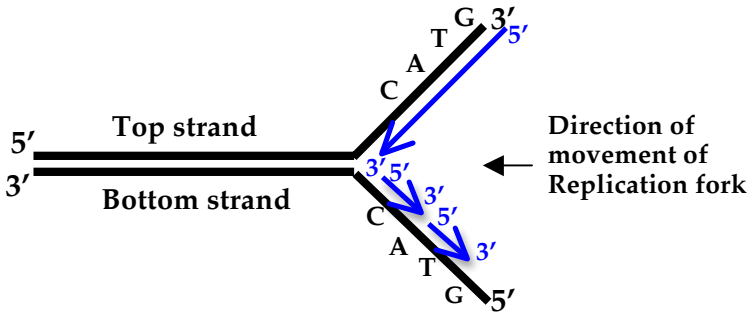
Solution Key- 7.013 EXAM 2 (4 / 3 / 13)

Question 1 (20 points)

a) In the table below, name the **sub-cellular location or organelle(s)** of the **eukaryotic cell** that will fluoresce when the following macromolecules are tagged with a fluorescent dye.

Macromolecules tagged with fluorescent dye	**Sub-cellular location or organelle(s) of cell that will fluoresce**
Proteins that add carbohydrates or lipids to the newly synthesized proteins	*Golgi body, ER*
Proteins that are a part of **functional** ribosomes	*Cytoplasm & ER*
DNA	*Nucleus & mitochondria*

b) Shown below is a segment of replicating DNA in the epidermal cells of the mice.



i. On the schematic, draw the **elongating DNA strands** and **label their 5′ and 3′ ends**.

ii. To which strand (*choose from* ***top****, bottom or both*) can primer **5′CATG3′** bind during replication?

iii. Which strand (*choose from* top *or* ***bottom***) is the template for **discontinuous (lagging) strand** synthesis?

iv. **Circle** the protein/ enzyme that relieves a replicating segment of DNA from super-coiling?

Helicases ***Topoisomerase*** **Primase** **Single stranded DNA binding proteins (SSDBP)**

d) You treat mouse epidermal cells in a plate with the drug, TAT-2. You observe that TAT-2 treated cells show reduced shortening of their chromosomes following each cell division and survive longer than the untreated cells.

i. Name the replication enzyme that serves as the target of TAT-2 and state whether TAT-2 *activates* or *inhibits* the function of this enzyme.

TAT-2 activates the telomerases

ii. Why does reduced shortening of chromosomes following each cell division promote long- term survival of a cell?

Telomerase aids in repairing the ends of chromosomes that progressively shorten after each replication cycle. This helps to preserve the genetic information that is crucial for the cell division, functioning and survival.

e) In a separate experiment, you irradiate the mouse epidermal cells, growing on a plate, with UV light. This treatment results in the formation of a **thymine dimer** (*a covalently joined pair of T-bases shown as bold and underlined*) as shown in the DNA segment below.

5′ ——— CT**TT**GCA ——— 3′
3′ ——— GAAACGT ——— 5′

Circle the process(s) (*choose from proofreading,* ***excision repair*** *or mismatch repair*) that will remove the thymine dimer and name the specific replication enzyme(s) that will **fill in and seal the gap** left after the removal of thymine dimer.

Enzyme that fills in the gap: *DNA polymerase* **Enzyme that seals the gap:** *Ligase*

Question 2 (18 points)

The following is the DNA sequence for the transcription initiation region of **Gene A that is expressed in epidermal cells** of mice. ***Note:*** *Part of the* ***promoter region*** *is boxed. Transcription* ***begins at and includes*** *the bold and underlined* ***A/T*** *base pair.*

```
5'----TGGACTGCTA  TAATAGCAGG  GCTGCCGAAT  GTGCTGCCAT  ACGGCCATGG TTCTTAAAGT----3'
3'----ACCTGACGAT  ATTATCGTCC  CGACGGCTTA  CACGACGGTA  TGCCGGTACC AAGAATTTCA----5'
```

a) Which DNA strand (*choose from top or* ***bottom***) serves as the **template strand** for transcription?

