Sierra College
Microbiology

Name: _____
Date: _____

**PHYSIOLOGICAL UNKNOWN #2**

Questions asked in association with this project are designed to help you review the steps involved in the identification of an unknown bacterial culture using DNA analysis. By answering these questions completely and accurately you will gain a better understanding of the process, and potentially improve your ability to score well on Laboratory Exam 2. These methods also apply to portions of semester projects involving DNA analysis for bacterial identification. **Data required for your semester project is indicated on page 4**.

**Information from Exercise 16, Application of the PCR in Bacterial Identification:**
The bacterial culture assigned to you was culture # _____ and was provided on an agar slant or plate. What steps or procedure did you follow in order to release chromosomal DNA (genomic DNA) from these bacteria?

What does PCR stand for? _____. What enzyme is most commonly used to catalyze in vitro DNA replication? _____. This enzyme is said to be thermostable because it is not degraded by exposure to high temperatures. Why is this significant, i.e., why is stability at high temperatures an important attribute for enzymes used in the PCR? _____
_____

The enzyme contained in our PCR master mix was originally produced by bacteria identified as _____ that were found initially in the near-boiling water of hot springs located in Yellowstone National Park. Bacteria that can live and grow under such conditions are referred to as _____ with respect to their temperature requirements.

What are oligonucleotide primers? _____
What were the two primers used in our PCR exercise? _____ and _____. Why are primers necessary in the PCR, i.e., what functions do they serve? _____
_____

What region of DNA (what gene type) was amplified during our PCR? _____
Why was this particular target selected as a means of identifying bacteria?

The energy required for DNA synthesis during the PCR is provided by _____ that are contained within the master mix. This is one of the reasons this mix is so expensive.

The initial step associated with a typical PCR involves breaking the _____ bonds holding the complementary and antiparallel strands of DNA molecules together. As these strands separate, the DNA is said to be _____. What event occurring within the thermal cycler causes the DNA strands to separate? _____ _____. During the second step in the PCR, primers are allowed to bind with or to hybridize with their complementary DNA sequences, so are said to _____. This occurs when the temperature within the thermal cycler is lowered to around _____. The specificity of the primers is essential to amplifying a particular target sequence. After the primers are in place, the temperature within the thermal cycler is increased to _____ for the third step in the process, called _____. The Taq-polymerase can now attach new nucleotides starting with the _____ end of each primer sequence and build new DNA strands (building always proceeds from 5' to 3').

**Information from Exercise 20, Gel Electrophoresis of DNA Samples:**
A procedure or method that allows DNA samples to be separated on the basis of size by placing them in an agarose gel and exposing them to an electric field is called _____ _____. This process allows DNA samples to be observed and potentially recognized/identified. DNA samples will travel toward the _____ electrode when exposed to an electric field because they have a slight _____ charge. Smaller DNA fragments travel _____ through the gel than larger ones because they can slip between the agarose particles more readily. Our DNA samples were mixed with a loading/tracking dye composed of _____ and _____ so they would be visible when loaded into the wells of an agarose gel. This dye also allows the movement of DNA within the gel to be tracked because the dye particles have a negative charge, and travel along with DNA fragments of a particular size. Which dye samples moves faster through the gel? _____ When this dye has moved about 3/4 of the way down the gel, the gel is removed from the chamber and the DNA samples are stained with _____, a chemical that produces an orange fluorescence when subjected to ultra violet light. Once stained, multiple DNA fragments of a particular size will appear grouped together in one region of the gel and will form a distinct band. By comparing the location of a DNA band to the banding pattern formed by a known standard, it is possible to determine the approximate size of the fragments contained within the band.

**Information from Exercise 21, Restriction Endonuclease Digestion of DNA & RFLP:**
Enzymes that cut double-stranded DNA molecules (break phosphodiester bonds) at specific locations are known as _____. These enzymes bind to DNA molecules at specific locations known as _____ sequences (palindromic or inverted repeat sequences) and cut within or near these to form blunt or cohesive termini. The third one of these enzymes found within bacteria identified as *Haemophilus influenzae* strain Rd is called _____ and binds to DNA regions with the base sequence 5'-AAGCTT-3'. This enzyme will cut DNA between the two adenines to form cohesive termini four bases in length. When DNA from bacteriophage lambda is cut with this enzyme, it yields DNA fragments that range in size from 23,130 base pairs to 125 base pairs (23,130, 9,146, 6,557, 4,361, 2,322, 2027, 564 and 125 bp). When run in an agarose gel, these fragments form a distinctive banding pattern that is easily recognized. Restriction enzymes can be used to cut DNA from various sources, and can generate banding patterns known as DNA _____ or RFLP patterns.

