Name: _____KEY_____

7.012 Final Exam -- 2006 KEY

You have 180 minutes to complete this exam.

There are 19 pages including this cover page, the AMINO ACID page, and the GENETIC CODE page at the end of the exam.

Please write your name on each page.

Only writing on the **FRONT** of every page will be graded.
(You may use the backs, but only as scratch paper.)

Question 1	**37 pts_____**
Question 2	**22 pts_____**
Question 3	**16 pts_____**
Question 4	**19 pts_____**
Question 5	**25 pts_____**
Question 6	**27 pts_____**
Question 7	**28 pts_____**
Question 8	**26 pts_____**
TOTAL	**200 pts_____**

Name: _____KEY_____

1. (37 pts) You discover 3 new yeast mutants, mutant A, mutant B, and mutant C, that all contain mutations in genes that encode enzymes that work in the pathway for tryptophan synthesis.
(a, 4 pts) When you did the genetic screen that allowed you to isolate these 3 mutants, what was the mutant phenotype you were looking for and how did you assay for that?

The mutant phenotype is the inability to synthesize tryptophan, or the inability to grow on medium lacking tryptophan. This is assayed by growing each colony on two types of medium (+ tryp and – tryp), and finding the mutants that grow on the former but not the latter.

You find that mutant A is recessive.
(b, 3 pts) What experiment could you have done to infer this?

Mate haploid mutant yeast to haploid wild-type yeast.

(c, 3 pts) What would have been the result of this experiment that led you to this conclusion?

The resulting diploid would be able to grow on medium lacking tryptophan.

[You also find that mutants B and C are also recessive, but you do not need to tell us how this was done.]
You find that mutant A and mutant B have mutations in the same gene.
(d, 3 pts) What experiment could you have done to infer this?

Mate haploid mutant A yeast to haploid mutant B yeast.

(e, 3 pts) What would have been the result of this experiment that led you to this conclusion?

The resulting diploid would be unable to grow on medium lacking tryptophan.

You find that mutant A and mutant C have mutations in different genes.
(f, 3 pts) What experiment could you have done to infer this?

Mate haploid mutant A yeast to haploid mutant C yeast.

(g, 3 pts) What would have been the result of this experiment that led you to this conclusion?

The resulting diploid would be able to grow on medium lacking tryptophan.

[You find that mutant A accumulates Compound X and mutant C accumulates Compound Y, but you do not need to tell us how this was done.]

Name: _____KEY_____

You find that gene A acts more upstream than gene C in the pathway.
(h, 3 pts) What experiment could you have done to infer this?

Make a double mutant haploid strain that contains both mutant A and mutant C.

(i, 3 pts) What would have been the result of this experiment that led you to this conclusion?

The double mutant would accumulate compound X.

You identify the gene that is mutated in mutant C.
(j, 3 pts) What kind of experiment could you have done to infer this?

Cloning by complementation. You would transform a plasmid library into mutant C yeast and find the one plasmid in the library that could convert a mutant C yeast into a yeast that will grow on medium lacking tryptophan.

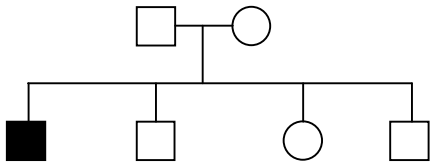
[You find that gene A and gene C have homologs in mice, and you obtain a mouse that is homozygous for loss-of-function mutations in both gene A and gene C, but you do not need to tell us how this was done.]
You find that gene A is very tightly linked to gene C.
(k, 3 pts) What experiment could you have done to infer this?

Mate the A–A– C–C– mouse to a wild-type mouse to make A–C–/A+C+ mice in the F1 generation. Then mate these F1 mice to A–C–/A–C– mice and look at the mice in the F2 generation.

(l, 3 pts) What would have been the result of this experiment that led you to this conclusion?

You would see only mice that accumulate compound X (A–C–/A–C– mice) and mice that are wild-type (A+C+/A–C– mice), because these are the two parental classes. You would not have seen any mice that accumulate compound Y, because those mice would be recombinants.

