Question 1

Over the past 30 years, trends have shown an increase in the population of the cownose ray along the east coast of the United States. The cownose ray preys on bay scallops and the ray is a prey of great sharks. Below are three possible models for why there is an increase in cownose rays:

Models:

#1. An increase in cownose ray fecundity

#2. A decrease in the population of the great sharks

#3. An increase in the population of bay scallops

For each of the models, describe why that change would cause an increase in the population of the cownose ray

1. **An increase in cownose ray fecundity would mean that the rays are having more offspring per generation resulting in a larger population.**

2. **The great sharks are predators of the cownose ray. A decrease in the population of these animals means that the ray will not be preyed on as much, resulting in an increase in the population.**

3. **The cownose ray feeds on bay scallops. If bay scallops increase, there will be more food for the cownose ray, which will increase the fitness of the ray resulting in an increase in population.**

You decide to test model #1 by following a cohort of cownose rays for several years.

(a) What is a cohort?

A cohort is a group of the same species, all of the same age.

(b) To study the cownose ray population, you set up a large region in the ocean that is caged such that sharks can come and go, but the cownose rays remain within the caged area. Describe, in general terms, the technique you would use to determine $N_x$ each year.

You would use the mark and recapture technique. This would involve capturing a number of cownose rays, marking them and then releasing. At a subsequent date close to the date of the first capture, you would capture another group of cownose rays and determine how many of your second capture have the mark that you made.
(c) Below is a chart of your collected data:

<table>
<thead>
<tr>
<th>x</th>
<th>N_x</th>
<th>m_x</th>
<th>l_x</th>
<th>q_x</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>200</td>
<td>0</td>
<td>1</td>
<td>.1</td>
</tr>
<tr>
<td>1</td>
<td>180</td>
<td>0.44</td>
<td>0.9</td>
<td>.08</td>
</tr>
<tr>
<td>2</td>
<td>165</td>
<td>0.60</td>
<td>0.825</td>
<td>.39</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>0.25</td>
<td>0.5</td>
<td>.6</td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>0</td>
<td>.2</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

(i) Fill in l_x for each year, x.
(ii) Fill in q_x, the death rate, for each year, x.
(iii) How many cownose rays (to the nearest whole number) were born to the cohort when x = 1? Assume your N_x represents females only. Show your work.

m_x is a birth rate. At x = 1, m_x = 0.44. Therefore, each cownose ray gave birth to 0.44 offspring. This means that: 0.44 x 180 = 80 offspring born.

(iv) Based on R_0, can model #1 be correct? Show any necessary calculations. Explain your reason.

\[ R_0 = \sum l_x m_x \]
\[ R_0 = 1.016 \approx 1 \]
When \( R_0 = 1 \), the population is exactly replacing itself. Therefore, the cownose ray population is not increasing its fecundity to create an increase in the population.

(v) Based on the intrinsic rate of increase, r, for when x = 1, determine how long it would take the cohort to double its population. Show your work. (hint: \( r = b - d \)).

\[ r = b - d = 0.44 - 0.08 = 0.36 \text{ years}^{-1} \]
\[ T_d = \frac{\ln 2}{r} = \frac{0.69}{.36} = 1.9 \text{ years.} \]
(d) Which type of equation, exponential or logistic, more accurately describes the growth of a population? Provide a justification.

The logistic equation more accurately represents the growth of a population. The logistic equation includes the variable $K$, or the carrying capacity and models the intraspecific growth of the population. The intraspecific competition is a result of environmental pressures (space requirements, nutrient deprivation, etc.) that affect a population as it grows.

(e) You now want to determine if models #2 or #3 could explain the increase in the population of the cownose ray. To do this, you gather information about the populations over the past 30 years of the great sharks, the cownose ray and the bay scallops from several different sources. The results are printed on the next page.

(i) Based on the data, which model, #2 or #3, more likely explains the current increase in cownose rays? Explain.

The data show that the great shark populations begin to decline before the increase in cownose rays, which is consistent with model 2. The data also show an increase in bay scallops with a peak in population around 1980, prior to the increase in cownose ray, which is consistent with model #3. It is important to note that around 1990, cownose rays increase and both bay scallops and great sharks decline. If model #3 were correct, it would make sense that the bay scallops decrease when cownose rays increase because the rays eat the scallops. However, if model #3 were correct then the population of great sharks should start increasing due to the increasing amount of cownose rays. This does not happen. Therefore, the data are most consistent with model #2.

(ii) Based on the data, which of the organisms (Great Sharks, Cownose Ray or Bay Scallops) is a keystone predator? Justify your answer with the data provided.

A keystone predator is one that has a large impact on its environment relative to its size. In this case, the number of great sharks is much less than the cownose rays or the bay scallops. The loss of the great shark also had a large impact on the environment (increase in cownose rays and the decimation of the bay scallop population). Therefore, the great shark is the keystone predator.

