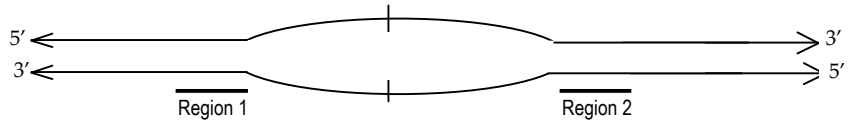
Solutions to 7.012 Quiz II
11/1/10

Class Ave = 56
Standard Dev = 14.5

Approximate grade	Range	%
A	74 - 100	11%
B	59 - 73	33%
C	44 - 58	38%
D	29 – 43	15%
F	0 – 28	3%

Question 1 (34 points)

The following segment of the DNA represents a replication bubble.



a) The DNA sequence in region 1 is:

```
5′ ACAGCGTC 3′  top
3′ TGTCGCAG 5′  bottom
```

Give the sequence of an 8 nucleotide RNA primer that would hybridize to region 1 and be elongated in a **continuous** manner. Label 5′ and 3′ ends.

5′ GACGCUGU 3′

b) The DNA sequence in region 2 is :

```
5′ ACAGCGTC 3′  top
3′ TGTCGCAG 5′  bottom
```

Give the sequence of an 8 nucleotide RNA primer that would hybridize to region 2 and be elongated in a **discontinuous** manner. Label 5′ and 3′ ends.

5′ GACGCUGU 3′

c) If DNA polymerase had a 50% decrease in its 5′ → 3′ polymerase activity, which of the following would you expect? Circle the **one** best answer.

A decrease in the speed of the lagging strand production.

A decrease in the speed of the leading strand production.

[A decrease in the speed of both lagging and leading strand production.]

A decrease in the accuracy of the lagging strand production.

A decrease in the accuracy of the leading strand production.

A decrease in the accuracy of both lagging and leading strand production.

d) If DNA polymerase had a 50% decrease in its 3′ → 5′ exonuclease activity, which of the following would you expect? Circle the **one** best answer.

A decrease in the speed of the lagging strand production.

A decrease in the speed of the leading strand production.

A decrease in the speed of both lagging and leading strand production.

A decrease in the accuracy of the lagging strand production.

A decrease in the accuracy of the leading strand production.

[A decrease in the accuracy of both lagging and leading strand production.]

e) The site on DNA to which proteins bind to initiate replication is called the

Origin of Replication

f) The site on DNA to which RNA polymerase binds to initiate transcription is called the

Promoter

Question 1, continued

g) In eukaryotes, the transcript initially produced is not recognized by the protein and RNA complex that carries out translation.

- Name the protein and RNA complex that carries out translation. *Ribosome*
- List three changes that are made to the initial eukaryotic transcript to produce a mature mRNA.
  addition of 7-methylguanosine (m7G) to the 5' end
  addition of a poly-A tail to the 3' end
  splicing to remove introns

h) If you know the sequence of a bacterial mRNA, you can correctly infer the primary structure of the resulting protein. If you know the primary structure of a bacterial protein, can you correctly infer the sequence of the mRNA that was translated? Explain your answer.
No, most amino acids are encoded by more than one codon, i.e., the genetic code is redundant.

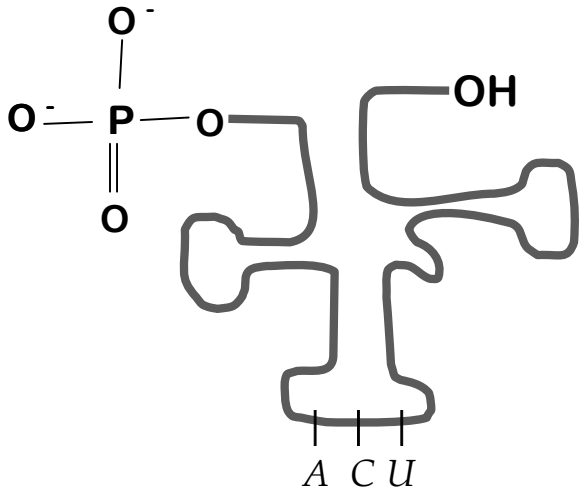
i) The following **double-stranded RNA** molecule was designed to use in a RNAi (RNA interference) experiment to decrease the function of gene X. The sequence represents the region of gene X that encodes the N-terminus of the protein.

```
            met
5' CUACUUCAAUGUGGACCAGUGGAGUGCCUUACAGAUCUGACGAAACGUCACGGAAUCUAUACUUUACAU 3'
3' GAUGAAGUUACACCUGGUCACCUCACGGAAUGUCUAGACUGCUUUGCAGUGCCUUAGAUAUGAAAUGUA 5'
```

- Give the sequence of the first 5 amino acids of the peptide produced from gene X. Label the N and C ends. *Please note: A codon chart is provided on the last page of this exam.*
  N Met-Trp-Thr-Ser-Gly C

- Below is a schematic of the molecule that will insert the fourth amino acid into the peptide encoded by gene X. Complete the diagram by giving the sequence of the anticodon.