2. (22 pts) The parents in the pedigree below already have four kids (as you can see), and they have just given birth to two more – a set of fraternal twins, one boy and one girl. Their oldest son displays a trait, indicated below by shading. This trait is caused by an inherited mutation in "gene Q." The parents haven't even seen their newborn twins yet, and the nurse is about to bring them in.



Name: _____KEY_____

(a, 8 pts) Assume the trait indicated by the shaded-in shape is autosomal recessive.

-- What is the chance that the girl twin will have the shaded-in trait?

25%.

-- What is the chance that the boy twin will have the shaded-in trait?

25%.

-- What is the genotype of the father?

Qq.

-- What is the genotype of the mother?

Qq.

(b, 8 pts) Assume the trait indicated by the shaded-in shape is X-linked recessive.

-- What is the chance that the girl twin will have the shaded-in trait?

Zero.

-- What is the chance that the boy twin will have the shaded-in trait?

50%.

-- What is the genotype of the father?

X^QY.

-- What is the genotype of the mother?

X^QX^q.

(c, 6 pts) Given the above pedigree, is it possible for the trait indicated by the shading to be caused by an inherited mutation…

-- on the Y chromosome?

No. If it were on the Y chromosome, none of the boys would have the trait, because the dad does not have the trait.

-- in the mitochondrial DNA?

No. If it were caused by a mitochondrial gene, the children would either be all wild-type or all mutant.

-- in a maternal effect gene that is found on an autosome?

No. If it were caused by a maternal effect gene, the children would either be all wild-type or all mutant.

3. (16 pts) You have part of the sequence of both a "wild-type" and a "mutant" DNA sequence from one region of chromosome 13 in mice.

```
Wild-type:   5′-ACTTGCAAGCGAATC-3′
Mutant:      5′-ACTTGCTAGCGAATC-3′
```

(a, 4 pts) The mutation seems to have created a stop codon. However you determine that the "mutant" mouse is not at all mutant in phenotype, and produces a protein of the exact same length and sequence as each of the corresponding proteins made by the wild-type mouse. List **two** different reasons for why this DNA sequence change may not shorten the length of any of the mouse proteins.

There are several possible answers:
-- this sequence change is not in a gene (most mouse DNA is non-coding)
-- this sequence change is in a gene but is in an intron
-- this sequence change is in an exon but is not on the correct strand
-- this sequence change is in an exon on the correct strand, but it is out of frame

(b, 5 pts) What would the effect be of a single nucleotide substitution in each of the following locations in a gene on the production of the protein product of that gene?

Mutation is in…	Effect on protein is…
the promoter	**The wrong amount of protein would be produced.**
the terminator	**There would be no effect on the protein.**
the start codon	**No protein (or a shorter protein) would be produced.**

Name: _____KEY_____

the stop codon	**The protein would be longer.**
the middle of an intron	**There would be no effect on the protein.**

(c, 3 pts) The insertion of how many nucleotides into the coding region of a gene affects the amino acid sequence of the product of that gene less than the insertion of the other numbers? Your choices are: 5, 6, or 7.

