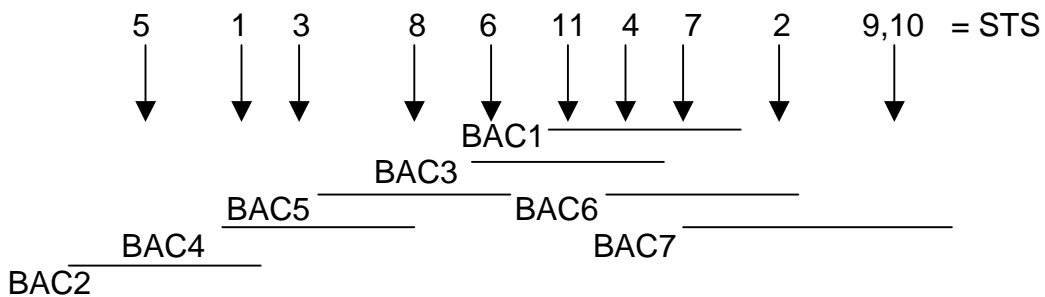
a) A rough estimation of the coding sequence of the human genome is 1.5%. There is roughly 1-1.5% regulatory and 1-1.5% coding. In base pairs:

$$3 \times 10^9 \text{ bp} * .015 = 4.5 \times 10^7 \text{ bp}$$

b) $4.5 \times 10^7 \text{ bp} / 30,000 \text{ genes} = 1500 \text{ bp/gene}$

c) $1500 \text{ bp} / 3 \text{ bp per amino acid} = 500 \text{ amino acids}$ or in molecular weight about 50 kiloDaltons

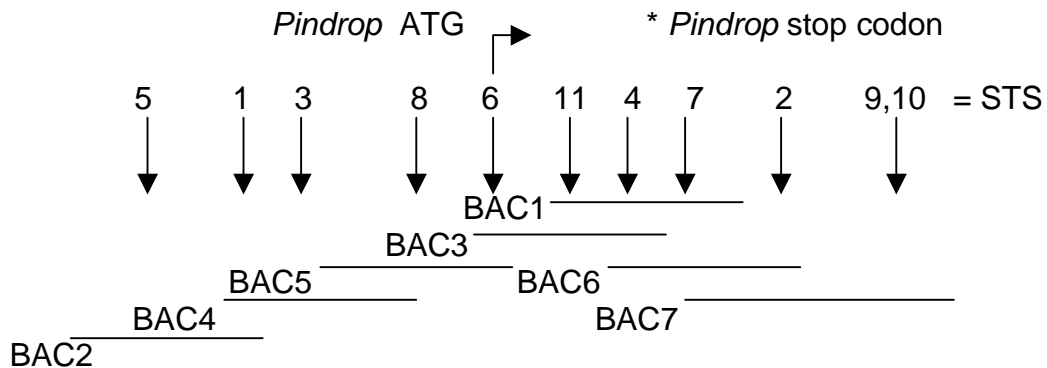
d)



The STSs are demarcated by arrows. The BAC numbers are written underneath and to the left of the line that they represent.

e) Much like the detection of STSs by PCR, we can use PCR to detect a sequence from the *Pindrop* mRNA on certain BACs. Design PCR primers to the 3' end of the mRNA sequence that includes the stop codon. This sequence will represent the last exon of the coding sequence. Then, collect DNA from each of our seven BACs and perform PCR on each with the primers to the last exon. Whichever BACs contain the stop codon will give a PCR product.

f)



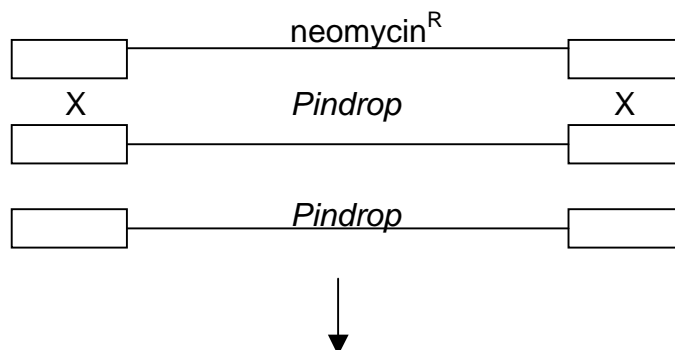
Our PCR based approach to find the translational stop codon would give the following results:

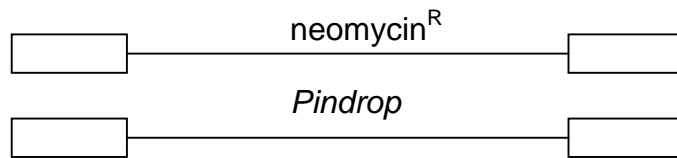
BAC1	Positive	BAC5	Negative
BAC2	Negative	BAC6	Positive
BAC3	Positive	BAC7	Negative
BAC4	Negative		

Since the PCR product is seen in three different BACs including BAC3, we know that transcription must go to the right. The translational stop must be somewhere between STS 4 and STS 7, but present on BACs 1, 3, and 6.