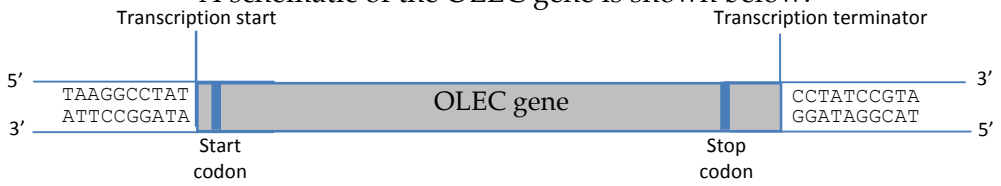
  This schematic represents a *tRNA* molecule.



Question 2 (30 points)
Plant lectins have been found to have anti-tumor properties. You have isolated a new lectin from the cytoplasm of orchid cells that may be useful as an anti-tumor agent. You hope to obtain large amounts of this lectin (called OLEC), by cloning the gene into yeast cells such that these cells produce and secrete OLEC into the media. To accomplish this you plan to:

Step 1) PCR the OLEC gene
Step 2) Attach DNA encoding a signal sequence onto the OLEC gene by cloning the OLEC gene into a vector that encodes a signal sequence.
Step 3) Transform yeast cells with the recombinant vector containing the signal sequence and the OLEC gene.
Step 4) Select for cells that have received a plasmid.

A schematic of the OLEC gene is shown below.



a) Give the DNA sequence of the 10 nucleotide primer or primers that would allow you to amplify the OLEC gene by polymerase chain reaction (PCR) technique. Label the 5′ and 3′ end(s).

5′ TAAGGCCTAT 3′ and 5′ TACGGATAGG 3′

b) You want to clone the PCR amplified OLEC gene into a vector that contains DNA encoding a signal sequence immediately following the promoter.

i) What is a signal sequence, and why is it important?
A signal sequence is a stretch of hydrophobic amino acids that is bound by SRP and specifies that this protein will be localized to the plasma membrane or secreted. The signal sequence directs the ribosome/mRNA complex to the ER where the protein is translocated across the membrane of the ER during translation.

ii) To accomplish the plan outlined above, the vector will contain a promoter from which organism?
Yeast

c) Your chosen vector also carries the wild type Arg 1^+ gene. The Arg 1^+ gene encodes an enzyme that allows yeast cells to synthesize arginine. Including this gene on the vector is important for the steps listed above.

i) Give the phenotype of the yeast cells (prior to transformation) that you will transform with your recombinant vectors. *These cells will be arg1- and unable to grow in the absence of arginine.*

ii) Onto what type of medium will you plate your transformation mix to distinguish between transformed and untransformed yeast cells? How will this medium allow you to identify transformed yeast cells? *Onto media lacking arginine*

d) Shown are the recognition sites for the restriction enzymes you have available. A vertical line (|) indicates exactly where each enzyme cuts.

Xba I:	*Nde* I:	*Sal* I:	*EcoR* I:	*Xho* I:
5′ T|CTAGA	5′ CA|TATG	5′ G|TCGAC	5′ G|AATTC	5′ C|TCGAG
3′ AGATC|T	3′ GTAT|AC	3′ CAGCT|G	3′ CTTAA|G	3′ GAGCT|C