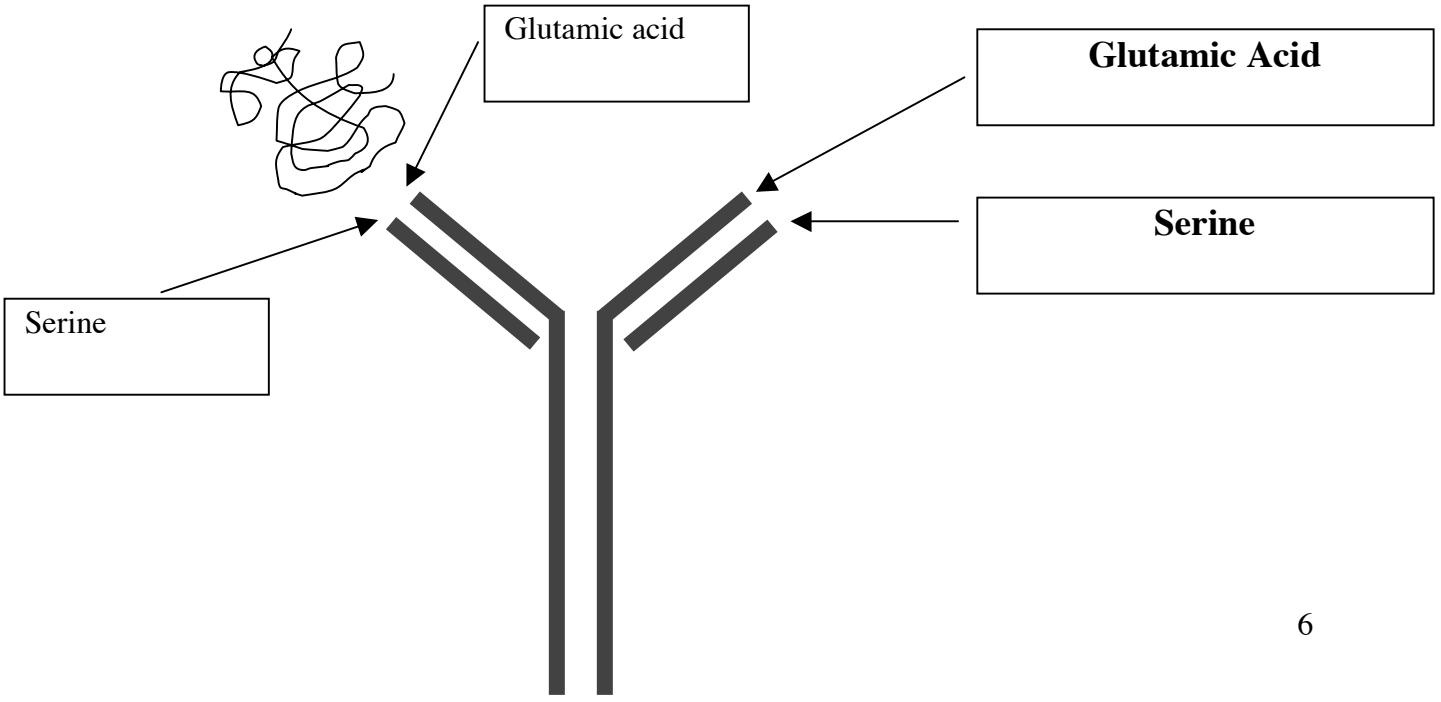
Six. If you insert six basepairs into the coding region of a gene, 2 amino acids will be added into the protein, which may or may not affect its activity. An insertion of 5 or 7 base pairs, however, will cause a frameshift mutation that will change every amino acid from the point of the insertion to the end of the protein.

(d, 4 pts) Answer "yes" or "no" to each of the four parts of this question. If a protein is normally 400 amino acids long, can you get a protein product that is 250 amino acids long as a result of:

a missense mutation? _____**NO**_____ a silent mutation? _____**NO**_____

a nonsense mutation? _____**YES**_____ a frameshift mutation? ____**YES**_____

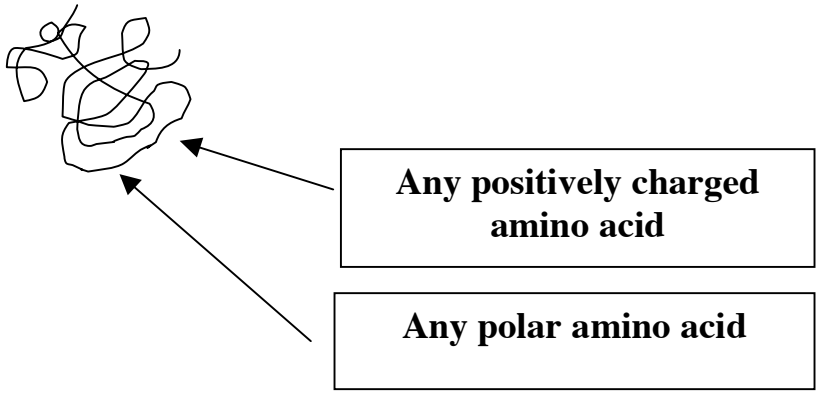
4. (19 pts) The following is a diagram of the structure of an antibody, with some of its amino acids labeled to show their identities. The antibody is shown bound to one molecule of antigen, diagrammed as a squiggle.



