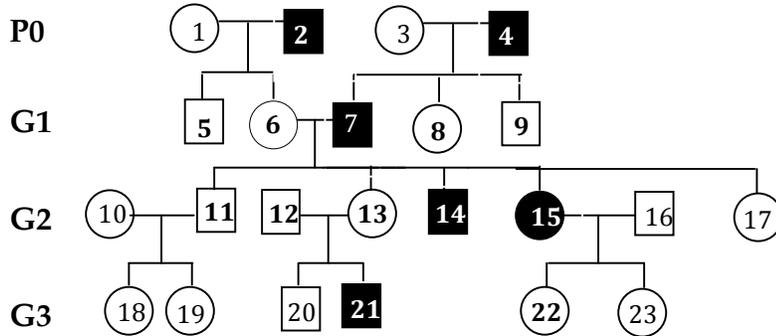


Solution Key- 7.013 Problem Set 3- 2013

Question 1

You are analyzing the following human pedigree for a disease. **Please note:** The filled squares or circles represent the abnormal phenotype. The disease related allele is rare in the population and individuals marrying into the family for generations G1- G3 do not have the disease allele. Assume that this disease is caused by the mutation of Gene A and no other mutation arises within the pedigree. Also assume complete penetrance.



Note: Use the symbol X^A , X^a , A or a where appropriate. In each case, use the letter "A" to represent the allele associated with the dominant phenotype and "a" to represent the allele associated with the recessive phenotype.

- a) Give the most likely **mode of inheritance of disease** shown in this pedigree. *X- Linked recessive*
- b) List **all** the possible genotypes of the following individuals in the above pedigree.

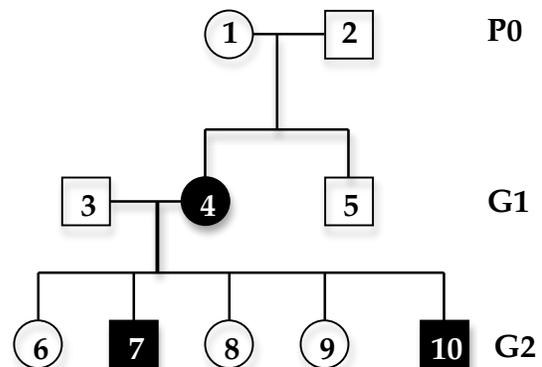
Individual 1: $X^A X^a$ or $X^A X^A$ **Individual 6:** $X^A X^a$ **Individual 20:** $X^A Y$ **Individual 22:** $X^A X^a$

c) **Individual #8** from this pedigree mates with a person who has wild-type allele(s) of Gene A associated with this disease. Together they have a son and a daughter.

- What is the probability of their **son** being **affected**? $\frac{1}{2}$ since #8 is a carrier
- What is the probability of their **daughter** being **affected**? 0
- What is the probability of their **daughter** being a carrier? $\frac{1}{2}$ since #8 is a carrier

d) Your friend shows you the following pedigree and says that the disease shown in this pedigree has the same mode of inheritance as the pedigree above. Would you agree or disagree with your friend? Provide an **explanation** for your choice.

You would disagree. Pedigree 1 is showing an X linked recessive mode of inheritance. If the 2nd pedigree was also showing the same mode of inheritance, the affected female (#4) should be receiving one copy of the disease associated allele, with its location on the X chromosome, from her mother (#1, making her a carrier) and the other copy from her father (#2, making him affected). But the father is not shaded/ affected in the pedigree. Thus this pedigree is not consistent with the X- linked recessive mode of inheritance.