What does RFLP stand for? _____.
(Note – the banding pattern generated by cutting bacteriophage lambda DNA with *Hin*dIII is a RFLP pattern.)

The PCR product DNA generated in this laboratory is 16S ribosomal DNA and is only about 1500 bp in length, much shorter than bacteriophage lambda DNA. However, if this DNA is cut with the restriction enzyme AluI it will yield a banding pattern that can be used to identify the bacteria from which the DNA was taken. What is the recognition sequence for AluI? _____. Where within this sequence will the DNA be cut? _____. If 16S ribosomal DNA from your physiological unknown #2 (**web page version of a word file**) were cut with AluI, how many times would it be cut? _____ What would be the sizes of the fragments generated? (please list the fragment sizes in sequence, from largest to smallest)
_____
To answer this question, access the 16S ribosomal DNA gene sequences available on the Bio. 4 Website and select your sequence by clicking on the grid number corresponding to your PUNK2 number. Use the "find" option of your word processor to find each cut site, and to count the bases in each fragment generated (see information in the laboratory manual).

Which gel (by number) and which lane within that gel (1-6 from top down) contains the RFLP for your unknown? _____ Does this pattern match the one you generated on paper? _____ What explanation might account for any variation observed? _____
_____

**Information from Exercise 17, Automated Nucleotide Sequencing & Electropherograms:**

What is an electropherogram? _____
A nucleoside triphosphate that is missing the hydroxyl group normally located on the 3' carbon of the sugar ribose is called a _____. What happens when a molecule with this configuration is incorporated into a nucleotide strand being synthesized by DNA polymerase and why? _____
_____.
When these nucleotides are labeled with fluorescent markers (different bases being represented by different colors), and are incorporated into populations of oligonucleotides of various sizes, these can be subjected to electrophoresis within a capillary tube and can be used to generate electropherograms (with the help of lazers and computers). Why does sequencing data generated by an automated sequencing machine and represented by an electropherogram require editing? _____
The three primers used to generate nucleotide sequences for use in this exercise were Bacteria 8-forward, Internal 533-forward and Enteric 1530-reverse. Which of these primers generated DNA sequences from the same template strand? _____
Why was it necessary to use the EDIT option "Reverse/Flip Sequence" with one of our electropherograms before copying it to the word file, i.e., what was unique about this sequence? _____

**Information form Exercise 18, Genomics, Proteomics & Bioinformatics:**

What is genomics? _____
What is proteomics? _____
Since the quantity of data generated by genomics and proteomics is immense, multiple interactive databases have been established, and systems developed for storing, accessing, managing, searching and analyzing the data accumulated.

A new discipline called _____ has emerged, and has application in both research and clinical settings.

What was the name of the public database accessed during the identification of physiological unknown #2? _____
What does BLAST stand for? _____

What did **you** use the BLAST algorithm for? _____
_____

Access the Microbiology Webpage, go to **Web-based Laboratory Assignments** and click on **PUNK2 Word Files**. Find the number of your physiological unknown #2 (PUNK2) in the grid and access the sequence **(word file)**. Copy the sequence and then click on the NCBI link at the bottom of the page. Follow instructions provided in your laboratory syllabus to access and utilize the NCBI nucleotide-nucleotide BLAST and compare the 16S Ribosomal-DNA gene sequence from your physiological unknown #2 culture with sequence information stored in the gene bank **[under 16S r-RNA sequences (Bacteria and Archaea), bottom of the option list]**. Complete procedural steps as instructed, and as indicated in your laboratory syllabus, to fill in the required information below:

Name of organism type showing the highest bit score _____
(include genus name, specific epithet, strain ID and subspecies name if listed), then record:

a. Length of Gene Bank sequence your query sequence is aligned with _____
b. Number of bases in your query sequence that are identical to the gene bank sequence _____ (this is presented as a ratio).
c. What is the percent identity between your sequence and the gene bank sequence? _____
d. What is the location and identity of unmatched bases? Note – some sequences have no unmatched bases and some have many. If your sequence has multiple unmatched bases, list the first two you find starting from the top and working downward. _____
   _____
e. What is the score (number before bits) for this alignment? _____

What is the Gene Bank accession number associated with the organism type with greatest sequence similarity to your unknown culture? _____ What is the taxonomic lineage information listed for these organisms? _____
_____

What is the name of the first author listed in association with this sequence, and what is the title of the reference article? Author is _____ Article title is _____
What is the Publication Name? _____
If there is no publication, record the name and address of the facility submitting the sequence.
_____

Return to the Results of BLAST (NCBI) page and look for the **[Distance Tree of Results.** Find your **query** on the cladogram **(highlighted in yellow).** Record the names and strain IDs of two different but closely related species (do not accept "leaves" in place of species identifications).
_____
_____