(iii) If model #2 were correct, predict how this might drive evolution of bay scallops.

If model #2 is correct, then there will not be a predator of the cownose ray and their population will continue to grow unchallenged (by a predator – not by intraspecific competition!). This means that the bay scallops will have to evolve to prevent being eaten by the increased cownose ray population. Perhaps they will evolve to grow in new environments that cownose rays can not access, or develop a molecule that produces a taste cownose ray wouldn’t like or grow thicker shells to prevent being broken by rays.
Question 1 continued

(f) Below is an energy diagram that represents the simple food web of sharks, rays and scallops:

(i) When determining the biomass for primary producers in an oceanic system, the area surveyed includes the surface of the ocean down to approximately 200 meters. Why is biomass surveyed down to this depth?

*The biomass includes all organisms that have carbon (which is all organisms). Sunlight can penetrate the ocean water to about 200 meters, which means that photosynthetic organisms, the primary producers, can support ecosystems down to that depth. Below that depth, there is still much life to be counted as biomass, but we do not include those organisms because (1) – it is difficult to measure and/or survey organisms at those depths (2) – organisms existing below that depth represent a very small portion of the biomass in the ocean.*

(ii) Oceanic biomass pyramids are often partially “inverted”. Explain how this is possible.

*Oceanic biomass pyramids are inverted because the biomass of the primary producers is MUCH smaller compared to the biomass of the organisms that they support.*

(iii) Are the bay scallops, the cownose ray or the great sharks responsible for the inverted biomass pyramid?

*None of these animals are primary producers, so none are responsible.*
Question 1 continued

(iv) What type of organism belongs in trophic level A? Circle the most appropriate answer.

- heterotroph
- carnivore
- autotroph
- detritovore
- herbivore

(v) What type of organism belongs in trophic level E?

- heterotroph
- carnivore
- autotroph
- detritovore
- herbivore

(vi) Which trophic level, A, B, C, D or E, represents the cownose ray?

C

(vii) How would you determine the ecological efficiency of trophic level C? You can use any of the below variables:

- $B = \text{Biomass}$
- $I = \text{Ingestion}$
- $A = \text{Assimilation}$
- $F = \text{Fecal matter}$
- $R = \text{Respiration}$
- $P = \text{Production}$

To indicate which trophic level you want to represent, use a subscript. Example:

$P_A = \text{production of trophic level A}$.

\[
\frac{I_C}{P_A} \times \frac{A_C}{I_C} \times \frac{P_C}{A_C} = \text{Ecological Efficiency}
\]
Question 2

You are spending your summer working for an agricultural company developing new plants that are resistant to a well known weed killer, Roundup. Plants that are genetically modified to be resistant to Roundup are known as “Roundup Ready” and include soy beans, corn, rice, sorghum and cotton. Your company wants to make Roundup Ready carrots (carrots\textsuperscript{RR}).

The common name for Roundup is glyphosphate. This molecule inhibits a metabolic pathway necessary for creating aromatic amino acids (tryptophan, tyrosine and phenylalanine). To make a plant Roundup Ready, the plant needs to express a particular gene (geneR) that encodes for a protein that prevents the glyphosphate from affecting plants.

Your company already generated Roundup Ready corn, so you have a copy of eukaryotic geneR in a plasmid for expressing geneR in corn. Below is a picture of the plasmid, pCRN. The arrow represents a corn promoter.

(a) Your advisor tells you that you need to move the geneR coding sequence to a new plasmid to express geneR in carrots. Explain two features that would be different between the plasmid shown above, pCRN, and the new plasmid used in carrot cells.

1. Change the promoter to a carrot-specific promoter
2. Change the origin to a carrot-specific origin
3. You don’t know what the selection marker is. It might work in carrots, but you might need to change it.

Do not write below this line.
(b) Below is a diagram of the region surrounding the geneR coding region in the corn plasmid and a diagram of the multi-cloning site of the new plasmid that has been designed to be used in carrots:

![Diagram of geneR coding sequence and multi-cloning site]

You have a chart of the DNA sequences recognized by each of the restriction enzymes. Triangles represent where the DNA is cut:

- **ClaI**: 5’-ATCGAT-3’
  
  3’-TAGCTA-5’

- **SmaI**: 5’-CCCGGG-3’
  
  3’-GGGCCC-5’

- **XbaI**: 5’-TCTAGA-3’
  
  3’-AGATCT-5’

- **PvuII**: 5’-CAGCTG-3’
  
  3’-GTCGAC-5’

- **EcoRI**: 5’-GAATTC-3’
  
  3’-CTTAAG-5’

- **KpnI**: 5’-GGTACC-3’
  
  3’-CCATGG-5’

You get advice from three people in your lab about which enzyme(s) you should use to move the geneX coding sequence and maintain the correct orientation of the gene with respect to the promoter. Choose the best possible recommendation. For each recommendation, indicate whether or not you would choose those enzyme(s) and support your choice.