Name: _____KEY_____

(a, 4 pts) On the above diagram of the antibody, fill in each empty box with one name of an amino acid that must be found at that site on the antibody.

(b, 4 pts) On the following diagram of the antigen, fill in each box with one name of an amino acid that would allow for the strongest possible antigen-antibody interaction at that site at pH7.



(c, 3 pts) List one amino acid that would fall into each of the three different categories of amino acids at pH7:

-- very hydrophobic:

Examples include leucine, isoleucine, valine, phenylalanine, methionine, and proline.

-- charged:

You could put glutamic acid, aspartic acid, histidine, lysine, or arginine.

-- polar but uncharged:

Examples include serine, threonine, asparagine, and glutamine.

(d, 4 pts) The antigens that some antibodies recognize are made of DNA instead of protein. Which specific amino acids might you find in the antigen-binding site of an antibody that recognized DNA instead of protein at pH7? List all that apply.

Histidine, arginine, and lysine. This is because these are the three positively charged amino acids, and the positively charged amino acids can form ionic bonds with the negatively charged phosphates on the DNA backbone.

(e, 4 pts) You are using a gel electrophoresis technique that separates proteins by both size and charge; this technique has good enough resolution to separate any two

polypeptide chains as long as they are not identical in both length and charge. How many separate bands do you think you would see if you loaded into each lane of your gel:

-- one antibody molecule:

One band. The four protein chains in an antibody are all covalently bound to each other by disulfide bonds.

-- a tube of polyclonal antibody:

Many bands. Each antibody would be a different length so each would run a different distance.

-- one antibody molecule treated with a chemical that disrupts disulfide bonds:

Two bands. The four protein chains in an antibody would be broken apart, but they would travel in pairs because the two long chains are the same length and the two short chains are the same length.

-- one antibody molecule treated with a chemical that disrupts all covalent bonds:

No bands. This would break the protein down into the five atoms C, H, O, N, and S, which are too small to be detected.

5. (25 pts) The extracellular signals that cells sense that tell them to grow and divide are proteins called "growth factors." Growth factor signals are typically sent to one cell type from another cell type.
(a, 3 pts) Where in a cell would growth factor receptors localize?

To the cell membrane.

(b, 3 pts) What experiment could you do that would allow you to conclude that one specific growth factor receptor protein did indeed localize to the place you named in part **(a)?**

Make a GFP fusion to the growth factor receptor protein.

(c, 3 pts) Why can't growth factors enter cells and affect proteins in the cytoplasm of those cells directly? (i.e. Why are growth factor receptors necessary at all?)

Growth factors are proteins, and proteins are large and hydrophilic. The lipid membrane will only allow things that are small and hydrophobic to pass through.

(d, 4 pts) Say you had a collection of cells that contained a mutation in a growth factor receptor that promotes the development of cancer. What kind of test could you do with these cells in a Petri dish to see that they had abnormal growth properties?

Test to see if the cells would grow in a growth-factor-independent manner, or test to see if the cells do not exhibit contact inhibition.

(e, 4 pts) Say you had a collection of cells that contained a mutation in a growth factor receptor that promotes the development of cancer. What kind of test could you do with these mutant cells using a mouse to see that the cells had abnormal growth properties?

Inject the cells into a mouse and see if they form a tumor.

(f, 4 pts) Say you have just discovered a new gene. What could you do (without stepping into a biology lab) to determine with a fairly high degree of certainty whether or not this new gene you discovered most likely encodes a growth factor receptor?