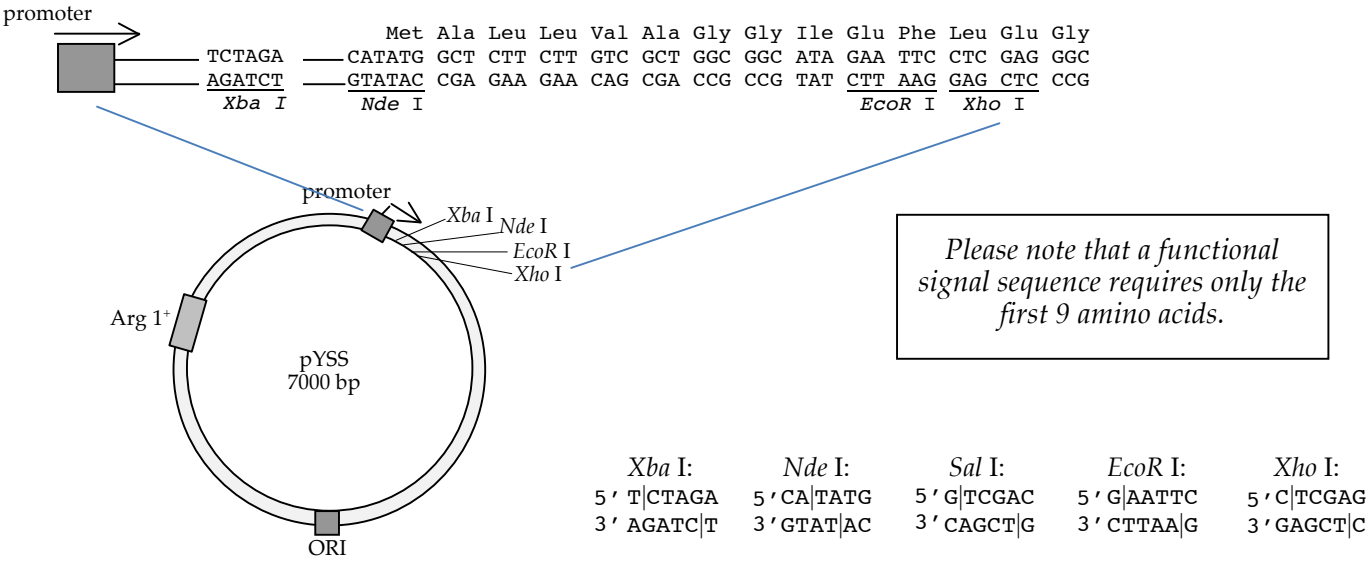
i)
Which two of these enzymes can cut to create ends that can be ligated to each other?
Sal I and Xho I

ii) Write the sequence that results from the ligation above **and** list all enzymes that can cut this sequence. Label 5′ and 3′ ends of all DNA molecules.

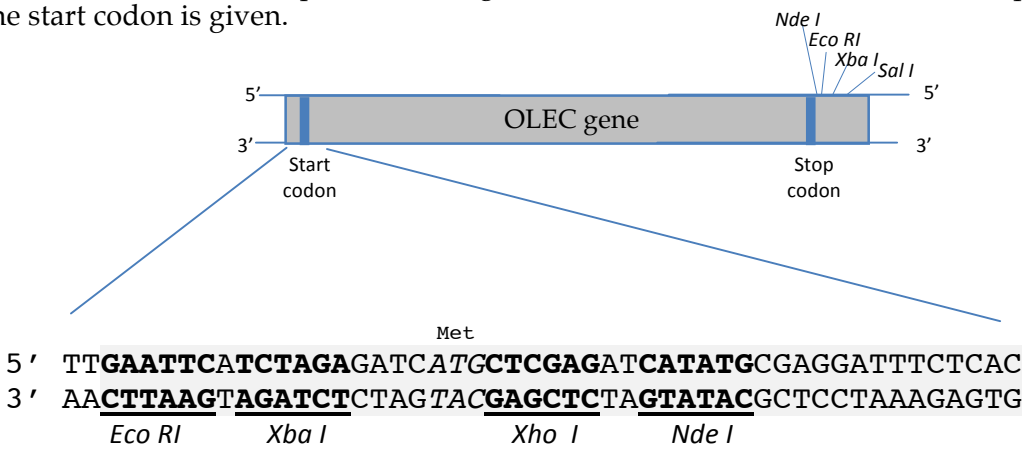
5′ GTCGAG 3′ *OR* *5′ CTCGAC 3′*
3′ CAGCTC 5′ *3′ GAGCTG 5′* *none of the above enzymes will cut this sequence*

Question 2, continued

A schematic of the cloning vector (including the signal sequence) is shown below. The sequence of the DNA encoding the signal sequence is given, including the first ATG after the promoter and the restriction enzyme sites.