Question 2

a) Complete the table below **for a eukaryotic cell**, by choosing from the following keywords. If there is more than one good choice, **include all**. Choose from *lysosomes, rough endoplasmic reticulum (RER), golgi body, nucleus, mitochondria, smooth endoplasmic reticulum (SER) cytoskeleton, extracellular matrix (ECM), ribosomes, nucleolus and cell membrane*.

| Description | Location or organelle in a eukaryotic cell |
|---|--|
| The location of circular deoxyribonucleic acid (DNA) | <i>Mitochondria</i> |
| The location of enzymes, in a eukaryotic cell, that are involved in digesting cell's own toxic material. The abnormal functioning of these enzymes may result in Tay Sachs disease. | <i>Lysosomes</i> |
| Membrane bound tubules that are involved in lipid metabolism and cholesterol synthesis and are NOT studded with ribosomes | <i>SER</i> |
| The component of cells where the ribosomes are synthesized | <i>Nucleolus</i> |
| The final location of proteins that glue the eukaryotic cells together and may participate in cell-cell communication | <i>ECM</i> |
| Membrane bound tubules that are involved in protein synthesis and are studded with ribosomes | <i>RER</i> |
| Stacked membrane bound tubules, in a eukaryotic cell, where a newly synthesized protein goes through post-translational modification i.e. glycosylation and phosphorylation. | <i>Golgi bodies</i> |
| Non-membrane bound structures that are involved in protein synthesis and comprised of RNA and proteins | <i>Ribosomes</i> |
| The final location of proteins that are a part of a dynamic network of fibers that maintain the shape and motility of a cell | <i>Cytoskeleton</i> |

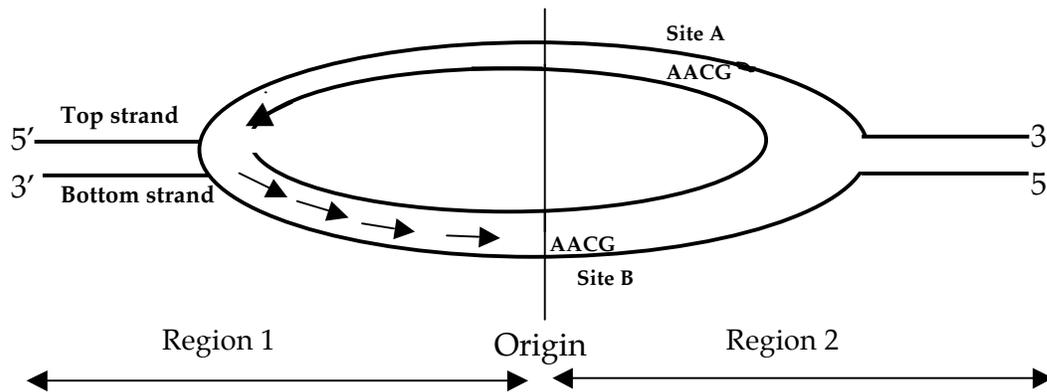
b) You label the DNA of a eukaryotic cells using BrdU (a thymidine analog). You then remove the nucleus of this cell (enucleated cell). You are surprised to see that the enucleated cell still has some DNA that is BrdU labeled. How can you **explain** this?

*As per the endosymbiotic theory, the mitochondria are primitive prokaryotic cells that got embedded into a bigger cell and continued staying in a symbiotic relationship with it. Therefore mitochondria have their own genome and hence the DNA, which is getting labeled with BrdU. It is this BrdU labeled DNA, which you are seeing in the cell even after the removal of the nucleus. **Note:** Its worth remembering that although we have paternal and maternal chromosomes we ALL receive our mitochondrial DNA from our moms.*

c) Ebola is a hemorrhagic virus, which disrupts the blood vessels. A recent publication demonstrated that this virus gains entry into the cells with the aid of the lysosomal protein NPC, found in the target cells. Briefly describe an experiment that supports the involvement of NPC proteins in Ebola infection. *There are multiple experiments that you can design. The following experimental design is based on what was discussed in the lecture. You knock out both the alleles of NPC genes (generate NPC⁻/NPC⁻ cells) from the genome of the target cells and see if these cells are more susceptible or/ resistant to the Ebola virus infection compared to the cells that are NPC⁺/NPC⁻ or NPC⁺/NPC⁺.*

Question 3

Shown below is a schematic of a replicating DNA in a bacterial cell.