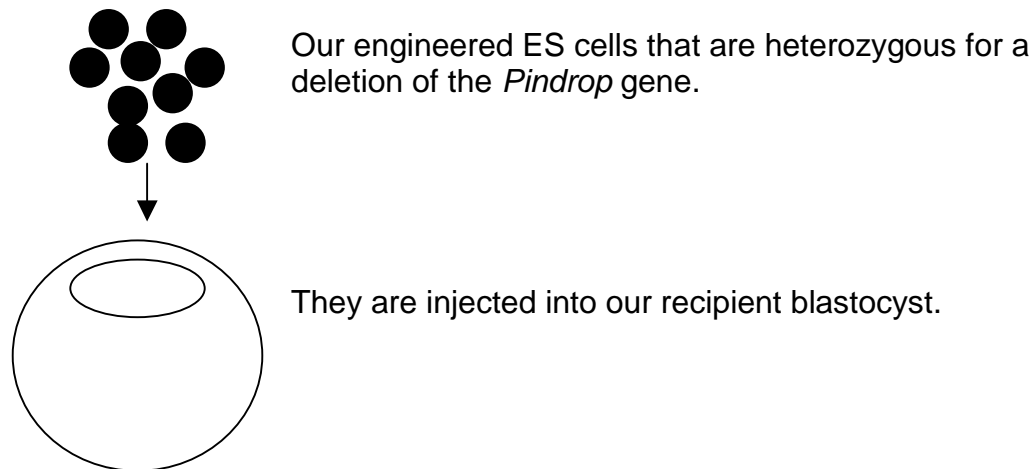
Problem 3

a) We would need to generate a knockout heterozygote to create a mouse with only one copy of the *Pindrop* gene. A DNA construct would be made that has homologous sequences flanking a neomycin resistance cassette (the homologous sequences are shown as boxes). Electroporation would be used to get the construct into an ES cell (derived from a black mouse line) and then homologous recombination (shown as an X) would replace one chromosome's *Pindrop* gene with the neomycin resistant cassette (thus deleting one copy of the gene).





Then, we would select for those ES cells with the neomycin resistant cassette. Those ES cells would be put into a blastocyst derived from a different visible strain (such as white coat color), and then that blastocyst would be put into a pseudopregnant female mouse. Once born that mouse will be a chimera derived from our engineered ES cells (black) and the host blastocyst (white).