A schematic of the PCR amplified OLEC gene is shown below, where the DNA sequence in the region of the start codon is given.



e) To clone the OLEC gene into the vector **such that the OLEC protein is produced and contains the signal sequence**…

i) What enzyme(s) would you use to cut the vector? *Xho I*

ii) What enzyme(s) would you use to cut the OLEC gene? *Xho I and Sal I*

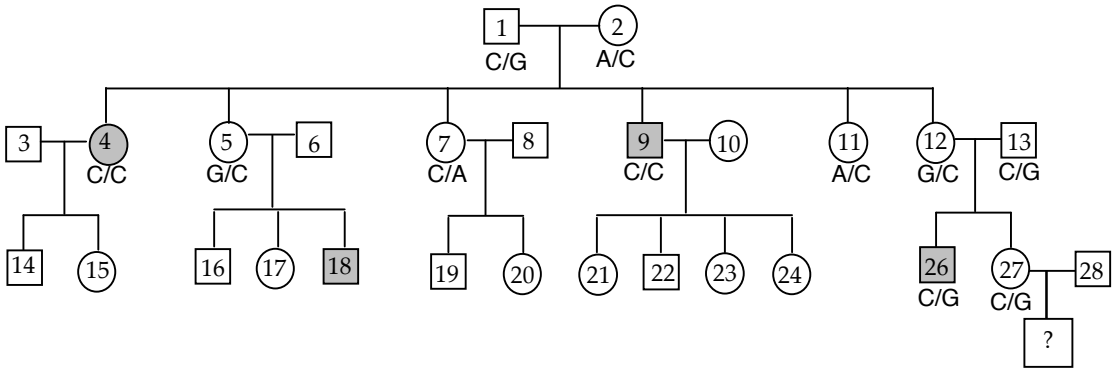
iii) When the OLEC gene is inserted in the correct orientation, what amino acids are encoded by the 12th -16th codons of the recombinant molecule?

Amino acid 12	Amino acid 13	Amino acid 14	Amino acid 15	Amino acid 16
Leu	*Glu*	*Ile*	*Ile*	*Cys*

Question 3 (19 points)

You are studying a newly characterized disease. You determine that SNP 1 and the disease locus are very tightly linked. The pedigree shows the pattern of inheritance of this disease. The individuals that are shaded show the disease phenotype, and all affected individuals are indicated.

a) What is the most likely mode of inheritance of this disease? *Autosomal recessive*



The two letters identify the two alleles of the SNP. For example, C/G indicates that on one of the chromosomes you would find a C (a C_{top} - G_{bottom} base pair), on the other chromosome you would find a G (a G_{top} - C_{bottom} base pair)

b) Also shown are the alleles of SNP 1 (for example C/G) for some individuals. Given the information provided by the pedigree, and assuming that no recombination occurs between the disease allele and SNP1, answer the following questions. Please note that the order of the SNP alleles is uninformative, *i*.e., C/A = A/C.

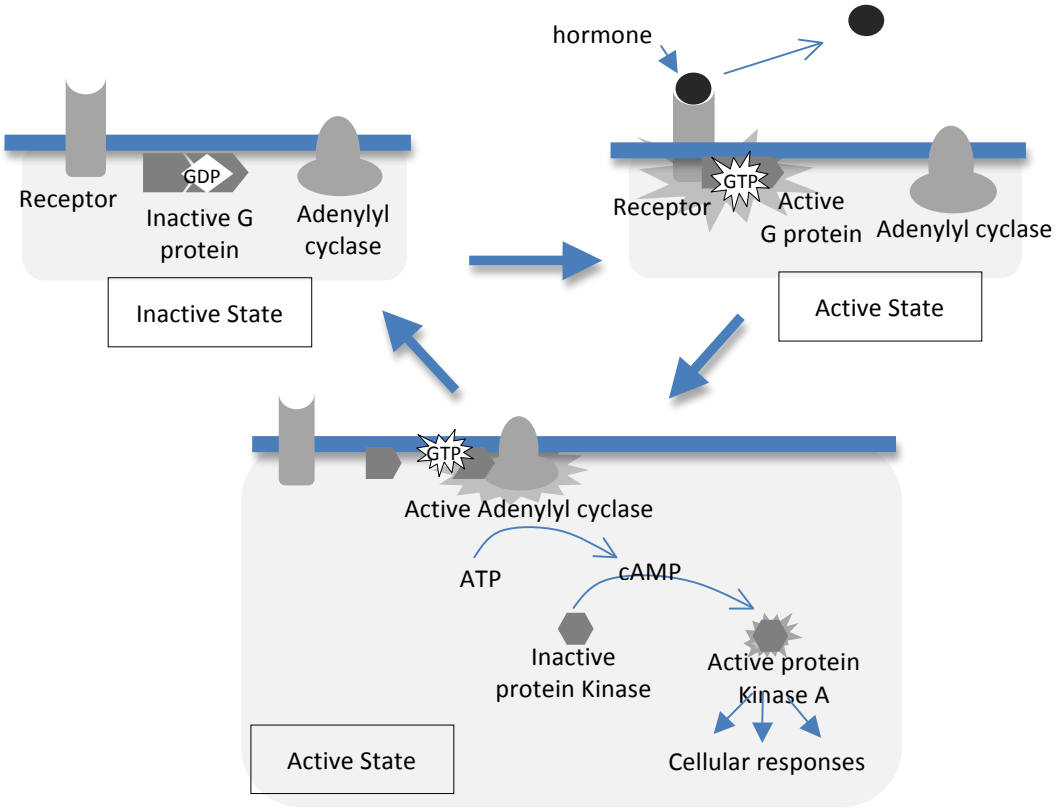
i) Assume that individual 8 is not a **carrier** of the disease allele. What is the chance that individual 19 will be a carrier of the disease allele?
50%