- On the diagram, label the 5' and the 3' ends of the parental DNA strands.
- Which parental DNA strand (*choose from top or bottom*) serves as a template for the synthesis of the **leading strand in Region 2**? *Bottom strand*
- To which site (*choose from A, B, or both*) can the primer 5' UUGC 3' bind? *Site B*
- You perform **one round** of DNA replication in a test tube (in vitro) **using a single-stranded linear DNA as the template and the appropriate DNA primer**. Complete the following table, for **one round** of DNA replication.

| Protein / Enzyme | Function | Required for one round of replication (Yes/ No)? |
|------------------|---|--|
| Primase | <i>It makes the RNA primers. The 3'OH ends of these primers are elongated by the DNA polymerase. This polymerase is a primase dependent enzyme.</i> | No , since the DNA primer is available |
| DNA polymerase | <i>It replicates the DNA in 5'->3' direction and proofreads it in a 3'->5' direction</i> | Yes , it will read the single stranded DNA in a 3'->5' direction to make a complementary strand in a 5'->3' direction |
| Ribonuclease | <i>It degrades the RNA primers once the DNA replicates</i> | No , since the replication is being done in vitro and there is no RNA primer to begin with. |
| DNA Ligase | <i>It joins or ligates the newly replicated DNA segments together.</i> | No , since you are making only one DNA strand continuously in a 5'->3' direction. |
| DNA Helicase | <i>It uncoils the DNA duplex thus making the two strands available to serve as a template for replication.</i> | No , since you are starting with only DNA strand that does not require further unwinding unlike the DNA duplex. |
| Topoisomerase | <i>It takes out the extra coils and allows the DNA to unwind</i> | No , since you are starting with only DNA strand that does not require further unwinding unlike the DNA duplex. |

Question 3 continued

e) While studying replication in a bacterial cell type you find a mutant that shows a 100 fold decrease in the fidelity of replication.

- i. Name the protein(s)/enzyme(s) that is mutated in this bacterial mutant and accounts for decreased replication fidelity.

DNA polymerase

- ii. What particular activity does this protein(s)/ enzyme(s) lack that regulates the fidelity of replication?

It lacks its 3'→5' exonuclease activity as a result of which it cannot proofread the replicating DNA.

f) The DNA polymerase enzyme binds to the “ori site” and starts elongating the 3'OH end of the RNA primer by adding the nucleotide bases that are complementary to the sequence of the template DNA strand. Each replication cycle results in the shortening of chromosomal ends.

- i. **Explain** why the **Ori site** in general is rich in adenine (A) and thymine (T) bases.

There are only two hydrogen bonds between an A/T base pair compared to three between every G/C base pair. The fewer hydrogen bonds allow the two strands to unwind easily at the A/T rich ori site.

- ii. The rate of mitotic cell division in an embryonic cell is much higher compared to a normal somatic cell in an adult. Briefly **explain** how the telomerase activity may regulate the number of mitotic divisions a cell can have during its life- time.

In the absence of a functional telomerase in adult somatic cells, the chromosomal ends continue to shorten with each cell division. This eventually results in the loss of genetic information, which is crucial for cell division and cell survival.

g) DNA, unlike RNA, is stable and therefore serves as the hereditary material for most organisms.

- i. What feature of RNA makes it less stable than DNA?

The hydroxyl group (-OH group) at the 2'C position of the ribose sugar in each ribonucleotide (rNTP) makes the RNA less stable and more reactive compared to the DNA.

- ii. The fidelity of transcription is always less than the fidelity of replication. **Explain** why is this so.