**#1 ClaI and EcoRI**

Yes or No? Yes

Why or why not?

_EcoRI and ClaI both result in overhangs that will help the geneX coding sequence DNA fragment insert in only one orientation into the new vector. Additionally, ClaI and EcoRI are in the same order in the new vector as they are in the corn vector, which will maintain the orientation of the coding sequence._

**#2 SmaI and PvuII**

Yes or No? No

Why or why not?

_SmaI and PvuII are in the correct order in the new vector as in the corn vector, but they do not result in overhangs. They both result in blunt cut sites, which will not allow for specific, directional annealing. A PvuII-cut piece of DNA could anneal to a SmaI-cut piece of DNA or any other piece of DNA that had been cut with a restriction enzyme that resulted in a blunt cut site._

**#3 ClaI and KpnI**

Yes or No? No

Why or why not?

_ClaI and KpnI both result in overhangs to help the coding sequence fragment insert into the new vector. The cut sites for these two enzymes are not, however, in the same order in the two vectors. In the new vector, the KpnI cut site comes before the ClaI site, so the fragment would be inserted in a backwards orientation with respect to the promoter._
You put your new plasmid into the carrot plant. You test if your carrots\textsuperscript{RR} are resistant to Roundup. To test for resistance, you grow carrots\textsuperscript{RR} in two plots: one with Roundup and one without. Below are your data:

(c) Under which condition do carrots\textsuperscript{RR} show greater intraspecific competition? Use the data from the graph to justify your answer.

Carrots\textsuperscript{RR} show great intraspecific competition when in the absence of Roundup than when in the presence of Roundup. This is indicated by the slope of the growth curve after the initial planting. If growth is not affected, then the carrots should grow proportionally to the increasing amount of seeds. The slope of the curve in the presence of Roundup is much steeper than the slope in the absence of Roundup. After planting the first seed density in the absence of Roundup, carrots\textsuperscript{RR} increase drastically, but plantings with increasing seed densities show little increase in net yield. In the presence of Roundup, carrots\textsuperscript{RR} do not grow as much initially, but this prevents some of the intraspecific competition, so net yield increases more proportionally to seed density.

**BONUS (1pt)** Your advisor says that she is not convinced by your data that carrots\textsuperscript{RR} are Roundup Ready. What other data do you need to collect to show that carrots\textsuperscript{RR} are Roundup Ready?

*To know for sure, you would need to grow normal carrots in the presence of Roundup to see how well they grow without the added advantage of geneR.*
Question 2 continued

(e) One of the concerns of genetically modified foods is how they will interact with the community. Will creating genetically modified carrots result in a carrot that grows differently than non-genetically modified carrots?

To answer this question, you grow normal carrots in the same plot as carrots \(^{RR}\) in the absence of Roundup. You also grow normal carrots and carrots \(^{RR}\) alone in the absence of Roundup.

If normal carrots showed interspecific competition, how would normal carrots grow in the same plot as carrots \(^{RR}\) as compared to how normal carrots grow in a plot by themselves?

*If normal carrots showed interspecific competition, you would expect normal carrots to grow worse in the presence of carrots \(^{RR}\) as compared to when grown by themselves.*

(f) It turns out that there is no interspecific competition! You hypothesize that the lack of interspecific competition is a result of the two types of plants having different nutrient requirements, and therefore, they do not occupy the same exact niches.

(i) If you were going to describe three dimensions of the niche space for carrots and carrots \(^{RR}\) what three might you choose?

Many answers: sunlight, nitrogen availability, moisture content, phosphorous availability, space, potassium availability, predators (animals that eat carrots), etc.

(ii) To test your hypothesis, you analyze the relative amounts of carbon, phosphorous, nitrogen and potassium (C, P, N, K) in each of the types of carrots.

<table>
<thead>
<tr>
<th></th>
<th>Carrots</th>
<th>Carrots (^{RR})</th>
</tr>
</thead>
</table>

Based on these data, indicate the likely soil composition used in the experiment in part (e) that allowed both plants to grow. Circle your choice. Ratios are C:P:N:K ppm.


(g) Based on your answer for part (f):

(i) For which resource is the carrot limited? \(P\)

(ii) For which resource is the carrot \(^{RR}\) limited? \(N\)
Question 3

The three major processes of the Central Dogma are DNA replication, transcription and translation. For each of these processes, circle the stage of the cell cycle during which this process occurs.

<table>
<thead>
<tr>
<th>Process</th>
<th>G1</th>
<th>G2</th>
<th>S</th>
<th>M</th>
<th>All</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA replication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transcription</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Translation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(b) DNA is a very special molecule. Below is a diagram of one of the four monomers that make up DNA:

(i) Circle a part of the monomer that takes part in hydrogen bonding to another monomer.
(ii) Draw a square around a part of the monomer that takes part in covalent bonding to another monomer.
(iii) Draw a triangle around a part of the monomer that could take part in ionic bonding.
(iv) Rank each of the bonds in order of increasing strength.

hydrogen bond, ionic bond, covalent bond

(c) Many experiments were important in discovering the structure, function and maintenance of DNA. Listed below are three of these experiments. Which property of DNA did each experiment (described below) help elucidate?

