

Name:  
Recitation Section:

## 7.03 Problem Set 3

### Due Friday, March 20, 2015 by 3 PM

Remember to show your work for all questions.

1. You are investigating the genetics of a newly discovered strain of *E. coli*. In order to do this, you infect the bacteria with  $\lambda$  Pam,  $\text{int}^-$  :: Tn5 phage to create knockout mutants. You grow the infected bacteria on kanamycin-containing media in order to select for the bacteria in which transposition occurred. Out of 10000 colonies that grow, you notice that 4 grow in the shape of a star.

a) You know that the genome size of this strain of *E. coli* is 5 Mb. Approximately what size is the gene causing the star-shaped colony phenotype? (Assume that a single gene is the cause, and that a transposon insertion anywhere in the gene would impede its function.)

You want to sequence the gene, but first you need to amplify it. In order to do this, you choose one of the star-shaped colonies to sequence the gene from. You plan to use restriction enzymes to cut its DNA and then circularize the DNA fragments and perform PCR.

b) You want the entire gene to be present on the fragment of DNA that contains your transposon in order to sequence it. How many base pairs long does the recognition sequence of the restriction enzyme you use need to be in order for this to be likely?

c) Once you've cut the genomic DNA into fragments, you circularize the DNA fragments and perform PCR to amplify your gene of interest. Draw the circular DNA fragment containing your gene, labeling all genes and regions on the fragment. Also draw the locations of the PCR primers you would use for this amplification. (Be sure to indicate the direction of the PCR primers.)

d) You sequence the PCR product of the gene. Now that you have the sequence of this gene in hand, you decide to investigate the other 3 colonies that showed the star-shaped phenotype. Using modern high-throughput methods, you obtain the RNA and protein content of each of 3 colonies. The first colony contains a truncated version of the protein that the gene codes for; the second colony contains mRNA for the gene, but no protein at all; and the third colony doesn't even contain any mRNA from the gene. For each of these colonies, where is the most likely location in the gene that the transposon was inserted into? Explain your reasoning.

2. To learn about genetic screens, you do a mutagenesis screen to create a collection of *E. coli* strains that contain a mutant LacZ gene. You choose to conduct your screen by treating your *E. coli* cells with 5-bromouracil, which is a derivative of uracil that can substitute for thymine in DNA. It is able to base pair with both adenine and guanine, leading to conversion of AT base pairs to GC and/or GC to AT during DNA replication.

A functional LacZ gene product will be able to hydrolyze the synthetic compound X-gal to produce a blue pigment yielding a blue colony. LacZ<sup>-</sup> colonies will remain white.

You screen 10<sup>4</sup> colonies and find 18 white colonies on X-gal plates and designate the mutants LacZ<sup>-</sup>(A-R).

You decide to sequence the LacZ gene of each of your mutants and one of them, LacZ<sup>-</sup>J, gives the resulting sequence (beginning at the start site):

5'ATGCACGGTTACGATGCGCCCATCTACACCAACGTGACCTAGCCCATTACGGTCAATCCGCCG  
TTTGT..... 3'

Bases marked in red differ from the wild-type consensus sequence.

a) What type of mutation would you suspect is causing the LacZ gene malfunction? What is the result of this mutation on the LacZ protein?

b) Describe a way to quickly test your hypothesis and interpret possible results.

c) You accidentally forget to put an X-gal plate of LacZ<sup>-</sup>F colonies into the incubator. When you come back to lab after a few days off, you find that all of your LacZ<sup>-</sup>F colonies have grown up and turned blue while sitting on your bench. What type of mutation does LacZ<sup>-</sup>F likely carry?

You decide to do a suppressor screen with your LacZ<sup>-</sup>F mutant. This time you use UV to mutagenize and find 9 suppressors, labeled 1-9.

d) How does UV exposure lead to mutagenesis?

e) As a first step to characterize these suppressor strains, you decide to sequence the LacZ gene in each strain. What types of suppressors will this allow you to identify?