Plug the sequence of the gene into the program BLAST. If all of the top hits in your search are growth factor receptor genes, then your new gene is most likely a growth factor receptor gene also.

(g, 4 pts) Say you had a collection of cells that contained a mutation in a growth factor receptor gene that promotes the development of cancer. However the growth factor receptor protein has the exact same sequence as it does in non-cancerous cells. What influence has the mutation had on the growth factor receptor gene?

The mutation influenced the amount of growth factor receptor that was produced from the gene, such that the protein product of the gene became overexpressed.

6. (27 pts) Answer the following questions about immunology.

(a, 4 pts) What is the main difference between proteins whose epitopes are displayed on MHC class I molecules and proteins whose epitopes are displayed on MHC class II molecules?

MHC class I molecules display proteins that are inside of cells. MHC class II molecules display proteins that were outside of cells, but were engulfed by either macrophages or B cells.

(b, 4 pts) Some viruses can block the host cells' ability to display proteins on MHC class I molecules. Explain how this would allow a virus to avoid the immune system.

MHC class I molecules display proteins that are inside of cells. This means that, if a virus has infected a cell and is inside of it, viral proteins would be displayed on the surface. These viral proteins displayed on MHC class I molecules would tell Tc cells to attack the

infected cell. If the virus blocks MHC class I display, then the Tc cells will no longer attack infected cells.

(c, 4 pts) Why do mature B cells need lots of endoplasmic reticulum?

Endoplasmic reticulum is an organelle involved in the production of secreted proteins and transmembrane proteins. A B cell's job is to make and secrete antibodies, so it needs lots of ER to be able to do that.

(d, 4 pts) There is a special kind of PCR reaction that is used to detect the number of HIV genomes present in a patient infected with the HIV virus. This special kind of reaction is called RT-PCR, because it consists of a PCR reaction that is preceded by a round of replication using the enzyme reverse transcriptase. Why can you not directly test the number of genomes of HIV present in a human using regular PCR? Why must you use this special RT-PCR reaction instead?

The genome of HIV is single-stranded RNA, which cannot be used as a template in PCR to do exponential amplification, because it is not double-stranded. Thus you need to convert the ssRNA genome into dsDNA first, so that you can then exponentially amplify the dsDNA in this PCR reaction.

(e, 4 pts) In class, we discussed an assay called a "plaque assay." In this assay, one detects the number of clearings of host cells (called "plaques") that result when you expose a layer of host cells to a dilute solution containing viral particles. Each plaque corresponds to the original infection of one host cell with one infectious viral particle. How do plaques end up being so large that they are visible by eye, if they result from the initial infection of one host cell by one viral particle?

One virus infects one cell, and then bursts it, releasing many baby viral particles. These baby viral particles then diffuse over to neighboring cells, infect them, burst them, and release even more viral particles. This cycle of infection after infection occurs until enough host cells are lysed and killed that you can actually see the clearing in the layer of host cells.

(f, 4 pts) The number of plaques formed in a plaque assay is invariably lower than the number of viral particles you see if you look at the solution under a microscope. Why is this? Why do these extra viruses that are obviously present in the solution not form plaques in the assay?

A plaque is formed by an infectious viral particle. If the viral particle contains a mutation that makes it uninfectious, the viral particle will not form a plaque. Many viral polymerases have very high mutation rates, which means that many, of not most, viruses contain mutations that make them uninfectious. (An alternative answer is that some of the viruses are hiding out inside the cells in the lysogenic phase.)

(g, 3 pts) You add an antibody (isolated from the blood of a person who has had an infection of that virus) to the solution of viral particles before beginning a plaque assay

experiment. You then expose a layer of host cells to this mixture. How many plaques do you expect to see in comparison to part **(e)**, when you did not add the antibody to the viral solution first? Your choices are more, less, the same, inconclusive.

Less. The antibodies from the person's blood will bind to the virus and prevent it from being able to bind to the host cells.