(i) Erwin Chargaff showed that all the organisms he tested had the same amount of adenine as thymine and the same amount of guanine as cytosine.  
*This experiment helped show that guanine pairs with cytosine and adenine pairs with thymine. Helped lead to notion of a double stranded, paired molecule.*

(ii) Meselson and Stahl showed that dense DNA (N\textsuperscript{15} substituted for N\textsuperscript{14} in the bases) became a density between dense DNA and normal DNA after one round of replication.

*Semi-conservative DNA replication*
Question 3 continued

(e) Below is a diagram of a DNA replication fork:

(i) Label the 5’ and 3’ ends of both strands
(ii) DNA polymerase requires a primer to be associated with the DNA before it can begin adding nucleotides. Draw the appropriate primer(s) on each template strand. Label the 5’ and 3’ ends of your primers.
(iii) Indicate the direction of replication on each strand by drawing an arrow.

BONUS (1pt): Fidelity is the property of DNA replication that keeps incorrect base-pairing during DNA replication to a minimum. How does DNA replication play a role in evolution?

Despite very good fidelity, some mistakes are not fixed during or after DNA replication. An incorrect nucleotide insertion will be passed on to one of the daughter cells (the other daughter will receive the correct base-pair). If the nucleotide change is within a gene, it has the possibility of altering the gene product. The altered gene product might affect the fitness of the organism in some way such that the nucleotide change is selected for (increase in fitness) or selected against (decrease in fitness).

Question 4

Recent data have shown that the neurological disease epilepsy can be treated by administering the drug 2-deoxy-glucose (2DG), which is an inhibitor of glycolysis. 2DG is an analog of glucose.

(a) Glucose is a subunit of what type of macromolecule?
   carbohydrate, sugar, polysaccharide

(b) This type of macromolecule is generally:
   hydrophobic          negatively charged          polar          positively charged
(c) Below is a schematic of glucose bound by hexokinase, the enzyme that performs the first step in glycolysis:

What type of chemical interaction could occur between glucose and the Serine amino acid in the binding pocket above? Circle all that apply.

- ionic bond
- non-polar
- hydrogen-bond
- covalent bond
- polar
- hydrophobic
- van Der Waals

**BONUS (1pt)** Below is the structure of 2DG. Based on the structure of 2DG and the binding pocket of hexokinase, predict if 2DG inhibits the first step in glycolysis. **Justify** your reason. You may assume that 2DG associates with hexokinase in the same manner that glucose does.

**Explanation:**

2DG has a very similar structure to glucose, except that it is missing the 2'–OH. This hydroxyl group is not involved in interacting with hexokinase, so we would predict the 2DG does not inhibit the first step in glycolysis. In fact, 2DG inhibits the second step in glycolysis.
(d) 2DG is also an anti-tumor drug. Many cells in tumors are in a hypoxic (little O\_2) environment. Why would inhibiting glycolysis be lethal for these hypoxic cells? Be specific in your explanation.

*Human cells use oxidative phosphorylation to produce ATP as an energy source. In this process, O\_2 acts as the final electron acceptor for the electron chain, which is important for creating the H^+-gradient necessary to drive ATPase. In the absence of O\_2, human cells can still function, but now they only receive ATP from glycolysis (which produces 2 net ATP molecules per glucose). Addition of 2DG inhibits glycolysis, which results in NO ATP being made. Thus, the cells die.*

You want to know how inhibiting glycolysis can prevent epilepsy. Before you can understand this mechanism, you need to identify what genes are affected in epileptic patients. Your lab has a mouse model of epilepsy, meaning that you have mice with a mutation that results in epileptic behavior. The epileptic phenotype is *recessive.*

You are going to use known “markers” in the mouse genome to determine on which chromosome you find the mutated gene. A “marker” is an allele of a gene whose location in the genome is known. You have two markers: an allele of gene B that results in brown fur and allele of gene R that results in red eyes. Both of these phenotypes are recessive to white fur and black eyes, respectively.

Use the following notation for alleles: E/e represents alleles for non-epileptic/epileptic mice, F/f represent alleles for white/brown fur and B/b represents alleles for black/red eyes.