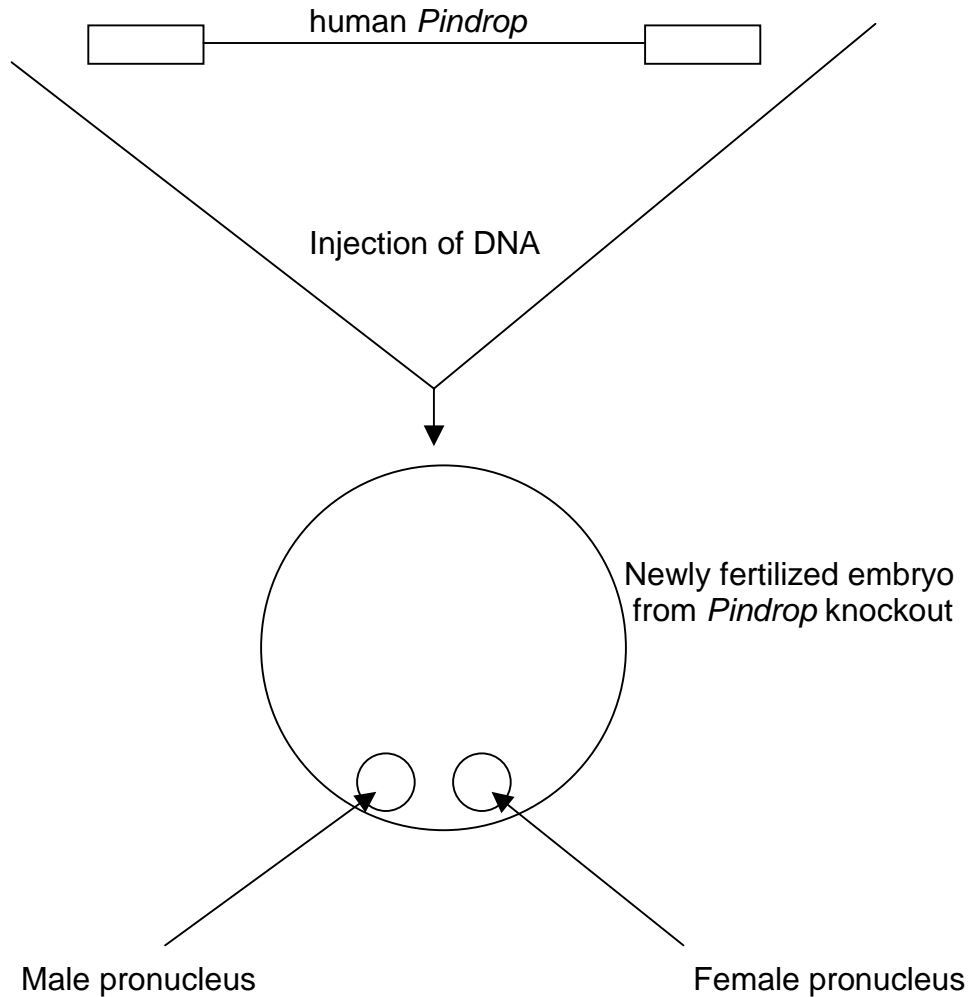


The chimeric mouse would then be mated to a white wild-type mouse to see if the engineered ES cells contributed to the germline (and thus will be passed on to the next generation). If black offspring are observed, then our ES cells contributed to the germline and half of all the black mice will be heterozygotes for the *Pindrop* gene (which we can test by PCR).

b) We can take two heterozygotes from part a and cross them to each other. One quarter of all their offspring will be homozygous for the deletion of the *Pindrop* gene. If *Pindrop* is required for hearing in mice, then those homozygous knockout mice will be hearing impaired.

c) We could do a complementation test between our *Pindrop* knockout and the *Ard* mutation. If the *Pindrop/Ard* trans-heterozygote is hearing impaired, then there is a strong possibility that they are mutations in the same gene.

d) We could use pronuclear injection of the human *Pindrop* gene to see if the human copy can rescue the deafness of the mouse *Pindrop* knockout. If the transgene array with human *Pindrop* can function in place of the mouse *Pindrop* gene, then the mouse will not be deaf.



7.03 Problem Set 6

Due before 5 PM on Tuesday, November 27, 2001

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

1. Consider two large but completely isolated populations of rabbits: population X (consisting of 100,000 randomly mating rabbits) and population Y (consisting of 50,000 randomly mating rabbits). In both populations there are two alleles for tail color: the B allele (associated with brown tails), and the b allele (associated with white tails). Brown tail color is dominant to white tail color. In population X, 1000 rabbits have white tails. In population Y, 5000 rabbits have white tails. Assume Hardy-Weinberg equilibrium. Estimate the following (showing your calculations):

(a) The frequency of the b allele in population X. The frequency of the b allele in population Y.

(b) The number of heterozygous rabbits in population X. The number of heterozygous rabbits in population Y.

(c) The probability that a rabbit will have a white tail if its mother is a randomly selected white-tail rabbit from population X and its father is a randomly selected brown-tail rabbit from population Y.

(d) The probability that a rabbit will have a brown tail if its mother is a randomly selected brown-tail rabbit from population X and its father is a randomly selected brown-tail rabbit from population Y.

Populations 1 and 2 have been separated for many years by a deep, raging river. Suppose that the river is now dammed upstream of the two populations, allowing for easy crossing of the riverbed and random mating between the two formerly isolated populations.

(e) Once Hardy-Weinberg equilibrium is established in the new, joint population (call it population XY), what will be the frequency of the b allele in population XY? Show your calculations.

(f) Once Hardy-Weinberg equilibrium is established in population XY, what fraction of the rabbits will have brown tails? Show your calculations.

(g) After the river is dammed, how many generations of random mating will be required for genotype frequencies to match Hardy-Weinberg expectations?

2. Cystic fibrosis is an autosomal recessive disease that currently affects about 1 in 1600 children in Europe. It has been hypothesized that heterozygotes for the cystic fibrosis allele may be less susceptible to some infectious disease (possibly plague) that was common in Europe hundreds of years ago, but that is seen rarely if ever today. Assume that today, and in the past, individuals with cystic fibrosis disease have about 5% as many offspring as average individuals in European populations. Assume that mating is random, and that the mutation rate is negligible.

- (a)** In Europe today, what is the frequency of the allele (call it CF) associated with cystic fibrosis?
- (b)** If the hypothesis stated above is correct, is the heterozygote advantage today higher, lower, or the same as it was hundreds of years ago? Explain your answer.
- (c)** If the hypothesis stated above is correct, is the frequency of CF heterozygotes today likely to be higher, lower, or the same as it was hundreds of years ago? Explain your answer.

Assume that the infectious disease in question disappeared abruptly and completely 400 years ago. Assume a human generation time of 20 years. Roughly estimate the following (making sure to state any additional simplifying assumptions that you employ, and showing your calculations):

- (d)** The frequency of the CF allele in Europe 400 years ago.
- (e)** The frequency of cystic fibrosis disease in Europe 400 years ago.
- (f)** The magnitude of the heterozygote advantage in Europe 400 years ago.