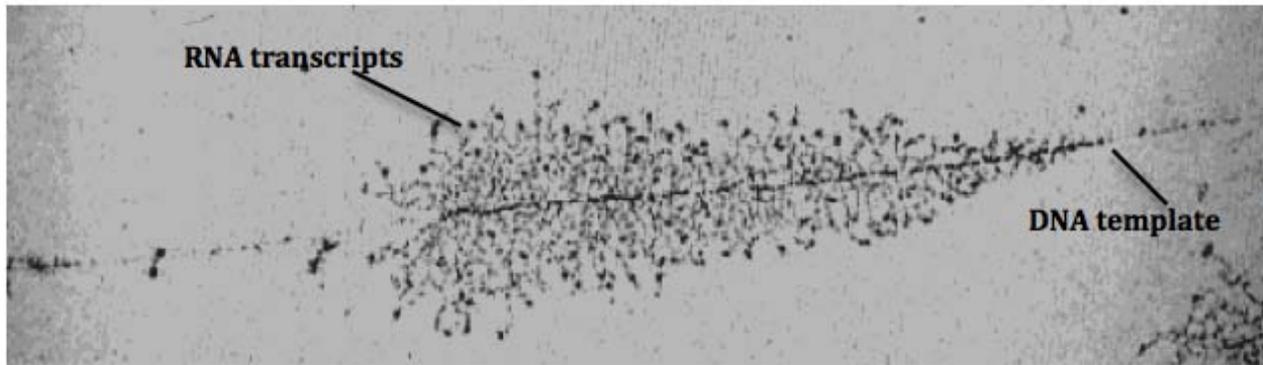
This is because DNA is transcribed into RNA by the RNA polymerase enzyme that lacks the proofreading ability unlike the DNA polymerase.

- iii. A cell much better tolerates errors during transcription than errors during DNA replication. **Explain** why is this so.

Transcription of a gene results in the generation of multiple copies of an mRNA transcript. Although some of these may carry mutations many others do not and hence they can be translated to make enough proteins, which allows the cell to maintain its normal structure and function. Besides, RNA is highly reactive and unstable compared to the DNA and hence the mRNA transcripts (including the ones carrying the mutations) will be degraded much rapidly thus will not hamper the structure and function of a cell.

Question 4

a) Below is an electron micrograph of a gene that is being transcribed. The DNA strand runs horizontally with RNA transcripts extending vertically outward.



- i. Draw an arrow to indicate the direction of movement of RNA polymerases along the length of DNA strand. Why did you choose this direction?

The longest mRNA will be associated with the RNA polymerases that have been transcribing the longest time.

- ii. Below is a partial sequence of the gene that is being transcribed in the schematic above. **Its orientation is the same as that shown in the schematic.** Which strand is the template strand, the top or the bottom strand? **Explain** your choice.

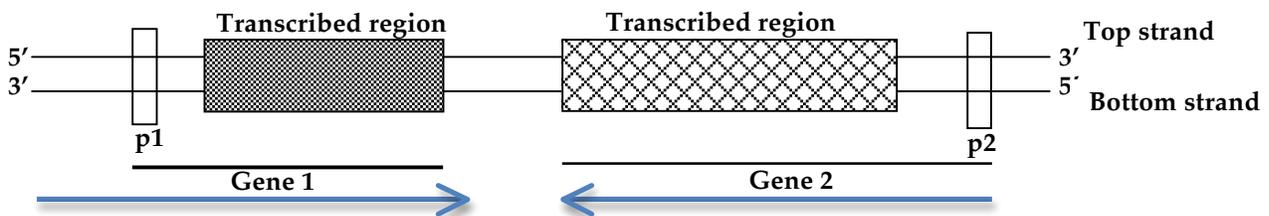
5' ACTCGATGCTAG3'
3' TGAGCTACGATC5'

The top strand is the template strand. From part (i) we already know that RNA polymerase is moving from right to left and we also know that mRNA is transcribed in a 5'→3' direction. Therefore the DNA template strand should be read from 3'→5' from right to left. The top strand satisfies this requirement.

- iii. Give the sequence of the mRNA that would be transcribed from the above sequence and label its 5' and 3' ends.

5'CUAGCAUCGUCGAGU3'

b) The following schematic shows the chromosomal location of Gene 1 and Gene 2. The corresponding promoters for Gene 1 (p1) and Gene 2 (p2) are shown.