(e) You cross a true-breeding epileptic mouse with a true-breeding brown fur mouse or a true-breeding red-eyed mouse. Indicate the genotypes for each of the parental strains as well as the F\textsubscript{1} strain. Also indicate the phenotype of the F\textsubscript{1} mice.

\begin{align*}
\text{CROSS 1} & & \text{CROSS 2} \\
\text{epileptic } eeFF & \times \text{ brown fur } EEff & \text{epileptic } eeBB & \times \text{ red-eyed } EEbb \\
\downarrow & & \downarrow \\
\text{F\textsubscript{1} phenotype } \text{wt} & \text{ genotype } EeFf & \text{F\textsubscript{1} phenotype } \text{wt} & \text{ genotype } EeBb
\end{align*}

(f) You then cross the F\textsubscript{1} mice from Cross 1 and Cross 2 with the appropriate recessive test strain. Indicate the genotype of the two test strains:

(i) genotype of test strain for F\textsubscript{1} mice from Cross 1 ___eef

(ii) genotype of test strain for F\textsubscript{1} mice from Cross 2 ___eebb
Question 4 continued

(iii) What are the possible genotypes (with respect to epilepsy and brown fur) of the four gametes produced by the F1 mice from **Cross 1**?

\[ Ef, EF, ef, eF \]

(iv) You get the following results (the number indicated number of offspring with that phenotype):

<table>
<thead>
<tr>
<th>Cross 1 F1 mice x test strain</th>
<th>Cross 2 F1 mice x test strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>non-epileptic, white – 17</td>
<td>non-epileptic, red – 28</td>
</tr>
<tr>
<td>non-epileptic, brown – 20</td>
<td>non-epileptic, black – 3</td>
</tr>
<tr>
<td>epileptic, white – 19</td>
<td>epileptic, red – 4</td>
</tr>
<tr>
<td>epileptic, brown – 20</td>
<td>epileptic, black – 26</td>
</tr>
</tbody>
</table>

Based on these numbers, which marker (the gene for red eyes or brown fur) is on the same chromosome as the mutated gene of interest? Explain your reasoning based on the data.

*The offspring from the cross with Cross 1 F1 mice are present in a 1:1:1:1 ratio for the possible phenotypes. A 1:1:1:1 ratio in this cross indicates that the two genes responsible for these phenotypes are unlinked. The offspring from the cross with Cross 2 F1 mice mostly display the parental phenotypes and a few show non-parental. This result indicates that the two genes are linked, but distant enough on the chromosome for a cross-over event to occur. Therefore, the mutant epileptic gene is on the same chromosome as the gene that encodes for eye color.*

(v) Circle the two classes of mice from the “Cross 1 F1 mice x test strain” cross that display the parental phenotype in part (iv) above.

(g) After many crosses with other markers, you identify the location of the mutated gene. You want to determine the sequence of the gene to determine the mutation.

(i) When you perform your sequencing, should you include NTPs, dNTPS or ddNTPS or a combination?

*`dNTPs and ddNTPs`*

(ii) What is the purpose of your choice(s)?

*dNTPS are for normal polymerization. Due to a missing hydroxyl on the 3’ C, addition of a ddNTP would inhibit DNA replication.*

(h) You find that the mutation is a single nucleotide change in the coding region of the gene. You want to study the effect this mutation has on the protein product. You sequenced the mutated gene from DNA isolated from skin cells and you initially plan to use these same skin cells to isolate the protein. However, epilepsy is a neurological disorder. Why must you use neurons instead of skin cells if you want to isolate protein?

*Because epilepsy is a neurological disorder, whatever is happening to cause epilepsy is affecting only neurons. One reason why a mutated gene would only affect one type of cell is because that gene is ONLY expressed in that type of cell. With this reasoning, the mutated gene product will only be found in neurons because those are the cells that express the protein. The DNA for sequencing could be isolated from skin cells because all cells have the same DNA, but not all express the same proteins.*
Question 4 continued

(i) You are able to isolate your protein of interest. Analysis shows that the protein is only one-half the size it should be! Explain how a single nucleotide substitution could result in a smaller protein.

There are three types of mutations that arise because of nucleotide substitutions:

1. **silent** – the nucleotide substitution does not change the amino acid for which the codon corresponds
2. **missense** – the nucleotide substitution results in an amino acid change
3. **nonsense** – the nucleotide substitution results in a stop codon.

If the nucleotide change is in the middle of the gene and the mutation results in a change to a stop codon, the mutant protein would be 50% as large as the wild-type protein.

Question 5

You are working in a lab that studies microbes isolated from deep-sea vents. You are working on a particular species called *Thermus chisholmalia* (named after your favorite ecology professor, of course!). *T. chisholmalia* has only recently been discovered.

(a) Explain how you would determine if *T. chisholmalia* is autotrophic or heterotrophic. Autotrophs are capable of producing their own energy (glucose) while heterotrophs are not. There are also different types of autotrophs: those that do photosynthesis and those that do chemosynthesis. You should grow *T. chisholmalia* under different conditions that test these properties. First grow in liquid media with light in a closed system that contains oxygen, nitrogen, phosphorous, all the salts in seawater, glucose and trace minerals. All aerobic organisms should grow in these conditions. Then grow in the same conditions but without glucose. *T. chisholmalia* will only survive if the species is an autotroph. You should also test whether *T. chisholmalia* can grow in the absence of glucose AND light. If the species grows, it is a chemotroph. If it does not grow, it is a phototroph.