b) Fill in the **first 6 nucleotides** of the **primary/ nascent mRNA** transcribed from Gene A. *5'AGGGCU3'*

c) Fill in the **first four amino acids** of Protein A encoded by Gene A. ***Note:*** *A codon chart is provided on the last page.* ***You can detach the last page.***

N- met-cys-cys-his-C

d) The last 5 **amino acids (amino acid^{105}- amino acid^{109}) at the C- terminus** of wild-type Protein A are indicated below. Each of these amino acids is critical for the proper folding of this protein.

N - pro^{105}-asn^{106}-ser^{107}-met^{108}-leu^{109}-C

The DNA sequence encoding the **above 5 amino acids** is included within the sequence below.

Wild-type

```
5'-AACCGAATTCCATGTTATAGC-3'
3'-TTGGCTTAAGGTACAATATCG-5'
```

You isolate and sequence the following two different mutant alleles of **Gene A** that encode the above 5 amino acids. Each mutant allele is due to a **point mutation** that is **bold** and **underlined**. Which of these mutants will **ALTER** the folding of Protein A (*Choose from mutant 1 or mutant 2*)?

Mutant 1

```
5'-AACCAAATTCCATGTTATAGC-3'
3'-TTGGTTTAAGGTACAATATCG-5'
```

Mutant 2

```
5'-AACCGTATTCCATGTTATAGC-3'
3'-TTGGCATAAGGTACAATATCG-5'
```

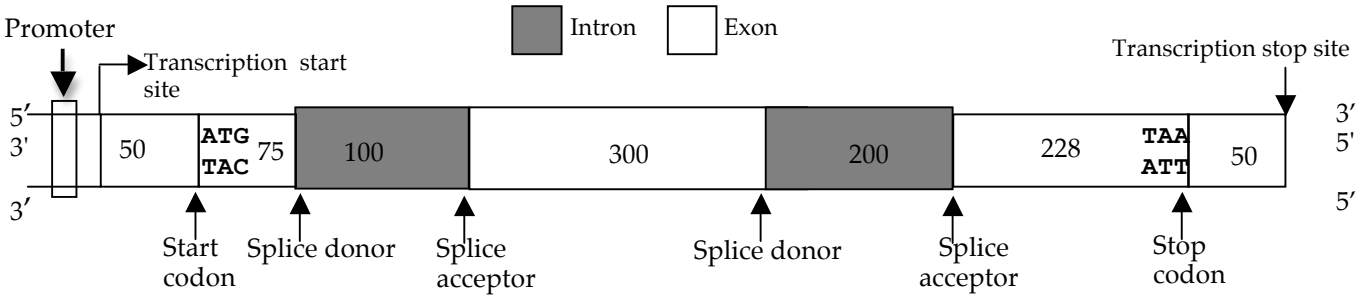
Explain, in terms of the change in the reading frame and/ or amino acid sequence, why you selected this mutant and **NOT** the other.

Mutant 1 will not alter the folding of Protein A since this is an example of silent point mutation that does not alter the amino acid sequence of the protein. In comparison, ***Mutant 2*** *is an example of missense point mutation that changes the codon 5'AAU3' (coding for asn^{106}) to 5'UAU3' (coding for tyr^{106}) thus altering the folding of this protein.*

e) You identify a disease of epidermal cells in mice in which **Gene A is not transcribed**. Further analyses reveals that the sequence of Gene A in affected and normal mice is the SAME. **Circle** the options, from the choices below that could explain why Gene A is **NOT** transcribed in the epidermal cells of the affected mice. In the epidermal cells of affected mice…

1. Mature mRNA corresponding to Gene A lacks the 5 methyl Cap and 3' Poly A tail
2. ***DNA around the promoter region of Gene A is methylated***
3. ***Epidermal cells of affected mice lack the transcription factors (TF) associated with Gene A.***
4. The Ribosome binding site in mature mRNA transcript corresponding to Gene A is mutated
5. The 3' untranslated region (3'UTRs) of mature mRNA corresponding to Gene A is mutated
6. Histones close to the Gene A are acetylated

