Name MEHMET TUGRUL SAVRAN

Fall 2014 7.012 Problem Set 3-20 Points total
DUE FRIDAY, 10/10/2014 AT 9:45AM ONLINE

TA: MUELLER

Question 1 (3 points)

Suppose you are working in a lab that studies the base composition of different genomes. You look at the percent base composition of 5 different genomes and you get the data shown in the table below. You know that three of the genomes are bacterial and 2 are viral.

	%A	%C	%T	%G
Genome 1	23	27	23	27
Genome 2	24	20	30	26
Genome 3	29	21	29	21
Genome 4	11	39	11	39
Genome 5	20	25	0	19

a) **Based on this information, determine which genomes are bacterial and which genomes are viral? Explain your answer.** Genomes 2 and 5 are viral. Genomes 1, 3, 4 are bacterial. 2 is viral because its base pair amounts are different. 5 is viral because there is no thymine in it. It most probably has uracyl. So it is viral. The rest are bacterial because their base-pair amounts are consisten

b) **Give a possible explanation for why you see 0% percent T for genome 5. Is it possible that this genome only has 3 bases?** Genome 5 is viral. This is why it has no thymine in it. It will have uracyl instead. So NO, it still has 4 bases.

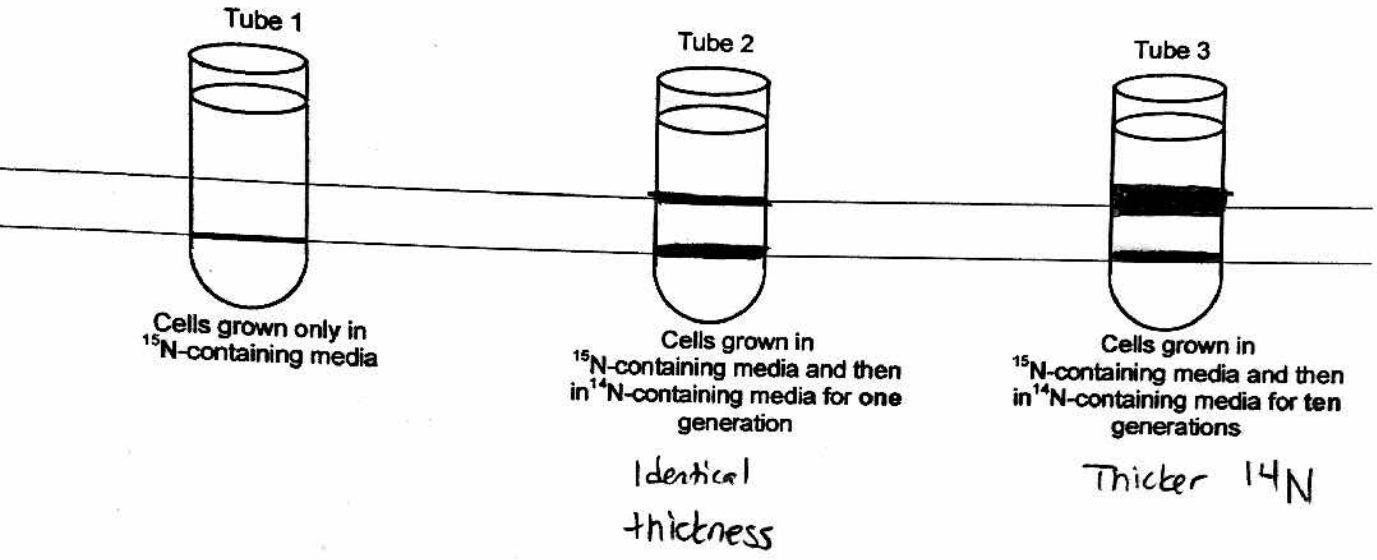
c) **How could you test your hypothesis in part b?**
The uracyl percentage can be checked.

d) **You calculate the total number of hydrogen bonds between bases for all bacterial genomes. You find out that one of the bacterial genomes has significantly more hydrogen bonds as compared to the other two genomes. Based on the data provided, can you identify this genome? Explain your answer.** Yes. Genome 4, because 3 hydrogen bonds are made between C-G bases. And Genome 4 has the most G-C percentage in it.