(b) You determine that *T. chisholmalia* is an autotrophic microbe. Which of the following would *T. chisholmalia* use as an energy source for its autotrophic process? Circle the appropriate choice.

- SO$_4^{2-}$
- sunlight
- ATP
- H$_2$S
- NO$_3^-$
- H$_2$O
- chlorophyll

(c) CO$_2$ is one of the substrates for the autotrophic process that *T. chisholmalia* undergoes. Is CO$_2$ an oxidizing agent or a reducing agent?

oxidizing reagent

(d) You are interested in isolating genes from *T. chisholmalia* that are important for allowing the microbe to exist at high temperatures.

Below is an outline of the experiment that you do to identify mutant *T. chisholmalia* that can no longer live at high temperatures. Fill in what you should do for Step 4.

1. Grow liquid culture and mutagenize cells
2. Plate mutagenized cells on media
3. Replica-plate cells
4. **grow at high temperature**
5. Identify mutants by comparing original plates and replica-plated plates.
(d) Using your method above, you isolate 6 mutants which you call M1 – M6. You perform a complementation test between each of your mutants. Below are your results (“+” means growth, “-” means no growth):

<table>
<thead>
<tr>
<th></th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
<th>M5</th>
<th>M6</th>
<th>WT</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>M2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>M3</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>M4</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>M5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>M6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>WT</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

(i) Does a complementation test indicate if two mutants have the same exact mutation? Explain.

No. Cells in the same complementation group have mutations in the same genes, but not necessarily the same mutation.

(ii) How many different genes did you identify as being important for allowing the microbe to exist at high temperatures? Indicate which mutants have mutations in the same genes.

2 genes:
- Complementation Group 1 – M1, M4
- Complementation Group 2 – M3, M5, M6

(iii) Give two reasons why there might be more genes involved than what you identified.

1. One of the mutants, M2, is dominant, so it might be in complementation group 1 or 2 or its OWN group. We can not tell.
2. The mutagenesis might not have created a mutation in every gene responsible for high temperature sustainability.
3. The mutation might be lethal

You perform the mutagenesis again and obtain even more mutants, called M7-M46. You perform more complementation tests and determine that two of these mutants, M18 and M32, have a mutation in the same gene and both have recessive phenotypes.

(e) Next you want to create a genomic DNA library of T. chishomalia DNA to clone the genes that are mutated in M18 and M32. You plan to clone by complementation. Should you make a library by isolating DNA from the wild-type T. chisholmialia or from one of the mutants? Explain.

Cloning by complementation requires that you identify your gene by transforming your library into a strain that has a mutant phenotype and identifying those plasmids in your library that can return the strain to a wild-type phenotype. Therefore, you would need to make your library out of wild-type DNA because the mutant DNA will not complement that mutant phenotype.
(f) You create your library by isolating DNA and then digesting the DNA with the restriction enzyme EcoRI. Your advisor suggests that you make another library by isolating more DNA and then digesting with KpnI. (See Question 2b for recognition sequences, if needed)

(i) Explain why it would be helpful to have multiple genomic DNA libraries each created by digesting DNA with different enzyme.

Initially, you do not know the sequence of your gene of interest and you do not know if the gene sequence contains a sequence recognized by the restriction enzyme you choose to digest your library DNA. If your gene does have that sequence, then your gene will be cut by the restriction enzyme and you will never recover the gene (it can’t be expressed if it is cut in two pieces!). Therefore, you should use more than one enzyme to avoid this possibility.

(ii) Why do you not need restriction enzymes to make DNA fragments when you create a cDNA library?

Because cDNA libraries are generated by isolating mRNA, which is reverse transcribed into DNA. This process results in cDNA fragments that are an appropriate length to be put inside a plasmid.

(h) You have cloned the gene that is mutated in M18 and M32, and you call it GeneTC3. You determined both the wild type and mutant sequences. Below is the wild-type sequence. Transcription begins at and includes the underlined, bolded base pair. The underlined sequences represent the promoter, and transcription proceeds to the right from the promoter.

```
5’-TAACCATGCTATGCAATTTTTGCAACATTATAAGAATTCTGACGAAGTTTTTTTACTTG-3’
1 --------------------------+---------------------------------+---------------------------+ 70
3’-ATGGGTAGAAGCTGCCACCAAGTTGTTGAATTTCTCTAACATGGTACTGTCCTGTTCAAAAAATGAAC-5’

5’-GTAGGATTATTTATCCTGGGCAGGCTGGCCCTGGAGTACCATGAGGCGATGCTCTGGCTCTTTGAAAATCGAGGG-3’
71 ---------------------------------------------------------------d-+-------------------+ 140
3’-CATCCTAAATAAGCTGGCTACCAACGGGGACCTACATGTCCTGCTACGTAACGGAAGCCAATCTAGTTTCC-5’

5’-CCGGGTCTCCTCTGACGGCCACCCACCTGGGCTGCCAACCTTCTGGCCCTGAAGGGCCCTACCCAGGAGC-3’
141- ---+-------------------------------+-----------------+------------ 210
3’-GGCCCGGAGGACTGCGGCTGGTGGAGCGGAGGATGGCTGCAGGGGCTGAGCCGCTCCGAGTGGTGGCTTG-5’
```