- 3.** “Double first cousins” are the result of either two brothers marrying two sisters, or of a brother/sister pair marrying another brother/sister pair. In some human populations where first-cousin marriages are common, double-first-cousin marriages also occur at a significant rate. Consider two large but completely isolated human populations (populations 1 and 2). In population 1, mating is random. In population 2, 15% of matings are between (conventional) first cousins and 5% of matings are between double first cousins. Assume that all other matings in population 2 are random. Consider two autosomal recessive diseases (diseases S and T). The incidence of disease S is the same (1 in 10,000 people) in both populations. In the case of disease T, the disease allele frequency is the same (0.002) in both populations. Assume that mutation rates and selection are negligible.
- (a) Calculate the inbreeding coefficient for matings between double first cousins.
- (b) What is the incidence of disease T in population 1? In population 2? Show your calculations.
- (c) What is the frequency of the allele associated with disease S in population 1? In population 2? Show your calculations.
- (d) In population 2, what fraction of individuals with disease S are products of either first-cousin or double-first-cousin matings? In population 2, what fraction of individuals with disease T are products of either first-cousin or double-first-cousin matings?

1. a) Let $f(b) = q$
 Population X: $q^2 = \frac{1000}{100000}$ $q = 0.1$

We can also determine the statistics for the rest of the population:

	$f(B) = p = 0.9$	$f(b) = q = 0.1$
$f(B) = p = 0.9$	BB $p^2 = 0.81$	Bb $pq = 0.09$
$f(b) = q = 0.1$	Bb $pq = 0.09$	bb $q^2 = 0.01$

$f(BB) = 0.81$; $f(Bb) = 0.18$; $f(bb) = 0.01$

Population Y: $q^2 = \frac{5000}{50000}$ $q = 0.316$

	$f(B) = p = 0.684$	$f(b) = q = 0.316$
$f(B) = p = 0.684$	BB $p^2 = 0.468$	Bb $pq = 0.216$
$f(b) = q = 0.316$	Bb $pq = 0.216$	bb $q^2 = 0.1$

$f(BB) = 0.468$; $f(Bb) = 0.432$; $f(bb) = 0.1$

- b) Heterozygous rabbits = $2pq \times$ population size
 Population X: $2(.9)(.1)(100000) = 18,000$ heterozygous rabbits
 Population Y: $2(.684)(.316)(50000) = 21,600$ heterozygous rabbits

c) In order to get a white-tailed offspring, the cross has to be Bb (father from Y) x bb (mother)
 $p(bb \text{ offspring}) = p(\text{brown-tailed rabbit in pop Y is Bb}) \times p(\text{having bb offspring})$
 $= \frac{.44}{.9} \times .5 = 0.24$

d) The probability that two rabbits will have brown-tailed offspring is the same as
 $1 - p(\text{white-tailed offspring})$
 $1 - p(\text{brown-tailed rabbit in pop X is Bb}) \times p(\text{brown-tailed rabbit in pop Y is Bb}) \times p(\text{bb offspring})$
 $1 - \frac{.18}{.99} \times \frac{.44}{.9} \times .25 = 0.98$

e) Since there is no force acting to change allele frequency (like mutation or selection), the allele frequency will remain equal to what it is when the two populations are combined.

$$\text{Total number of b alleles in XY} = 200000(0.1) + 100000(0.316)$$

$$\text{Divide this by the total number alleles in XY (300000)} \cdot \mathbf{0.17} = q_{XY}$$

f) Number of rabbits with brown tails = $1 - q^2 = 1 - (.17)^2 = \mathbf{0.97}$

g) Since mating is random, it takes only **one generation** for genotype frequencies to adjust to expected Hardy-Weinberg equilibrium states.

2. a) Let $q = f(\text{CF})$ $q^2 = \frac{1}{1600}$ $q = \frac{1}{40} = \mathbf{.025}$

b) The heterozygote advantage today is **lower** than it was hundreds of years ago. The reason for the advantage was resistance to an infectious disease that is no longer present today. Therefore, there is no selective force giving a heterozygote advantage today.

c) Since there is no longer any (or perhaps very little) heterozygote advantage, the main force acting on CF allele frequency is selection against homozygotes. This will lower the CF allele frequency. Since CF is rare, the heterozygote frequency is approximately $2q$, so the heterozygote frequency will be **lowered** as $f(\text{CF})$ decreases.

d) 400 years ago, there was a case of balanced polymorphism. The CF allele was rare, and mutation was negligible, so $\Delta q = hq - Sq^2 = 0$. Then, the disease disappeared, so we can say that h became 0. Consequently, q began to change at a rate of $-Sq^2$ per generation. $S = 0.95$ in this case, since CF individuals have only 5% as many offspring as normal individuals (fitness = 5%).