- i. Would the direction of transcription of Gene 1 be the same as that for Gene 2 (Yes/ No)? **Explain** your choice with reference to the promoters.

No, it will not be the same. The direction of transcription is determined by the location of the promoter, which is always located prior to the transcribed region of the gene. Therefore for Gene 1, the bottom strand is the template DNA strand that is read 3'→5' by the RNA polymerase to make the mRNA in a 5'→3' direction. In comparison, for Gene 2, the top strand is the template DNA strand that is read 3'→5' by the RNA polymerase to make the mRNA in a 5'→3' direction.

- ii. On the schematic above, show the direction of transcription of Gene 1 and Gene 2 by drawing the arrows.
- iii. **Circle** the region(s) on each gene where the transcription factors would **most likely** bind and regulate transcription.

You should circle the promoters i.e. P1 for Gene 1 and P2 for Gene2.

Question 4 continued

c) You realize that Gene 1 is expressed in intestinal cells of mice and mutations in this gene are associated with Crohn's disease. Further analysis reveals...

- That the coding sequence of Gene 1 is the same in affected and wild-type mice.
- There is much less of mature mRNA transcribed from Gene 1 in the affected mice relative to that in the wild-type mice.
- The intestinal cells of the affected mice show no detectable protein unlike the wild-type mice.

Give two possible modifications of RNA, whose absence may explain the differences that you have observed between the affected and wild-type mice.

The mature mRNA lacks the 5'CAP and/ or 3'Poly A tail which stabilize the mRNA and promote its translocation from the nucleus into the cytosol where it can be translated. In the absence of these modifications, the transcribed mRNA will be very unstable and will not be available to get translated into a protein.

d) Gene 2 is expressed in mouse heart muscle cells and its over-expression (too much of it) is associated with heart disease. You try to reduce the expression of Gene 2 using a particular micro RNA (miRNA-32) that you think may regulate Gene 2.

- Circle** the sequence (choose from promoter, nascent (primary) RNA, mature mRNA, 5' UTR and 3' UTR) to which the miRNA is most likely to bind? 3'UTR

- Briefly describe how the miRNA may inhibit expression of Gene 2.

These are 21-22 bases long stretches of RNA which hydrogen bond with specific complementary bases at the 3'untranslated regions (3'UTR) of the transcribed mRNA and thereby inhibit its translation.

e) From your 7.013 lectures, you have learnt that although almost all cells in an organism contain an identical genome, gene expression differs in different cell types and is highly regulated by chromatin structure. Given the following treatments, explain whether chromatin structure **compacts** or **becomes less compact** and whether gene expression is **enhanced** or **reduced**.

- Cells are treated with DNA methylase, an enzyme that adds methyl groups to cytosine bases in DNA.

*The chromatin will be in the **condensed state** since enhanced methylation prevents histone removal resulting in more **condensed packaging/folding of chromatin** structure and reduced transcription. **Note:** It is worth noting that methylation at ALL positions do not have this effect, instead it is usually the methylation of base at 5C position that promotes condensation and close packing.*

- Cells are treated with histone acetylase, an enzyme that adds acetyl groups to the histone proteins thus making them less basic.

*The chromatin will be in the **open state** since enhanced methylation prevents histone removal resulting in more **unpackaging/unfolding of chromatin** structure and reduced transcription.*

f) As we mentioned in lecture, Prader-Willi (PW) syndrome is a rare disorder that results in intellectual disability, a constant sense of hunger and obesity. This syndrome is caused by deletion or inactivation of seven genes that lie on Chromosome 15. Normally, these genes are INACTIVE on the maternally-contributed chromosome, but ACTIVE on the paternally-contributed chromosome. PW syndrome occurs because the paternal genes that should be expressed are not for one of the following reasons:

- Paternal genes on chromosome 15 are missing or defective.
 - The patient has inherited two copies of chromosome 15 from the mother and no chromosome 15 from the father.
- Of what is the PW syndrome an example? **Explain** the reasoning behind this answer.