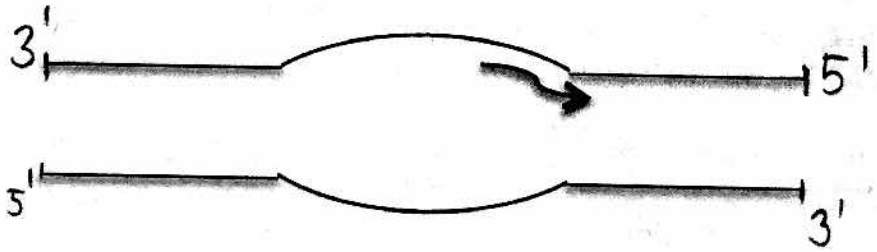
Question 2 (5 points)

DNA replication is semiconservative. Each DNA helix contains one old strand and one newly synthesized strand. The conservative model of DNA replication was also considered before the Meselson-Stahl experiment. According to this model, the double stranded DNA helix serves as a template to produce a new helix. Replication produces two DNA helixes, one comprised entirely of old DNA and the other one comprised entirely of new DNA. **Note-this model is not correct.**

a) The diagram below depicts tubes containing the denoted samples in a cesium chloride gradient after ultracentrifugation. Draw the bands you expect to see in tube 2 and 3, if DNA replicates via a conservative model. If you notice more than one band in tube 2 and/or 3, comment on their relative thickness.



b) A schematic of an origin of replication is shown below. The arrow indicates the direction of DNA replication. Label the 5' end and 3' end of all DNA strands.



c) Suppose you are carrying out the replication of this fragment of DNA in a tube. You add all the necessary components and enzymes for this process, but you leave out DNA ligase. Which strand would the absence of ligase affect? Your answer options are top, bottom, neither or both. Explain your answer.

Ligase joins Okazaki fragments, therefore its absence will affect both the top and bottom strands. Because both strands will have ~~dis~~ uncompleted fragments.

Name SAVRAN

d) You are considering deleting the gene encoding the ligase enzyme in bacteria. What would you expect the phenotype of these bacterial cells to be?
DNA replication cannot occur! So it won't be able to reproduce.

e) You find out that if you grow *E. coli* using a growth medium supplemented with a certain substrate A, they appear blue. This happens because bacteria produce an enzyme that converts substrate A into the blue color.

You then come across a mutated strain of *E. coli* in which the exonuclease activity of the DNA polymerase has been compromised. You plate these cells in medium containing substrate A and notice that the majority of the single colonies are blue. However, you also observe a few white colonies. Give a possible explanation for the appearance of the white colonies.
Exonuclease activity fixes the mistakes of base-pairing. If this feature is compromised, one will observe mutations. This is why white colonies are seen.

Question 3 (7.5 points)

Ampicillin is an antibiotic used widely in treating bacterial infections. Ampicillin kills bacteria by interfering with the synthesis of the cell wall. Without a proper cell wall, bacterial cells will lyse and die. However, bacteria that contain the gene, ***bla***, are resistant to the antibiotic ampicillin. This happens because the protein encoded by *bla*, beta-lactamase, destroys ampicillin and allows bacteria to survive and propagate in its presence. Below is shown a fragment of the *bla* gene. A part of its promoter has been also labeled with a box.

```
5′-CTTTTTGCGTTTCTACAAACTCTTTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGC-3′
3′-GAAAAACGCAAAGATGTTTGAGAAAACAAATAAAAAGATTTATGTAAGTTTATACATAGGCG-5′

5′-TCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATT-3′
3′-AGTACTCTGTTATTGGGACTATTTACGAAGTTATTATAACTTTTTCCTTCTCATACTCATAA-5′

5′-CAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCA-3′
3′-GTTGTAAAGGCACAGCGGGAATAAGGGAAAAAACGCCGTAAAACGGAAGGACAAAAACGAGT-5′

5′-CCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGT-3′
3′-GGGTCTTTGCGACCACTTTCATTTTCTACGACTTCTAGTCAACCCACGTGCTCA-5′
```

a) From the information above, do you know the direction in which RNA polymerase moves in transcribing the *bla* gene? If the answer is yes, does it transcribe to the left of the promoter or to the right of the promoter? If the answer is no, explain why.
No. We cannot know. This is because they both have TA.TA sequence, which is where RNA polymerase usually binds to. 3
∴ Either can be the template.

Name SAVRAN

b) You are told that the first transcribed nucleotide is a uracil and this is transcribed from the first nucleotide outside the boxed region. Indicate which strand (top or bottom) is the coding strand.

Top strand.

c) Which strand is the template strand for transcription?

Bottom strand.

d) Write down the first ten nucleotides of the mRNA. Remember that the first nucleotide is a uracil transcribed from the first nucleotide outside the boxed region. Clearly label the 5' and 3' ends.

5' UUGAAAAAGG 3' mRNA

e) Write down the first three amino acids produced from the beta-lactamase gene. The amino acid codon chart is provided on the last page of this problem set for your convenience.