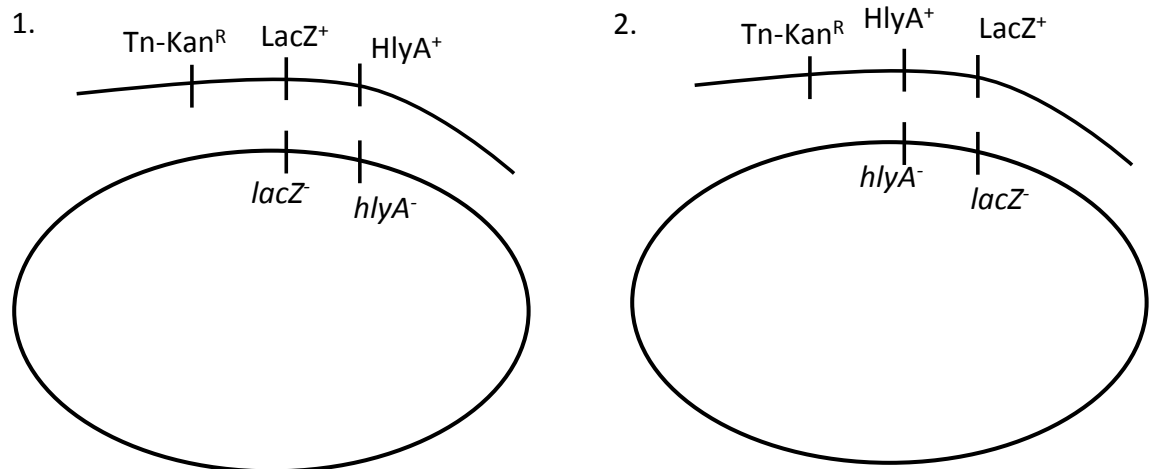
3. You are studying virulence factors in an *E. coli* strain. You are interested in genes that produce hemolysins, factors that promote lysis of red blood cells, which can be a cause of pathogenicity. Hemolytic activity can be detected by plating bacteria on blood agar, which contains mammalian red blood cells. Presence of hemolysins in a strain will cause a white clearing around the resulting colonies.

a) You have strain in which one of the 3 hemolysin genes *E. coli* carries,  $\alpha$ -hemolysin, is mutated ( $hlyA^-$ ) and cannot cause a white clearing ( $hly^-$  phenotype). You also have an  $HlyA^+$  strain that carries a Tn5 insertion near HlyA. You grow P1 phage on this strain and then collect phage lysate that you use to infect the  $hlyA^-$  strain. You grow the infected strain on media containing kanamycin to select for kanamycin resistant ( $Kan^R$ ) transductants. Out of 200  $Kan^R$  transductants, you isolate 158 that are  $hly^-$  and 42 that are  $Hly^+$ . What is the distance between the Tn5 site and the  $hem^-$  mutation in terms of cotransduction frequency?

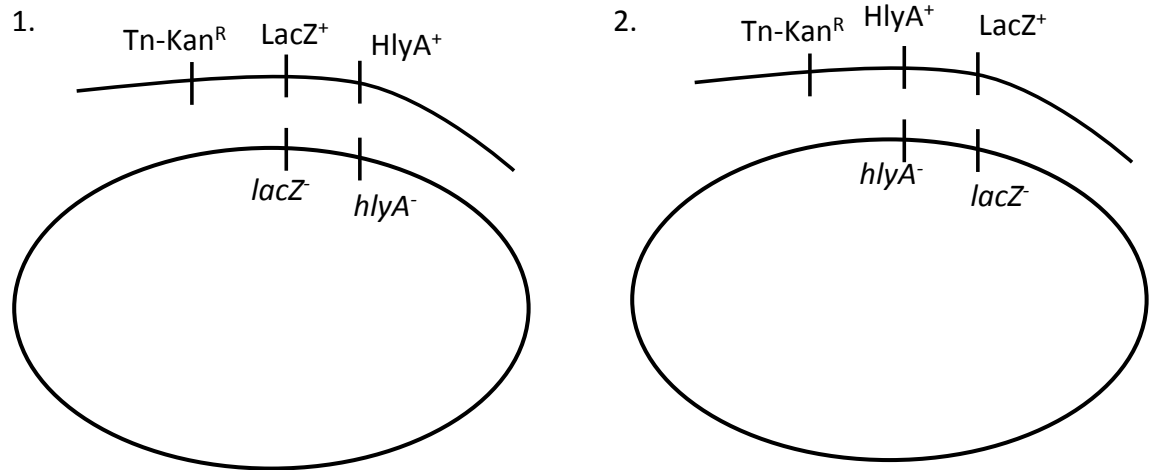
The LacZ gene is also nearby and you wish to determine the order of the three loci (the Tn5 insertion site, HlyA, and LacZ). You repeat the above experiment and grow P1 phage on an  $HlyA^+ LacZ^+$  but this time use your phage lysate to infect an  $hlyA^- lacZ^-$  strain and score for ability to grow on lactose (ability to hydrolyze X-gal to make blue colonies).

b) Consider 2 potential orders: #1: Tn5 – LacZ – HemA  
#2: Tn5 – HemA – LacZ

(i) Draw the crossover events that will yield  $Kan^R$  bacteria that can hydrolyze red blood cells but not grow on lactose.