7. (28 pts) You are studying olfaction (smelling) in a dog. This dog's genome contains 2,000 genes for olfactory receptors, but each olfactory neuron in the dog's body expresses only a single olfactory receptor that binds to a single odorant (smelly compound). You hypothesize that the expression of a single receptor in each neuron is **either** due to transcriptional regulation ("Hypothesis One"), **or** is due to the fact that 1,999 genes encoding olfactory receptors are physically removed from the genome of each olfactory neuron during differentiation ("Hypothesis Two").

(a, 5 pts) List the steps of an experiment you would do using organismal cloning that would distinguish between these two hypotheses.

Take an adult dog's olfactory neuron and remove its nucleus. Take an unfertilized dog egg and remove its nucleus. Put the olfactory neuron nucleus into the enucleated egg. Implant this newly nucleated egg into a pseudopregnant dog. Let the cloned dog grow up.

(b, 4 pts) Describe the two potential results of your experiment and state which hypothesis each potential result supports.

Result One = the cloned dog can smell anything. This result supports Hypothesis One.

Result Two = the cloned dog can smell only one odor. This result supports Hypothesis Two.

(c, 4 pts) In mice, gray fur is dominant over white fur, and is determined by a single fur color gene (-- let's call it "gene H"). You isolate ES cells from a mouse embryo that would have grown up to be a gray mouse and was homozygous at the fur color gene. You inject these ES cells into the blastocyst of a mouse embryo that would have grown up to be a white mouse. You then implant this new, mixed blastocyst into a mouse mother, and wait for the mouse to be born. The mouse that is born is a chimera with some white fur and some gray fur. What are the genotypes of:

-- the gray cells in this mouse

HH.

-- the white cells in this mouse

hh.

Name: _____KEY_____

You generate four chimeric spotted mice in the exact same way as described in part **(c)**. We will call them Mouse 1, Mouse 2, Mouse 3, and Mouse 4. You breed each mouse to a white mouse, and see the following results.

Parent bred to white mouse	**% offspring that are gray**
Mouse #1	0%
Mouse #2	35%
Mouse #3	15%
Mouse #4	???

(d, 4 pts) Mouse 4 had the maximum percentage of gray offspring that is possible from this experiment. What percentage of offspring from Mouse #4 were gray?

100%. This would occur if the entire germline of the chimeric mouse consisted of gray HH cells.

(e, 4 pts) What factor determined the different percentages listed above? Why did mouse #1 have no gray offspring whereas the other mice all had some gray offspring?

The factor that determines it is what percentage of the gremline is made up of gray HH cells. Mouse #1's germline must have consisted of all white hh cells.

(f, 3 pts) Is it possible that any of the offspring of any of the four mice were gray-and-white-spotted, like their one parent was? If so, which of the four mice could have spotted offspring?

No, none of these mice will give birth to spotted offspring. The mouse you breed to is hh and will thus give a "h" allele to the baby mouse. Mouse #1, #2, #3, or #4 will either give "H" or "h." Thus the baby mouse will either end up H/h (gray) or hh (white).

(g, 4 pts) If you took two gray mouse babies whose parent was Mouse #2 (and thus are siblings of each other) and mated them to each other when they reached adulthood, what would percentage of their offspring would be gray?

75%. You would be mating an H/h mouse to an H/h mouse.