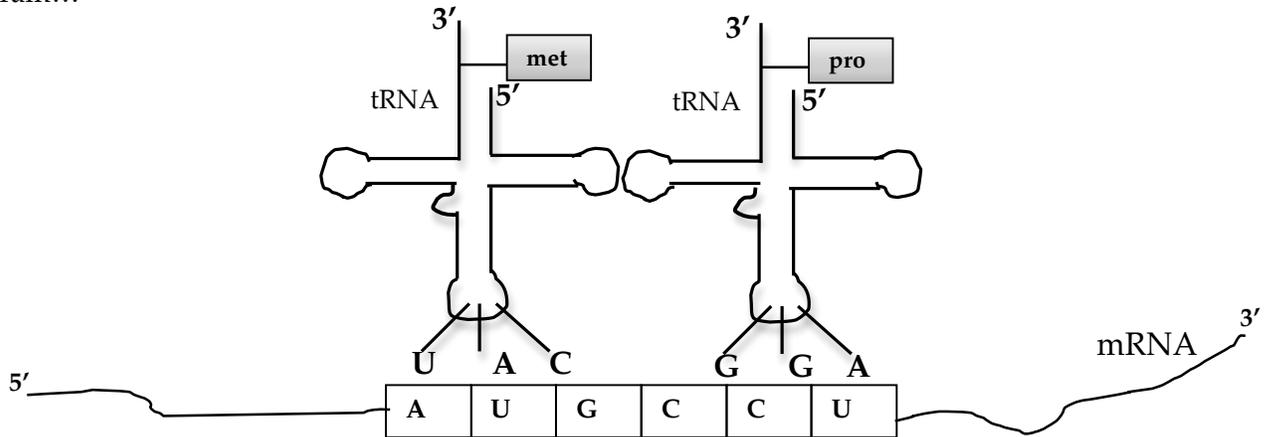
This is an example of mammalian imprinting i.e. inheritance of maternal or paternal specific methylation pattern. In the affected individuals, the maternally- contributed chromosome is silenced (similar to that in normal individual) but the paternal chromosome is either inactive or not present (unlike the normal healthy individuals) thus leading to the disease.

Question 4 continued

- ii. You isolate cells from a PW patient and try to activate the Chromosome 15 genes. Circle the option from below that you would try. **Explain** your reasoning.
- Treat cells with 5-azacytidine that inhibits DNA methylase
 - Treat cells with 2-chloroethyl-N-nitrosourea (CNU) that activates DNA methylase.

Question 5

a) Below is a diagram of two tRNAs and an mRNA in the active site of the ribosome during translation of the mRNA into protein. Three nucleotides from the sequence of each tRNA are shown for you. In the diagram...



- Label the 5' and 3' ends of each tRNA.
- Fill in the **empty boxes** in the mRNA with the 6 nucleotides that would be present there.
- Fill in the **shaded box** attached to one end of each tRNA with the name of the amino acid that would be attached there. Note: A codon chart is provided on the last page of this problem set.
- Which tRNA is about to transfer its attached amino acid over to the other tRNA to elongate the peptide chain (choose from the tRNA on the left or the tRNA on the right)?
The tRNA on the left. This allows the tRNA to the left to leave the ribosome and allow the peptide to continue being synthesized on the ribosomes.

Question 5 continued

b) The following schematic represents the 3D- conformation of a tRNA molecule. One end of the tRNA has an anti-codon that base pairs with the appropriate codon on the mRNA and the other end has an amino acid arm that binds to a specific amino acid. Below are three anti-codon sequences for three tRNAs. Fill in the corresponding amino acid on the blanks. *Note: A codon chart is provided on the last page of this problem set.*



| anticodon found on tRNA | amino acid attached to tRNA |
|-------------------------|-----------------------------|
| 5' AGU 3' | <i>threonine</i> |
| 5' AUG 3' | <i>histidine</i> |
| 5' CUG 3' | <i>glutamine</i> |

i. Give the anticodon used in the tRNA encoding tryptophan (trp) and label its 5' and 3'. 5'CCA3'

ii. Would a substitution within a codon for trp **always change** the resulting protein sequence? **Explain** your answer.