(i) What are the first 10 nucleotides of the mRNA that results from transcription of this gene? Make sure to label the ends.

```
5’GGCAGGUGGCC 3’
```

(ii) Draw a box around the nucleotides in the sequence above that encode the translation start site.
Question 5 continued

The site of each of the mutations is marked with a letter. The position of the mutation in M18 is marked with the letter “a” and the position of the mutation in M32 is marked with a “b”.

(i) The mutation at a is a change from a C:G (top:bottom) pair to a G:C pair.

(i) How does this change affect the primary level of protein structure?
Changes an arginine amino acid to a glycine amino acid.

(ii) How might this change affect the tertiary level of protein structure? Specifically discuss how the change in the primary structure would have an effect.
glycine is a small amino acid (and can act as polar or non-polar) and arginine is a large, positively-charged amino acid. This change could affect the structure by removing a positive charge, which might have been involved in an ionic bond necessary for the 3-dimensional structure of the protein. The arginine amino acid might also be involved in a hydrogen bond. While the glycine might be able to be involved in a hydrogen bond, it might not be large enough to contact the necessary bonding partner.

(iii) What if the tertiary structure is not affected at all? How could this mutation affect the activity of the protein?
It is possible that the arginine amino acid is not important for structure, but is important for the activity of the protein (DNA binding, enzymatic activity, RNA binding etc.). The loss of the positive charge could then result in a loss in activity.

(j) The mutation at b is a deletion of that nucleotide pair. You analyze the mRNA transcript produced in this mutant strain as well as the protein translated from that mRNA transcript. The length of the mRNA transcript from M32 for this gene is identical (except for one nucleotide) to the transcript from the wild-type T. chisholmalia. However, the mutant protein product is shorter than the wild-type protein by 13 amino acids. The protein product is entirely non-functional. How does the deletion of a single nucleotide pair result in the above situation?
The loss of a nucleotide does not affect the length of the transcript because RNA polymerase will continue making an RNA transcript until the transcription terminator is reached. The ribosome will continue adding amino acids to create a protein until a stop codon is reached. A deletion of a nucleotide will change the reading frame of the ribosome. A stop codon to be present earlier in the mutant transcript than it should be in the wild-type transcript as a result of the reading frame shift. If we sequenced the mutant protein, most of the amino acids after the deletion should also be different than in the wild-type protein.

(k) You think that other microbial species from the deep-sea vent would have GeneTC3.

(i) Is GeneTC3 appropriate for determining the evolutionary relationships between microbes around MIT? Explain.

No, GeneTC1 would not be appropriate. Microbes around MIT do not have to sustain themselves in such high heat, so it is less likely that they will have evolved to have genes necessary to exist at high temperatures.
ii) You sequence a portion of GeneTC3 from four other microbes from the deep-sea vent. Below are your results for the sequence of one of the strands of DNA:

<table>
<thead>
<tr>
<th>Microbe</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>T. chisholmalia</td>
<td>5’ – CCACCTGCCCTACCCG – 3’</td>
</tr>
<tr>
<td>T. aquaticus</td>
<td>5’ – CCACCTGCTCTTCCCG – 3’</td>
</tr>
<tr>
<td>T. thermophilus</td>
<td>5’ – CCGCCTGGCTACCCG – 3’</td>
</tr>
<tr>
<td>G. uralicus</td>
<td>5’ – TCACTGCTCCTCCG – 3’</td>
</tr>
<tr>
<td>P. furiosus</td>
<td>5’ – CCGCCTGGCCTACCTT – 3’</td>
</tr>
</tbody>
</table>

- Which species is the most evolutionarily distant from T. chisholmalia?

The most distant is G. uralicus.

- How did you determine your answer above?

To determine evolutionary distance, you must determine the number of nucleotides that are different between your sequence of interest and the other sequences. The larger the number, the more evolutionarily distant.

(i) You scan the genome of a microbe that grows in the dirt outside 54-100, M. institutous. You discover that M. institutous has a copy of GeneTC3! GeneTC3 in T. chisholmalia is expressed all of the time. In M. institutous, however, GeneTC3 is in an operon and is regulated. Below is a diagram of the operon from M. institutous.

![Operon Diagram]

Your lab does some analysis on the operon and determines that the regulating protein is an activator, GeneX.

(i) Based on the role of GeneTC3 in T. Chisholmalia, under what conditions would you expect genes TC3 and TC4 to be expressed?

You would expect TC1 and TC2 to be expressed when M. insitutous is exposed to extremely high temperatures.