Leu Lys Lys

f) Would you expect the TAA sequence boxed below to lead to a stop in translation?

No, because top strand is not the template strand for transcription.

```
5′-CTTTTTGCGTTTCTACAAACTCTTTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGC-3′
3′-GAAAAACGCAAAGATGTTTGAGAAAACAAATAAAAAGATTTATGTAAGTTTATACATAGGCG-5′

5′-TCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATT-3′
3′-AGTACTCTGTTATTGGGACTATTTACGAAGTTATTATAACTTTTTCCTTCTCATACTCATAA-5′

5′-CAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCA-3′
3′-GTTGTAAAGGCACAGCGGGAATAAGGGAAAAAACGCCGTAAAACGGAAGGACAAAAACGAGT-5′

5′-CCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGT-3′
3′-GGGTCTTTGCGACCACTTTCATTTTCTACGACTTCTAGTCAACCCACGTGCTCA-5′
```

Name SAVRAN

g) Imagine you mutate the given sequence at the following locations 1-4. Each mutation is described in more detail below:

1-The G nucleotide is deleted
2-The A nucleotide is converted to a C
3-The A nucleotide is converted to a G
4- The T nucleotide is deleted

Describe the specific effect each mutation will have on the protein sequence. Be as detailed as possible.

```
5′-CTTTTTGCGTTTCTACAAACTCTTTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGC-3′
3′-GAAAAACGCAAAGATGTTTGAGAAAACAAATAAAAAGATTTATGTAAGTTTATACATAGGCG-5′

                          1                                   2
5′-TCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATT-3′
3′-AGTACTCTGTTATTGGGACTATTTACGAAGTTATTATAACTTTTTCCTTCTCATACTCATAA-5′
    3   4
5′-CAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCA-3′
3′-GTTGTAAAGGCACAGCGGGAATAAGGGAAAAAACGCCGTAAAACGGAAGGACAAAAACGAGT-5′

5′-CCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGT-3′
3′-GGGTCTTTGCGACCACTTTCATTTTCTACGACTTCTAGTCAACCCACGTGCTCA-5′
```

1) Not effect at all; it's before promoter.

~~2) It changes purine to pyrimidine base. The codon will become UAU → UCU. Changes the a.a expressed to ser. : misinterpretation~~

~~3) Changes purine to purine base. The codon will become ACA →~~

2) It is changing the start codon, AUG. So transcription will not take place.

3) This will have no effect. Because the new codon will also stand for the same a Glutamine.

4) Deletion causes the entire frame to be shifted. Basically, every amino acid expressed w be different!

h) Suppose you generate 4 strains of bacteria each containing one of the mutations from part g. Bacteria strain 1 contains mutation 1, bacteria strain 2 contains mutation 2 etc. Which bacteria would you expect to survive if treated with ampicillin and which bacteria would you expect to die. Explain your answer.

Survivors

Bact. 1) Because mutation 1 does not affect anything.

Bact. 3) Because the resulting a.a will be the same, Bac 3 won't be affected.

Dead

Bact 2) Because it won't be able to synthesize protein.

Bact 4) Because it will result in very diffrent proteins, Bact 4 will die as well.

Name SAVRAN

Question 4 (2.5 points)

Consider a hypothetical gene. You know that in the coding strand the distance between the first ATG and the first in-frame TAA is 1200 base pairs (this number includes the first ATG bases, but not the TAA bases).

a) How many amino acids would you expect the protein product of this gene to be? ~~1200 / 3 = 400~~ $\frac{1200}{3} = 400$

b) You find out that the protein product is 200 amino acids long. Based on this information, could you tell whether this is a prokaryotic gene or eukaryotic gene? Explain your answer.
Eukaryotic. As the gene has introns, many pairs would be deleted.

c) You create a mutant of this gene that is missing 2 base pairs between the ATG and the TAA sequences from part a. What would you expect the effect of this mutation to be on the protein sequence? It will cause a shift in reference frame, which could lead to nonfunctional proteins. Thus, wrong amino acid chain will be synthesized.

d) You find out that this mutation has no effect at all on the resulting protein. The protein contains all the same 200 aminoacids in part b. How could you explain this?
Oh, So this mutation actually happened in introns!

Question 5 (2 points)

Go to the MITx/edX site https://lms.mitx.mit.edu/courses/MITx/7.012/2014_Fall/about and log into your 7.012X account. Navigate to "Courseware". Within week 5 you will find a section labeled Problem Set 3-REQUIRED. You must complete this section.
Your TA can check that you completed this through the 7.012x site, so you do not need to submit any work here.