Question 3 (20 points)
Shown below is the schematic of Gene A. The numbers within the boxes indicate the length (in base pairs) of each region. The DNA sequence corresponding to the translational start and the stop codons and the splice donor and splice acceptor sites are indicated.



a) You observe that Gene A is transcribed both in epidermal and muscle cells to produce a nascent / primary mRNA transcript. This mRNA directs the synthesis of two different proteins in these two different cell types.

- *In the muscle cells Gene A encodes a **protein** (100 amino acids long) that functions as a **nuclear protein (TF-1)***
- *In epidermal cells, Gene A encodes a **protein** (200 amino acids long) that functions as a cell membrane protein.*

Could Gene A direct the synthesis of **two different proteins** due to the…

Difference in….	**Explain** why you selected this option
Splicing (***Yes***/ *No*)?	*These proteins can be produced from the same gene due to alternative splicing of introns i.e. if the splice donor site of Intron1 base pairs with splice acceptor site of Intron 2 you get a mature mRNA corresponding toTF-1. In comparison, if both Introns 1 & 2 are spliced out as two separate exons you get a mature mRNA transcript that encodes the cell membrane protein.*
Protein processing (***Yes***/ ***No***)?	*Yes, if you assume that the nascent polypeptide chain in muscle cells is post-translationally cleaved to form functional protein of 100KD but it does not get cleaved in epidermal cells. No, if you say that post translational modifications such as glycosylation or addition of lipids may alter the molecular weight of the proteins but will not have any impact on the primary amino acid length of the proteins.*
Promoter sequence (*Yes*/ ***No***)?	*No, both proteins are encoded by the same gene and hence have the same promoter. A change in the promoter sequence affects the amount of gene expression but does not influence the type of gene products encoded by a gene.*

b) You want to study **another nuclear protein, TF-2** in mouse muscle cells. You identify a mutant cell line, which shows a **cytosolic location of TF-2** in muscle cells.

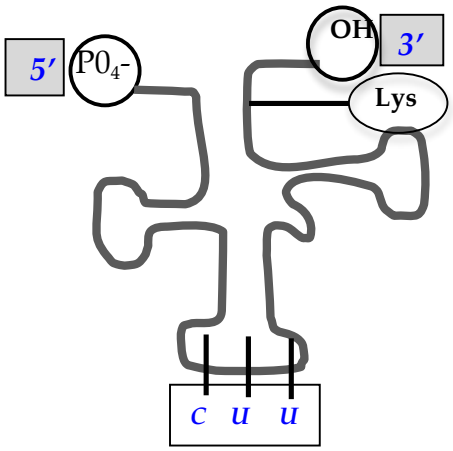
i. Name a stretch of amino acid sequence that the TF-2 in mutant muscle cell line lacks.
Nuclear localization sequence

ii. In the wild-type muscle cells, if this stretch of amino acid sequence is located at the N-terminus of TF-2, where in the mature mRNA transcript (*choose from the **5′ end** or the 3′ end*) would the corresponding base sequence be?

iii. The proteasome is a multi-protein complex that degrades any misfolded protein in a cell. How does the proteasome recognize which proteins in the cell are misfolded?
The proteasomes recognize the misfolded proteins once they are ubiquitinylated

Question 3 continued

c) The lys[60] (encoded by **5′AAG3′ codon**) of TF-2 is critical for its binding with its target sequence. The following is a schematic of tRNA specific for lys[60] (encoded by **5′AAG3′ codon**).