Going from one generation to the next: $q' = q - Sq^2$

We need to go backwards from the present, so we know q' and S and need to solve for q :

$$\begin{aligned} .025 &= q - .95q^2 \\ .95q^2 - q + .025 &= 0 \\ q &= \frac{1 \pm \sqrt{1 - 4(.95)(.025)}}{2(.95)} \quad \cdot \mathbf{.0256} \end{aligned}$$

This is what $f(\text{CF})$ was one generation, or 20 years, ago. To get an exact answer for $f(\text{CF})$ 400 years ago, do this calculation 20 times - once for each generation:

Years	$f(\text{CF})$
0 (present)	0.025
-20	0.0256
-40	0.0263
-60	0.0270
-80	0.0277
-100	0.0285
-120	0.0293
-140	0.0301
-160	0.0311
-180	0.0320

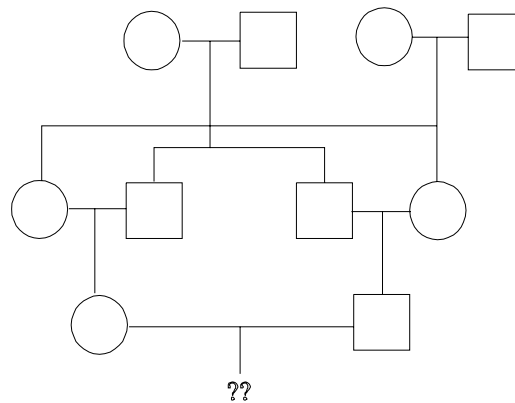
-200	0.0331
-220	0.0342
-240	0.0354
-260	0.0367
-280	0.0380
-300	0.0395
-320	0.0411
-340	0.0429
-360	0.0448
-380	0.0469
-400	0.0491

Therefore, $f(\text{CF})$ 400 years ago was approximately **0.0491**.

e) $f(\text{CF}/\text{CF}) = q^2 = \mathbf{0.0024}$

f) The CF allele is rare, and mutation is negligible, so we can approximate h :
 $h = S q = (.95)(.0491) = \mathbf{.047}$

3. a) Just like in first-cousin matings, the ?? offspring has a 1/16 probability of being heterozygous by descent at a given locus from one set of great-grandparents. However, there are two such sets of great-grandparents, so the inbreeding coefficient is **1/8**. Another way of looking at this is the ?? offspring has a 1/64 chance of being homozygous for one allele from a great grandparent. There are 8 such alleles at a given locus, so the inbreeding coefficient is $8 \times 1/64 = \mathbf{1/8}$.



b) Population 1 is in Hardy-Weinberg equilibrium, while population 2 has non-random mating.

$$\text{Population 1: Disease incidence} = q^2 = (.002)^2 = \mathbf{4 \times 10^{-6}}$$

Population 2: Disease incidence = sum of probabilities from each type of mating

$$p(\text{disease from given mating}) = \text{inbreeding coefficient} \times q \times p(\text{that type of mating})$$

$$p(\text{disease from first-cousin mating}) = (1/16)(.002)(.15) = 1.875 \times 10^{-5}$$

$$p(\text{disease from double f-c mating}) = (1/8)(.002)(.05) = 1.25 \times 10^{-5}$$

$$p(\text{disease from random mating}) = (.002)^2(.80) = 3.2 \times 10^{-6}$$

$$\text{Disease incidence} = 1.875 \times 10^{-5} + 1.25 \times 10^{-5} + 3.2 \times 10^{-6} = \mathbf{3.445 \times 10^{-5}}$$

c) Population 1: $q^2 = .0001$ $q = \mathbf{0.01}$

Population 2: Same calculations as above, except we now know disease incidence instead of allele frequency.

$$(1/16)(.15)q + (1/8)(.05)q + .8q^2 = .0001$$

$$.8q^2 + .015625q - .0001 = 0$$

$$q = \frac{-.015625 \pm \sqrt{(.015625)^2 - 4(.8)(.0001)}}{2(.8)} = \mathbf{5.08 \times 10^{-3}}$$

d) (Disease incidence due to first-cousin or double first-cousin mating)/(total disease incidence)

$$\text{Disease S: } \frac{(1/16)(.15)(5.08 \times 10^{-3}) + (1/8)(.05)(5.08 \times 10^{-3})}{.0001} = \frac{7.94 \times 10^{-5}}{.0001} = \mathbf{0.79}$$

$$\text{Disease T: } \frac{(1/16)(.15)(.002) + (1/8)(.05)(.002)}{3.445 \times 10^{-5}} = \frac{3.125 \times 10^{-5}}{3.445 \times 10^{-5}} = \mathbf{0.91}$$