(ii) Where, on the diagram above, does the activator bind?

The activator binds to the O, or operator, region.

(iii) Under conditions when the operon is expressed, does the activator bind to this segment of DNA?

It binds in the presence of the inducer.
HERE IS WHAT YOU DIDN’T HAVE TO DO ON THE EXAM.

(iv) Given the following diploid genotypes, indicate whether TC3 and/or TC4 would be expressed (+ means wild-type, - means loss-of-function mutation, and each line represents a chromosome):

<table>
<thead>
<tr>
<th>Genotype</th>
<th>TC3</th>
<th>TC4</th>
</tr>
</thead>
<tbody>
<tr>
<td>GeneX$^+$ P$<em>x^+$ P$</em>{TC}^+$ O$^+$ TC1$^-$ TC2$^*$</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>GeneX$^+$ P$<em>x^+$ P$</em>{TC}^+$ O$^+$ TC1$^+$ TC2$^+$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GeneX$^+$ P$<em>x^+$ P$</em>{TC}^+$ O$^+$ TC1$^-$ TC2$^*$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GeneX$^+$ P$<em>x^+$ P$</em>{TC}^+$ O$^+$ TC1$^+$ TC2$^*$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GeneX$^+$ P$<em>x^+$ P$</em>{TC}^+$ O$^+$ TC1$^+$ TC2$^*$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GeneX$^+$ P$<em>x^+$ P$</em>{TC}^+$ O$^+$ TC1$^+$ TC2$^*$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
STRUCTURES OF AMINO ACIDS

GENERIC AMINO ACID:

Individual amino acids are linked through these groups to form the backbone of the protein. Peptide bonds

Protein Synthesis

Side chain, unique to each different amino acid

Peptide bonds

GENERIC AMINO ACID:

- ALANINE (ala) **A**
- ARGinine (arg) **R**
- ASPARAGINE (asN) **N**
- ASPARTIC ACID (asp) **D**
- CYSTEINE (cys) **C**
- GLUTAMIC ACID (glu) **E**
- GLUTAMINE (glN) **Q**
- GLYCINE (gly) **G**
- HISTIDINE (his) **H**
- Isoleucine (ile) **I**
- LEucine (leu) **L**
- LYSINE (lys) **K**
- METHionine (met) **M**
- PHENYLALANINE (phe) **F**
- PROline (pro) **P**
- SERine (ser) **S**
- THREONINE (thr) **T**
- TYROSINE (tyr) **Y**
- VALINE (val) **V**

Individual amino acids are linked through these groups to form the backbone of the protein.
<table>
<thead>
<tr>
<th>First letter</th>
<th>Second letter</th>
<th>Third letter</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>U</strong></td>
<td>UUU ⟷ Phe</td>
<td>UCA ⟷ Ser</td>
</tr>
<tr>
<td></td>
<td>UUC ⟷ Leu</td>
<td>UCC ⟷ Stop p</td>
</tr>
<tr>
<td></td>
<td>UUA ⟷ Leu</td>
<td>UCA ⟷ Stop p</td>
</tr>
<tr>
<td></td>
<td>UUG ⟷ Leu</td>
<td>UCG ⟷ Stop p</td>
</tr>
<tr>
<td><strong>C</strong></td>
<td>CUU ⟷ Leu</td>
<td>CCA ⟷ Gln</td>
</tr>
<tr>
<td></td>
<td>CUC ⟷ Pro</td>
<td>CCA ⟷ Gln</td>
</tr>
<tr>
<td></td>
<td>CUA ⟷ His</td>
<td>CUA ⟷ Gln</td>
</tr>
<tr>
<td></td>
<td>CUG ⟷ Arg</td>
<td>CUG ⟷ Arg</td>
</tr>
<tr>
<td><strong>A</strong></td>
<td>AUU ⟷ Ile</td>
<td>ACC ⟷ Lys</td>
</tr>
<tr>
<td></td>
<td>AUC ⟷ Thr</td>
<td>ACA ⟷ Lys</td>
</tr>
<tr>
<td></td>
<td>AUA ⟷ Lys</td>
<td>AAA ⟷ Lys</td>
</tr>
<tr>
<td></td>
<td>AUG ⟷ Arg</td>
<td>AAG ⟷ Lys</td>
</tr>
<tr>
<td><strong>G</strong></td>
<td>GUU ⟷ Val</td>
<td>GCU ⟷ Asp</td>
</tr>
<tr>
<td></td>
<td>GUC ⟷ Ala</td>
<td>GCA ⟷ Asp</td>
</tr>
<tr>
<td></td>
<td>GUA ⟷ Arg</td>
<td>GCC ⟷ Asp</td>
</tr>
<tr>
<td></td>
<td>GUG ⟷ Arg</td>
<td>GCG ⟷ Asp</td>
</tr>
<tr>
<td></td>
<td>Val</td>
<td>Ala</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GCA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GCG</td>
</tr>
</tbody>
</table>