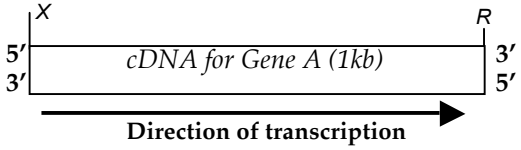
i. In the **blank box,** write the **anti- codon** that base pairs with the 5′AAG3′ codon for Lys**.**

ii. Label the 5' and 3' ends of the tRNA by filling in the shaded boxes.

d) Is the tRNA in the schematic ***charged*** or *uncharged*? **Explain** why you selected this option.
It is charged since it is covalently bonded to an amino acid.

Question 4 (24 points)

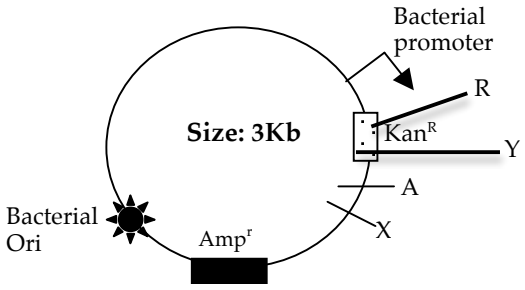
You decide to further characterize the TF-1 protein encoded by Gene A in muscle cells of mice. You adopt the following strategy to get a large amount of TF-1 protein for characterization. You make the cDNA using the mRNA derived from the wild- type allele of Gene A and by adding oligo- dT primers. The Gene A cDNA has the recognition sites for restriction enzymes X and R as shown below.



a) From the choices below, **circle** the sequences that are a part of the Gene A but are NOT contained in the corresponding cDNA.

Promoter **Exons** ***Introns*** **Enhancers** **3′UTR**

b) You want to clone the cDNA for Gene A into the following plasmid that has recognition sites for restriction enzyme Y, R, X and A as shown. ***Note****: A vertical line (|) represents the cutting site for each restriction enzyme.*

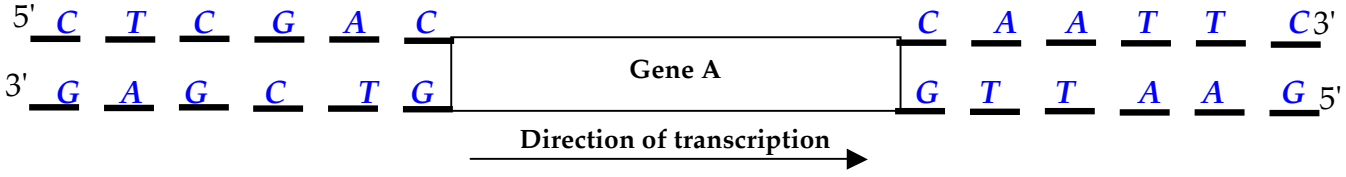


Note: Amp^R and Kan^R represent the ampicillin and kanamycin resistance genes.

X	A	R	Y
5′G/TCGAC3′	5′GAATT/C3′	5′CAATT/G3′	5′C/TCGAG3′
3′CAGCT/G5′	3′C/TTAAG5′	3′G/TTAAC5′	3′GAGCT/C5′

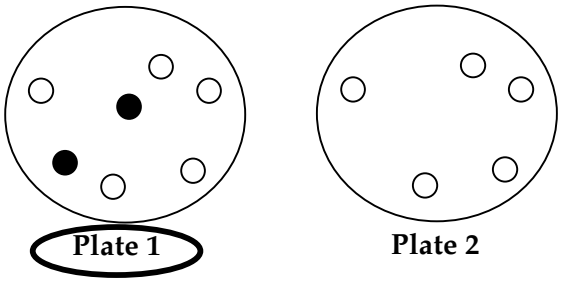
i. Which enzyme(s) (*choose from X, **A**, R or **Y***) would you use to cut the plasmid that would allow **directional cloning and expression** of Gene A from the bacterial promoter?

ii. Write the resulting **6- base pair sequence** at the **two points of ligation** after Gene A inserts into the plasmid.