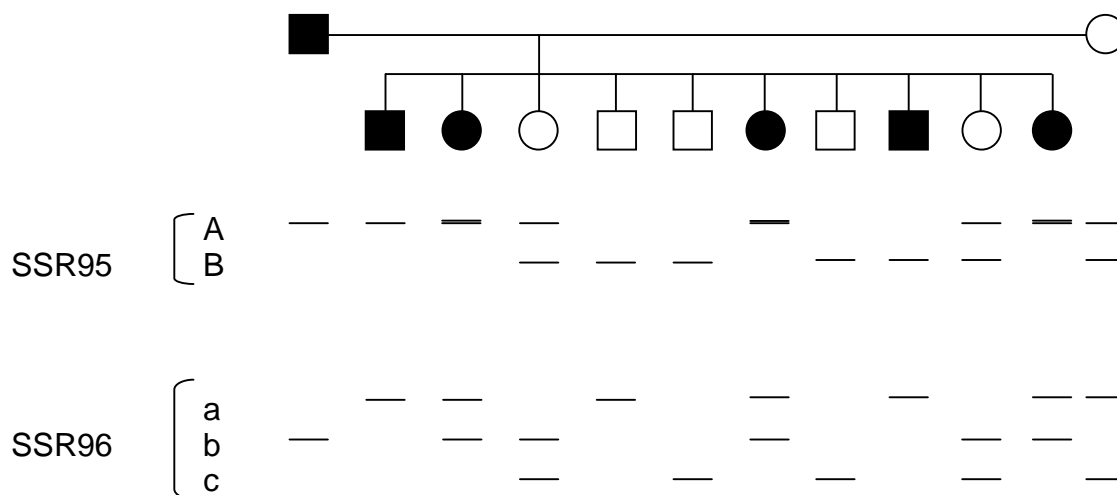
7.03 Problem Set 7

Only problems 1 and 2 will be graded. Your answers to problems 1 and 2 are due before **4 PM** on **Friday**, December 7, 2001. Hand in answers to problems 1 and 2 in recitation section or in the box outside the class. Problem 3 is for practice only. Do not hand in your answer to Problem 3.

1. You set out to genetically map color blindness with respect to SSR markers. Color blindness shows X-linked recessive inheritance and therefore is usually found in males. However, the mutant allele frequency is sufficiently high that colorblind females do occur.

Alleles: + (normal) cb (associated with color blindness)

Here is a family in which some individuals are affected:



- (a) Diagram the two possible phase relationships between the SSR95 and SSR96 alleles in the mother.
- (b) Calculate the LOD score for linkage at $\theta = 0.1$ between SSR95 and SSR96 in this family.
- (c) Identify a value of θ at which this family will yield a higher LOD score for linkage between SSR95 and SSR96. Calculate the LOD score for linkage between SSR95 and SSR96 at that new θ value.
- (d) Diagram the two possible phase relationships between the SSR95 and color blindness alleles in the mother.

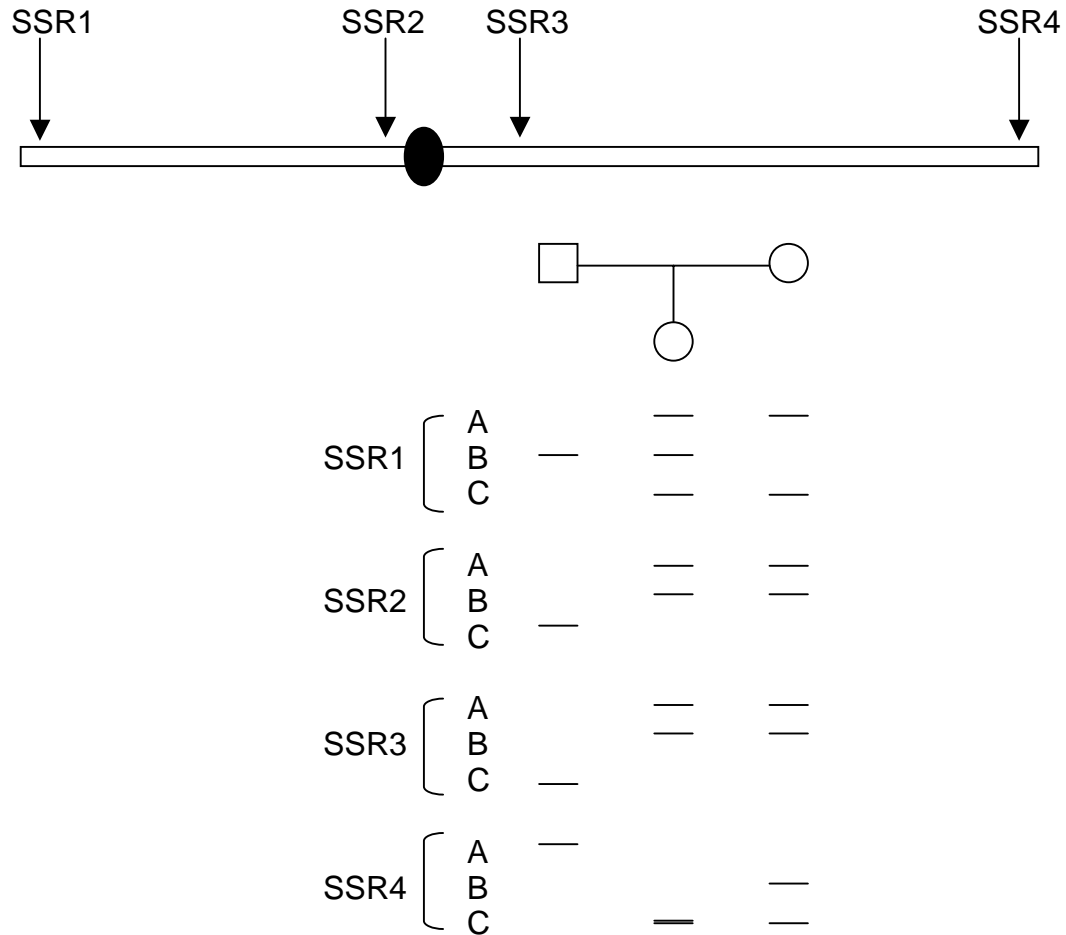
(e) Calculate a LOD score for linkage at $\theta = 0.1$ between SSR95 and color blindness in this family.