Yes, since there is only one codon for tryptophan. Hence a base substitution in either 1, 2 or 3 position of the codon would result either in a different amino acid or a stop codon.

iii. Would a substitution within a codon for threonine (thr) always change the resulting protein sequence? **Explain** your answer.

No, since threonine can be encoded by four different codons. All these four codons have the same base at 1st and 2nd position. However they all have a different base at the 3rd position. Therefore a base substitution at 3rd position will not change the protein sequence since the codon will still code of Threonine.

Question 6

a) The following is the partial DNA sequence of Gene 1. *Note: The underlined sequence (from position 20-54) represents the promoter for Gene 1 and the underlined and italicized sequence (from position 71-90) represents the ribosome binding (RBS) site of Gene 1. Transcription begins at and includes the bold T/A base pair at position 60.*

```

1           10          20          30          40          50          60          70
I-----I-----I-----I-----I-----I-----I-----I
5' ATCGGTCTCGGCTACTACATAAAACGCGCGCATATATCGATATCTAGCTAGCTATCGGTCTAGGCTACTAC
3' TAGCCAGAGCCGATGATGTATTTGCGCGCGTATATAGCTATAGATCGATCGATAGCCAGATCCGATGATG
                                     Promoter

           80          90          100         110         120         130         140
-----I-----I-----I-----I-----I-----I-----I
5' CAGGTATCGGTCTGATCTAGCTAGCTTCTCTTCTCTCTCCCCGCGGGGCTGACTATCATGCGTCG
3' GTCCATAGCCAGACTAGATCGATCGAAGAGAAGAGAGAGAGGGGGCGCCCCGACATGATAGTACGCAGC
                                     RBS

           150         160         170         180         190         200         210
-----I-----I-----I-----I-----I-----I-----I
5' TCTCGGCTACTACGTAACGCGCGCATATATCGATATCTAGCTAGCTATCGGTCTCGGCTACTACGTA
3' AGAGCCGATGATGCATTTGCGCGCGTATATAGCTATAGATCGATCGATAGCCAGAGCCGATGATGCATTT

```

i. Which strand (Top/ bottom) is the template strand for transcription of Gene 1? Bottom

Question 6 continued

ii. What are the first 6 nucleotides of the mRNA transcribed from Gene 1?

5'UAGGCU3'

iii. What are the first 4 amino acids encoded by Gene 1? *Note: A codon chart is provided on the last page of this problem set.*

N-met-arg-arg-leu-C

b) You have found mutant alleles of Gene 1 that have insertion of the following three base pairs immediately **after** the C/G base pair at positions, 100 and 150. Complete the table below for each of the following mutations.

5' CGA 3'
3' GCT 5'

| | Positions of insertions of the above sequence | |
|---|---|---|
| | After 100 th base pair | After 150 th base pair |
| Effect on the length of mRNA (longer than/ shorter than/ same as the mRNA produced by the wild- type allele of Gene 1). Explain. | It will be longer than the original mRNA by 3b since the insertion is after the transcription start site. | It will be longer than the original mRNA by 3b since the insertion is after the transcription start site. |
| Effect on the length of protein (longer than/ shorter than/ same as the protein produced by the wild- type allele of Gene 1). Explain. | Same as the original protein. Although this insertion is after the ribosomal binding site it is prior to the start codon and hence is not a part of the open reading frame. | Longer than the original protein by one amino acid since it is after the start codon. But it does not change the reading frame. |
| If the protein is produced, will it be functional? Explain. | Yes, since the translation of the protein following this insertion remains unaffected for the reasons described above. | Functional, if the extra amino acid does not change the 3D conformation of the protein. Nonfunctional if you do not make this assumption. |

Question 7

Following are three genes that encode small peptides. The splice site donor **always** has the composition 5'-GUG-3' (i.e. for the 5' end of the splice site) and the splice site acceptor always has the sequence 5'-UUG-3' (i.e. for the 3' end of the splice site), meaning that these sequences and everything between them will be removed from the mature mRNA. Only the **non-template/ mRNA like strands of DNA** for each gene are shown. *Note: A vertical line is drawn after every 10th base in each sequence. The bold and underlined ATG represents the correct reading frame.*

a) Box the **intron** in each sequence.