Question 4 continued

c) You then plan to amplify the recombinant plasmid in *E. coli* bacterial cells. You transform the *E. coli* with the ligation mix and plate them on a master plate (*growth medium with no antibiotics*). You then replica- plate the colonies on **plate 1** (*growth medium + ampicillin*) and **plate 2** (*growth medium containing both ampicillin and kanamycin*). You obtain the following colonies.



i. **Circle** the plate (*1 or 2*) in the schematic that contains bacterial colonies that have the **recombinant plasmid.**

ii. In the plate that you circled, **fill in/ color** the colonies that contain the **recombinant plasmid**. **Explain** why you selected these colonies.

The recombinant plasmid will have an intact and functional Amp^R gene but a disrupted Kan^R gene where Gene A is inserted. So any bacterial cell that will receive the recombinant plasmid will be Amp^R Kan^S and will therefore grow on plate 1 but not on plate 2.

d) Another group is competing with you. Although they have the bacterial clone that contains the recombinant plasmid with Gene A in correct orientation they cannot express Gene A in bacteria. To understand what the issue is they decide to PCR amplify and sequence Gene A isolated from their bacterial clone (**mutant**) and compare it with the sequence of Gene A that you published (**wild- type**). Shown below is the sequence flanking the **mutant allele** of Gene A.

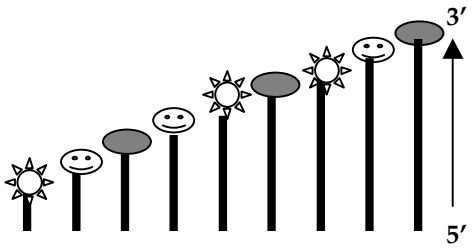
5′GAAATC Gene A GGACTT3′ **Top strand**
3′CTTTAG CCTGAA5′ **Bottom strand**

They have the following primers for amplifying Gene A by PCR reaction. **Circle** the **primers** that they would use to PCR amplify Gene A.

#1: *5′GAAATC3′* **#2: 5′TTCAGG3′** **#3: 5′CTTTAG3′** **#4:** *5′AAGTCC3′*

e) Sequencing results show that the mutant version has **one point mutation** compared to the wild- type version of Gene A. Shown below is a portion of the fluorescence dideoxy- sequencing gel that gives the sequence of the **mRNA like strand/ non-template strand** of the DNA that corresponds to amino acids 5-7.

Fluorescent dideoxy- C
Fluorescent dideoxy- G
Fluorescent dideoxy- T
Fluorescent dideoxy- A



i. Write the sequence of the **mRNA** that corresponds to amino acids 5-7 of the mutant allele of Gene A.

5′AUGUAGAUG3′

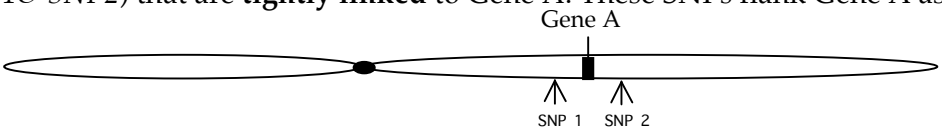
ii. If the wild –type allele has the amino acids **N- met[5]-trp[6]-met[7]-C, circle** the base in the sequence that you gave in part (i) that has undergone point mutation in the mutant allele. ***Note:*** *Codon chart is provided on the last page.* ***You can detach the last page.***