Name: _____KEY_____

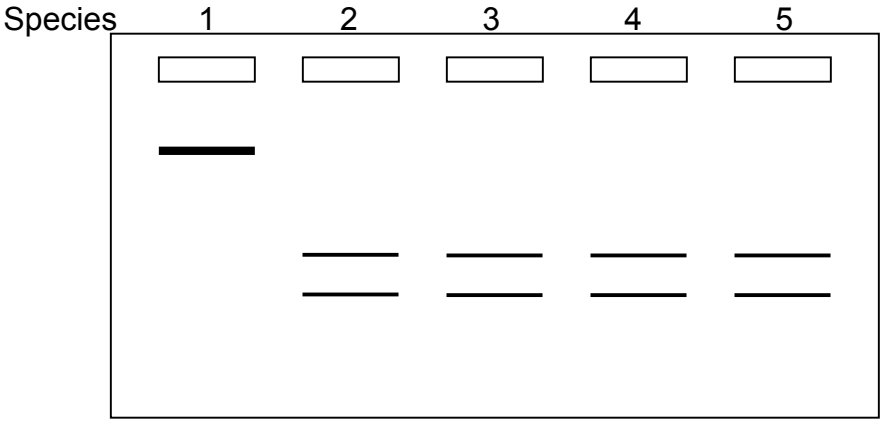
8. (26 pts) You have discovered 5 new species of bacteria, and your job is to make a phylogenetic tree showing the evolutionary relatedness of these 5 new bacteria to each other. Other biologists have already shown that there are two main groups of bacteria, and most species fall into one of these two groups. There is a certain sequence in a particular region of the genome that is associated with being in one group, and a sequence that differs by a single nucleotide that is associated with being in the other group. This sequence change happens to affect a site that is cut by a restriction enzyme that recognizes the sequence 5'-TTATAA-3' and cleaves in the center of this site. Below is shown the DNA sequence of one strand of DNA from this region from Species 5. The regions of DNA that are bound by the PCR primers you use in this assay are marked and shown in bold. You design the primers to regions that are shared in sequence by all bacteria.

Left primer →

Species 5: 5′-**GGCATATGCTATGCT**AGGGTTATAAGCACGT**GCGCATACTGCTCTTAGC**-3′

← Right primer

You do a PCR reaction on the DNA of each of the 5 bacterial species, treat the PCR product with this restriction enzyme, and separate the products using gel electrophoresis. You see the following results.



(a, 3 pts) What are all of the components you have to put in the test tube in order to perform a successful PCR reaction?

You need template DNA, DNA polymerase, primers, and dNTPs.

(b, 6 pts) What is the sequence of the two primers you used to do this successful PCR reaction? Make sure to label the 5' and 3' ends of each of your primers. Make sure the primer on the left is 15 nucleotides long and the primer on the right is 18 nucleotides long.

Name: _____KEY_____

Left: 5'-GGCATATGCTATGCT-3'
Right: 3'-CGCGTATGACGAGAATCG-3'

The sequence of the double stranded piece of DNA is:

5′-**GGCATATGCTATGCT**AGGGTTATAAGCACGT**GCGCATACTGCTCTTAGC**-3′
3′-CCGTATACGATACGA……CGCGTATGACGAGAATCG-5′

When you denature this DNA and add the two primers, the two primers must be facing each other and the two 3' ends must be pointing towards the inside of the stretch of DNA (i.e. the stretch you want to amplify):

5′-**GGCATATGCTATGCT**AGGGTTATAAGCACGT**GCGCATACTGCTCTTAGC**-3′
3′-**CGCGTATGACGAGAATCG**-5′

5′-GGCATATGCTATGCT-3′
3′-CCGTATACGATACGA……CGCGTATGACGAGAATCG-5′

(c, 3 pts) What is true about bacteria that ensures that you will never see three bands in any one lane of the gel when doing this analysis?

Bacteria are haploid. Thus they only have one allele of every gene. The allele will either be cut by the restriction enzyme, in which case you will see two bands, or it will not be cut, in which case you will see one band. (A diploid organism could display three bands in such a reaction if it were heterozygous with one allele that got cut and the other that didn't.)

You now determine the DNA sequence of the region that is internal to the two primers shown above in each of the 5 species of bacteria. You get the following results.