2. Congenital pyloric stenosis (an obstruction to the stomach's outlet to the small intestine) has a population incidence of 0.5% in newborn boys and of 0.1% in newborn girls. When a disease is more common in newborn boys than in newborn girls, two possibilities come to mind: 1) X-linked recessive inheritance, 2) disease susceptibility modified by sex hormones (e.g. increased by high levels of testosterone).

- (a) Which of these 2 possibilities are consistent with the observed population incidence in newborn males and that in females? Explain your answer.
- (b) Can you be confident, based on this data alone, that there is a genetic component to the risk of congenital pyloric stenosis? Explain your answer.
- (c) The concordance rate in MZ twins is 22% while that in DZ twins is 2%. Are you now confident that this is a genetic component, and, if so, how many genes are involved? Is there an environmental component?
- (d) 5.5% of the sons and 2.4% of the daughters of males with congenital pyloric stenosis are affected. By contrast, 19.4% of the sons and 7.3% of the daughters of females with congenital pyloric stenosis are affected. Why might the offspring of affected females be at higher risk than the offspring of affected males?

3. While working as a medical geneticist, you encounter an unusual patient: a 47,XXY girl.

You prepare DNA samples from the girl and from her parents. You confirm that the stated father is in fact the biological father by testing the family for a large number of autosomal SSRs. You also test the family 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) What might account for this girl having developed as a female despite the presence of a Y chromosome? Explain how you would test your hypothesis.

(e 5 pts.) How would you account for the presence in the XXY girl of a paternal allele for SSR1?

7.03 PS #7 Solutions

Fall 2001

$$1. \text{ a) } \frac{A \quad a}{B \quad c} \quad \frac{A \quad c}{B \quad a}$$

These are the alleles that mom passed to her kids.

SSR95	A	A	B	B	B	A	B	B	B	A
SSR96	a	a	c	a	c	a	c	a	c	a
cb	cb	cb	+	+	+	cb	+	cb	+	cb

- b) As we can see from the table above, there was recombination in 2 of the children between SSR95 and SSR96.

$$\text{LOD}_{0.1} = \log_{10} \frac{\frac{1}{2}(0.45)^8(0.05)^2 + \frac{1}{2}(0.05)^8(0.45)^2}{(0.25)^{10}} = \mathbf{0.343}$$

- c) Recombination in 2 out of 10 children suggests a recombination fraction of 1/5. $\theta = 0.2$

$$\text{LOD}_{0.2} = \log_{10} \frac{\frac{1}{2}(0.4)^8(0.1)^2 + \frac{1}{2}(0.1)^8(0.4)^2}{(0.25)^{10}} = \mathbf{0.536}$$

$$\text{d) } \frac{A \quad cb}{B \quad +} \quad \frac{A \quad +}{B \quad cb}$$

- e) Recombination between SSR95 and color blindness only happened in one of the children.

$$\text{LOD}_{0.1} = \log_{10} \frac{\frac{1}{2}(0.45)^9(0.05) + \frac{1}{2}(0.05)^9(0.45)}{(0.25)^{10}} = \mathbf{1.297}$$

2. a) **Disease susceptibility modified by sex hormones** is the more likely explanation. This is because an X-linked recessive trait would show square the incidence in females than in males. If frequency in boys = $q = 0.005$, then frequency in girls would be $q^2 = 0.000025$, which is considerably less than what is observed.

- b) **No**, this data is not enough to make any sort of conclusion about genetic components. We need more information about how being related to an affected individual affects a person's probability of being affected themselves. Studies such as twin studies would give us this information

- c) **Yes, there is a genetic component.** We know this because $MZ > DZ$. The ratio of the two is 11:1, so this indicates **multiple genes**.
Yes, there is an environmental component. We know this because $MZ < 100\%$.