5′UGG3′ (coding for trp^6) is changed to 5′UAG3′ which is a stop codon

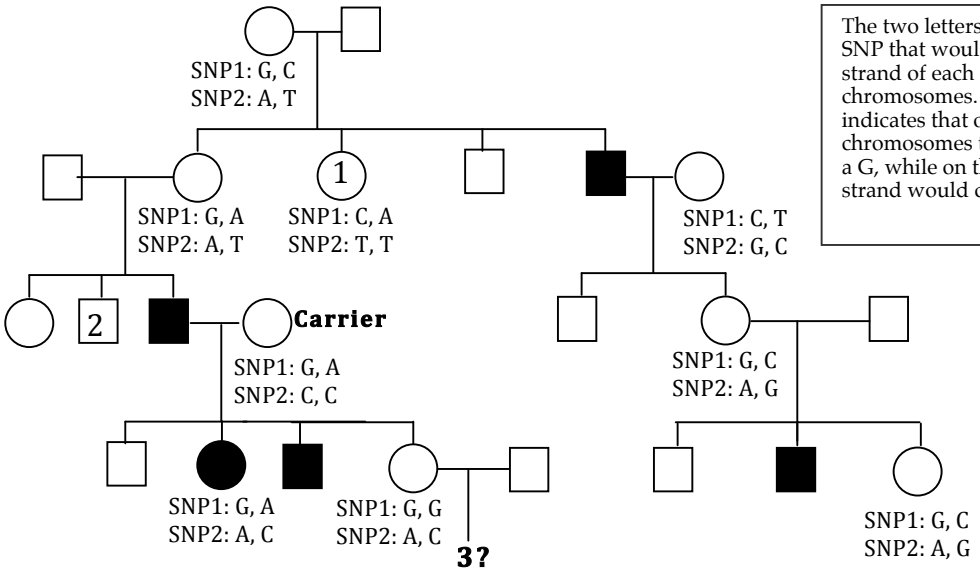
iii. **Circle** the type of mutation that you see in the mutant allele from the choices below (*choose from silent, missense,* ***nonsense*** *or frameshift*).

Question 5 (18 points)

You are studying a genetic disorder that is associated with a mutation at the Gene A locus. You identify two SNPs (*SNP1& SNP2*) that are **tightly linked** to Gene A. These SNPs flank Gene A as shown below.



Please note: *All the individuals with the disease phenotype are shaded. People marrying into the family have only the wild-type alleles of Gene A unless* ***indicated as a carrier****. Also listed are the alleles of SNP1 and SNP2 for some individuals. Assume* ***complete penetrance of the disease phenotype*** *and* ***NO Recombination*** *between SNP1, SNP2 and Gene A.*



The two letters identify the alleles of the SNP that would be found on the "top" strand of each of the two homologous chromosomes. For example, "SNP 1: G,A" indicates that on one of the homologous chromosomes the top strand would contain a G, while on the other chromosome the top strand would contain an A.

a) Give the most likely **mode of inheritance** of this disease. *X- linked recessive*

b) Give the **genotype(s)** of **Individual 1** at Gene A locus based on her SNP1 & SNP2. ***Note:*** *Use the letter "A" or X^A to represent the allele associated with the dominant phenotype and 'a" or X^a to represent the allele associated with the recessive phenotype.* ***X^AX^A***

c) Give the alleles of SNP1 & SNP2 that are tightly linked with the allele of Gene A in **Individual 2**.

SNP1: ***A or X^A or X^AY*** **SNP2:** ***T or X^T or X^TY***

d) If **Individual #3 is a male**, what is the probability that he will be affected? ***50%***

e) The above disease is also observed in mice. You mate an affected female with an affected male to get a fertilized ovum. You then successfully introduce a wild- type allele of Gene A into the fertilized ovum and implant it into a pseudo-pregnant female mouse. You observe that it develops into a **newborn male**.

i. What would be the genotype of **all** somatic cell-types in this newborn? ***AX^a or $X^{Aa}Y$***

ii. Will the introduced gene be passed on to the subsequent generations by the transgenic mouse (*choose from yes, no or may be*)? **Explain**.

The introduced transgene is stably integrated in the fertilized ovum. If integrates into an autosome it will be passed on to subsequent generation. But if it is introduced into the sex chromosome, then depending on whether the gametes receive the transgene it may or may not be passed on to subsequent generations.

f) Why two SNPs flanking a Gene are regarded as better markers to predict its mode of inheritance compared to only one SNP located at one end of the Gene?

This is because if the person inherits two SNPs that are tightly linked to and flank a gene, there is a much reduced chance of two recombination events. Or in other words the probability that the person has inherited the allele of the gene that is flanked between the two SNPs is much higher.

7.013 Introductory Biology
Spring 2013