Gene 1

5' - GACTAAG ATG | ATTTGGGCAAACACAG GTGCA | CTTTCCCAGT | TTTTCTTTCT | TGGCGAACGC | CTGA - 3'

Gene 2

5' - ACA ATG ATAC | GGGGAATG GTGGAACGATATC | ACGCTAGAGATTAATTGTG | GATCGTAGAGGACGCCAACGC | CATGA - 3'

Gene 3

5' - GACT ATG ATAACTATC | GTGCAGCAGACCCG | SAAGCCCAGAAAACCACTGG | ATTG | TCTGGCAGAGAGGCCA | CATGATC - 3'

Question 7 continued

b) Which of the above genes (choose from Genes 1, 2 and 3) encodes the smallest peptide? Give the amino acid sequence of this peptide in an N → C direction. **Note:** A codon chart is provided on the last page of this problem set.

Peptide sequence from Gene 1: NH₂- M I WANT AN A-COOH (9 amino acids)

Peptide sequence from Gene 2: NH₂- M I RECEIVED AN A-COOH (13 amino acids)

Peptide sequence from Gene 3: NH₂- MIT IS GREAT -COOH (10 amino acids)

c) You create the following mutants that have changes in one of the three gene sequences above. State the effect of each of these mutations on the peptides produced.

i. The 27th base in the **non-template strand** of **Gene 1** was changed from **T** to **A**.

The 5'GUG3' splice site is lost. Accordingly, the intron will not be spliced out. The peptide produced will be longer, misfolded and most likely non-functional.

ii. The 49th base in the **non-template strand** of **Gene 2** was changed from **G** to **T**.

A pre-mature stop codon will be inserted resulting in a truncated non-functional peptide. This is an example of nonsense mutation.

iii. The 4th base in the **non-template strand** of **Gene 2** was changed from **A** to **T**.

This will result in the loss of first start codon. The 2nd start codon will be used to make the protein. The mRNA will be read in a different frame. Therefore the resulting peptide will have a different amino acid sequence and will be non-functional.

iv. The 47th base in the **non-template strand** of **Gene 3** was changed from **C** to **T**.

This mutation creates a new 3' splice acceptor site (3'UUG). This is an example of a mutation that changes the reading frame of mature mRNA and therefore the amino acid sequence of the resulting peptide.

| | U | C | A | G |
|----------|---|--|--|--|
| U | UUU Phe (F) UUC " UUA Leu (L) UUG " | UCU Ser (S) UCC " UCA " UCG " | UAU Tyr (Y) UAC " UAA Stop UAG Stop | UGU Cys (C) UGC " UGA Stop UGG Trp (W) |
| C | CUU Leu (L) CUC " CUA " CUG " | CCU Pro (P) CCC " CCA " CCG " | CAU His (H) CAC " CAA Gln (Q) CAG " | CGU Arg (R) CGC " CGA " CGG " |
| A | AUU Ile (I) AUC " AUA " AUG Met (M) | ACU Thr (T) ACC " ACA " ACG " | AAU Asn (N) AAC " AAA Lys (K) AAG " | AGU Ser (S) AGC " AGA Arg (R) AGG " |
| G | GUU Val (V) GUC " GUA " GUG " | GCU Ala (A) GCC " GCA " GCG " | GAU Asp (D) GAC " GAA Glu (E) GAG " | GGU Gly (G) GGC " GGA " GGG " |

MIT OpenCourseWare
<http://ocw.mit.edu>

7.013 Introductory Biology
Spring 2013

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.