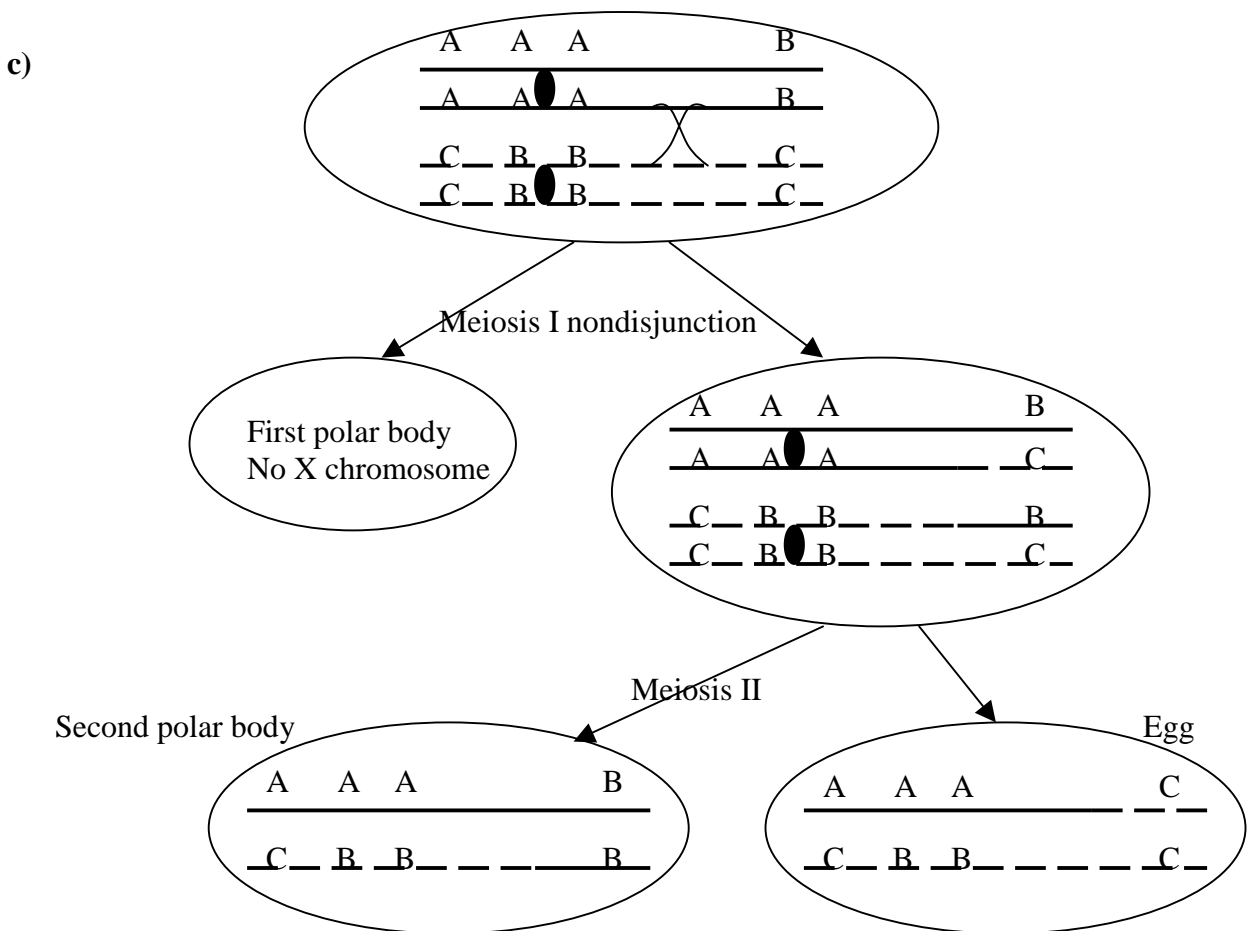
- d) The most common explanation is that these results are indicative of multiple genes and some disease susceptibility modified by sex hormones. Since males are automatically more susceptible to this condition, then they would not need as many mutant genes as

women would to cause the disease. Thus, an affected mother will have more genes to pass on to her offspring than an affected father.

A second possibility that the development of the fetus is affected by the environment in which it develops, namely the womb. If the mother is affected, there may be a greater chance that the fetus is affected.

Maternal affect is another possibility. Since a sperm and an egg contribute unequally to the zygote's cytoplasm, there might be some cytoplasmic factor mutated in an affected mother that predisposes the offspring to the condition.

3. a) Nondisjunction occurred in the **mother**.
- b) Since the daughter is heterozygous at the SSRs that are closely linked to the centromere, this shows that nondisjunction occurred in **meiosis I**.



Note: This is only one possibility for phase of the alleles. However, we know that crossing over occurred between SSR3 and SSR4, and we know which alleles are in the egg.

- d) If the daughter got a Y chromosome from her father that lacked SRY, then she could develop as a female.
- e) The presence of SSR1 from the father suggests that there might have been crossing over between the X and the Y chromosomes in the father. Consequently, the Y lost SRY and gained the SSR1 from the X chromosome.