ii) Individuals 5 and 6 have an affected child. What is the genotype for individual 6 at the disease gene locus? Choose from DD, Dd, dd, or can't tell. *Dd*

iii) Individuals 5 and 6 have an affected child. What is the genotype for individual 6 at SNP 1? Choose from: A/A, A/C, A/G, C/C, G/C, G/G, or can't tell.
Because individual 6 is marrying into this family, there is no way to tell what alleles of the SNP he will have, all of the above are possible.

c) Assume that individual 28 is not a carrier of the disease allele. What is the chance that the offspring of individuals 27 and 28 is a **carrier** of the disease allele? H**ow did you use the SNP information to make this determination?**
0%. There are only two ways to have the C/G alleles of the SNP given the parents (#12 and #13).
1) One could get the C SNP from #12, which is linked to the d allele, and the G SNP from #13, which (given the phenotype of #26) is also linked to the d allele. In this case the genotype at the disease gene locus is dd, so this individual would have the disease.
2) One could get the C SNP from #13, which is linked to the D allele (given the phenotype of #26), and the G SNP from #12, which is also linked to the D allele. In this case the genotype at the disease gene locus is DD.
Given that #27 does not have the disease, we know she is DD and thus her child cannot be a carrier.

Question 4 (17 points)



Above is a schematic of a signaling cascade that is stimulated by a hormone. Hormone binds to the receptor, which activates the G protein. The activated G protein, in turn, activates Adenylyl cyclase. Activated Adenylyl cyclase stimulates the formation of cAMP from ATP. cAMP binds to and activates Protein Kinase A, which stimulates a kinase signaling cascade and results in the cellular responses.

a) The cellular responses indicated usually occur only in the presence of hormone. However, certain mutations can eliminate the need for hormone and increase the response, other mutations can cause an activation of the signaling pathway that also results in an increased cellular response. Put an **X** next to any of the following mutations that could result in an increase cellular response.

_X_ A mutation such that Adenylyl cyclase produces cAMP independent of the G protein
_____ A mutation such that the proteins after Protein Kinase A cannot be activated
_____ A mutation such that the receptors cannot bind hormone
_X_ A mutation such that G-protein fails to hydrolyze GTP
_X_ A mutation such that the receptor activates the G protein independent of hormone
_____ A mutation such that the protein on which Protein Kinase A acts cannot be phosphorylated.

b) Protein Kinase A adds phosphate groups to target proteins.

- Name the **three** amino acids in target proteins that have potential to be phosphorylated. Why can these be phosphorylated but the other 17 amino acids cannot?
  Ser, Thr, Tyr. These amino acids have an OH group that can accept the phosphate.
- What molecule donates the phosphate group to these amino acids?
  ATP

Question 4, continued

DNA microarray technology has many applications. It can be used to determine how gene expression differs between cell types or between treated and untreated cells. DNA Microarray technology can also be used to genotype individuals, (*i.e.,* determine which alleles are found at sites across the genome).

You design a DNA microarray that represents each gene in the genome of human cells. You use this microarray to determine how gene expression differs between hormone treated and untreated cells.

c) Would the array you designed for the above experiment be the same array that you would design to determine the genotype of two different humans?
No.

d) In the space given below…

- if your answer to (c) was yes, explain why the one array can be used for both experiments.
- if your answer to (c) was no, explain how the two DNA microarrays would be designed differently.

To use an array to genotype any individual or individuals, the array must contain short DNA sequences that are specific for each allele of the gene or genes (or each allele of the SNP or SNPs) that you are concerned about. In this case your hybridizing sample will be genomic DNA.

To compare gene expression between different cells, each position on the array would represent a conserved region of the gene. The sequence chosen to represent each gene must be coding sequence because your hybridizing sample will be mRNA or cDNA.