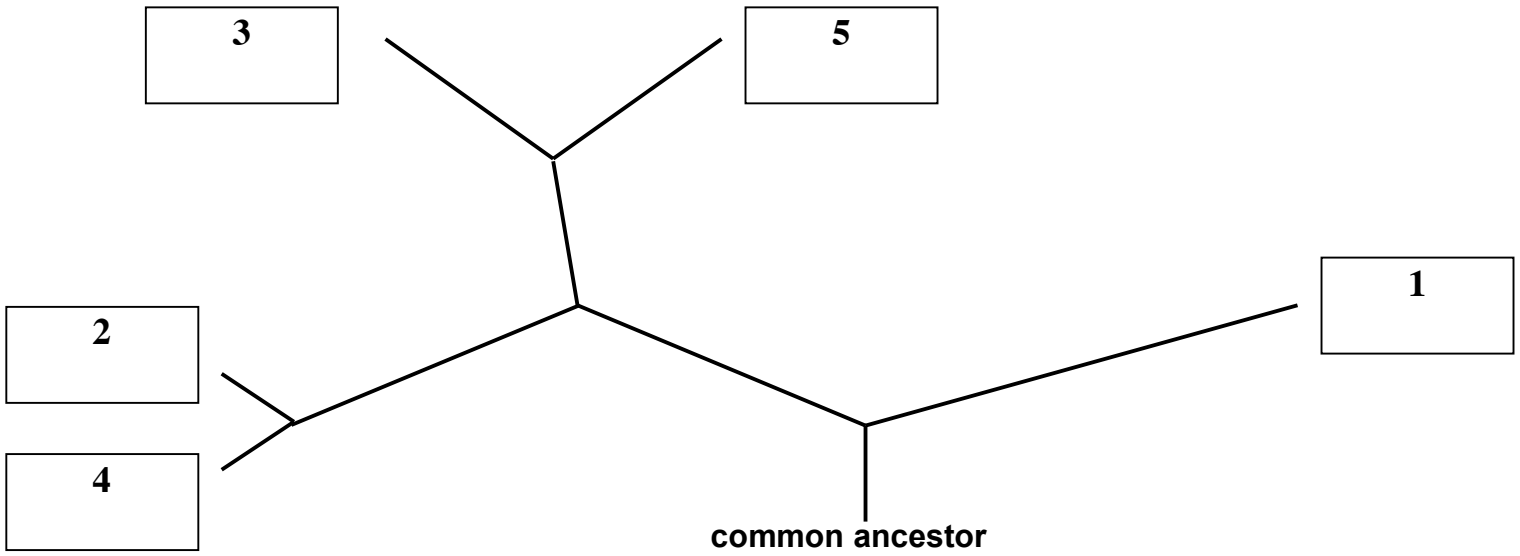
Species 1: **5′-**TCGTTGATAAGGACCC**-3′**
Species 2: **5′-**CGGATTATAACTACGT**-3′**
Species 3: **5′-**AGGGTTATAAGGACTT**-3′**
Species 4: **5′-**CGGATTATAACAACGT**-3′**
Species 5: **5′-**AGGGTTATAAGCACGT**-3′**

(d, 3 pts) What is the one component that you do put in the test tube in order to perform a successful sequencing reaction that you do not put in when performing a successful PCR reaction?

ddNTPs. These are nucleotides that terminate the chain during replication because they don't have a 3'OH.

Name: _____KEY_____

(e, 5 pts) Given the results of the gel analysis and the sequencing analysis, place the 5 species of bacteria onto the following phylogenetic tree. Do this by writing the number of one bacterial species into each blank box.



This is done by comparing how similar in sequence each organism is to each other. You know that #1 is the most distantly related, because the restriction digest data show you that #1 is in one group of bacteria, and the other four are in another group of bacteria.

(f, 3 pts) Based on your tree drawn above, which two species of bacteria are the two most closely related organisms?

2 and 4. In a phylogenetic tree, the closer in distance two organisms are to each other, the closer related they are. In this tree, #2 and #4 are the closest two organisms to each other.

(g, 3 pts) You notice that the overall genome content of the bacteria is as follows:

	%G	%A	%T	%C
Species 1:	30	20	20	30
Species 2:	41	9	9	41

Which species do you think would survive the best in a very hot environment?

Species #2. Species #2 has a high GC content, and GC basepairs are stronger than AT basepairs (because every G and C form three H bonds whereas every A and T form two H bonds). Thus an organism with more GC basepairs can survive higher temperatures better, since its DNA is harder to denature.