(ii) Draw the crossover events that will yield Kan<sup>R</sup> bacteria that cannot hydrolyze red blood cells but can grow on lactose.



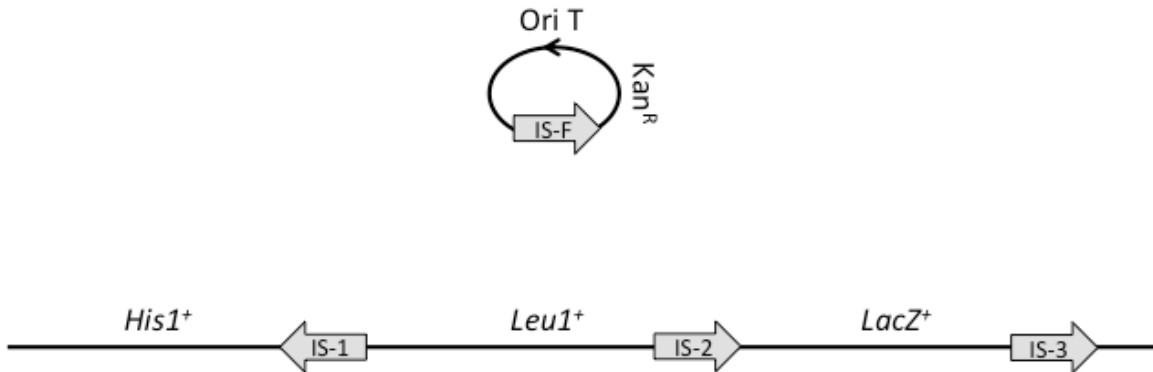
c) Your lab ran out of X-gal; how else could you score for the ability to grow on lactose? Draw out how you would identify cells that are lac<sup>-</sup>.

d) Among 1000 Kan<sup>R</sup> transductants scored, you find:

- 487 can hydrolyze red blood cells and grow on lactose
- 469 cannot hydrolyze red blood cells and cannot grow on lactose
- 41 can hydrolyze red blood cells and cannot grow on lactose
- 3 cannot hydrolyze red blood cells and can grow on lactose

Which of the two orders we've considered is more likely to be correct and why?

4. You have *E. coli* that contain an F plasmid with an insertion sequence, labeled IS-F in the diagram below. The F plasmid also contains a kanamycin resistance gene. On the *E. coli* chromosome are three copies of the insertion sequence, labeled IS-1, IS-2, and IS-3. In the region containing these three insertion sequences are three genes: *His1*, which is necessary for histidine synthesis; *Leu1*, which is necessary for leucine synthesis; and *LacZ*, which is necessary for lactose hydrolysis. The bacteria are functional at all three genes. The locations of these three genes relative to the insertion sequences are also shown in the diagram.



a) The plasmid can integrate into any of the three chromosomal insertion sequences through homologous recombination to form an Hfr bacterium. Draw the Hfr bacteria that result from integration at all three chromosomal insertion sequences. Be sure to show all insertion sequences, genes, and the F plasmid's origin of transfer.

(i) Insertion at IS-1

(ii) Insertion at IS-2



(iii) Insertion at IS-3

You take one Hfr strain that you generated and mix it with a strain of *E. coli* that is  $F^-$ ,  $His1^-$ ,  $leu1^-$ ,  $LacZ^-$ , susceptible to kanamycin, and resistant to ampicillin. You leave the two strains together for varying amounts of time before replica plating them on several different types of media, all of which contain both kanamycin and ampicillin. You notice that some colonies are  $his^- leu^- Lac^+$ , some are  $his^- leu^+ Lac^+$ , and some are  $his^+ leu^+ Lac^+$ . No other phenotypes are observed.

b) Based on this information, at which insertion sequence did the F plasmid integrate in the Hfr strain? Explain your reasoning.

c) In the experiment above, why were all the bacteria grown on media that contained both kanamycin and ampicillin?

After repeating the experiment in part B several times with the same Hfr strain, you suddenly get a different result in one experiment: the three phenotypes generated are  $his^- leu^+ Lac^-$ ,  $his^- leu^+ Lac^+$ , and  $his^+ leu^+ Lac^+$ .

d. What single recombination event in the Hfr strain could have caused this result? Draw both the *E. coli* chromosome during the recombination event and after the recombination event.

While doing yet another replication of the experiment in part B with the original Hfr strain, you get yet another variant result: no matter how long you allow the Hfr and  $F^-$  bacteria to sit together before plating on kanamycin- and ampicillin-containing media, you can only ever grow bacteria that are  $his^- leu^- Lac^+$ .

e) What single recombination event in the Hfr strain could have caused this result? Draw both the *E. coli* chromosome during the recombination event and after the recombination event.