

Inquiry Approach to Molecular Biology Using Salt Marsh Bacteria

This laboratory exercise is composed of six parts to be completed over the course of a semester. The bacterial cultures used in this lab were obtained by last year's AP Biology class during their Sapelo Island trip. This lab reviews many of the molecular biology techniques that we need to get familiar with before the AP exam (see Chapter 20: Biotechnology).

After completion of all parts you will prepare a written report summarizing your activities during this laboratory exercise. The lab report must be typewritten and should be 3-5 pages in length (single spaced). If you worked with a partner during the lab, do not write the lab report together...it must be entirely your own work (it will be a test grade). Include the following sections in your report:

Introduction: What was the reason for carrying out this study? You might choose to focus your introduction on topics such as global biodiversity, or the role of bacteria in salt marshes, or any other framework that you think is relevant to the lab. There is no single right reason...you just need to choose something reasonable and present it well. Include at least 3 outside references (books, scientific papers), 2 of which must be different from the papers provided to you in lab. (Length of introduction: approximately 1 page).

Bacterial Sequencing:

Describe the laboratory methods you used to obtain a sequence of the 16S rRNA gene.

Sequence Analysis:

Describe what you learned about your organism and its relatives from the BLAST analysis at NCBI.

Part I

Objectives:

Students will isolate a single bacterial colony from agar plates made by last year's students.

Materials:

Plates
Sterile loops
Agar plates with bacterial culture

Obtaining a Pure Culture:

Each group will receive an agar plate containing a culture of bacteria isolated by last year's class. To assure us that the colonies are "healthy" we will do one more transfer before the extraction of DNA.

To transfer bacteria:

Use a sterile plastic loop and pick a single colony from the plate assigned to you. Streak onto a fresh agar plate and follow the method for streaking to single colonies (see Fig. 1).

Method for streaking to single colonies: A plastic loop containing cells is touched to the plate at location ① and the cells are then gradually diluted by streaking the loop back and forth across the plate. A second sterile loop is positioned at location ②, passed through the previous streak one time and then diluted further by streaking back and forth. A third sterile loop is positioned at location ③ and the process is repeated.

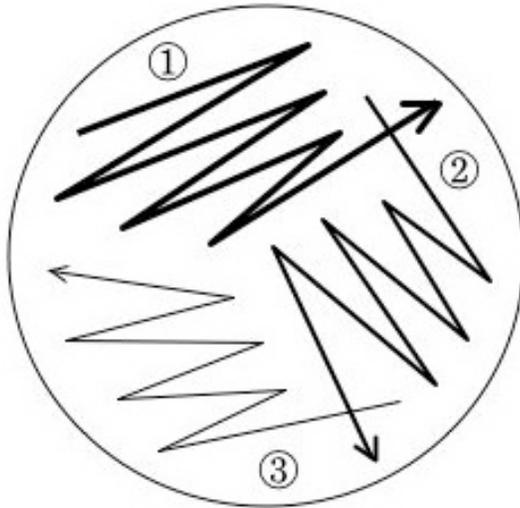


Figure 1. Leave your plate at room temperature 2-3 days and then place it in the refrigerator.

Part II

Objectives:

Students will extract DNA from a pure culture of marine bacteria isolated previously.

Materials:

Pure culture of bacteria
Sterile plastic loop
Pipets
Microcentrifuge tubes
Sterile water
Vortex
Boiling water bath

DNA Extraction:

- 1) For each bacterial isolate, use a sterile plastic loop to scrape a small blob of cells from the agar plate.
- 2) Stir the loop into a microcentrifuge tube containing 500 μl of sterile water. Close the tube and vortex very well (about 1 minute or until the cell clump is dispersed).
- 3) Place the tubes into a boiling water bath for 10 min. Use the special tube holder to prevent the tops of the tubes from opening during boiling.
- 4) Remove the tubes from the bath and centrifuge the contents for 10 min at 13,000 rpm in a microcentrifuge.
- 5) Transfer approximately half the supernatant to clean, labeled microcentrifuge tubes using an automatic pipettor. This will serve as the DNA template for your PCR reaction. If you need to stop at this point, you can store the DNA preparation (i.e., supernatant) in the freezer.

Part III

Objectives:

Students will set up a PCR reaction.

Materials:

Microcentrifuge tubes containing a bead (reaction mix and enzyme)

Sterile water

Sharpie

Ice bucket

DNA preparation

Primers

Vortex

Microcentrifuge

Thermal cycler

Set up the PCR Reaction:

- 1) Obtain two microcentrifuge tubes containing a bead (of reaction mix and enzyme) in the bottom, and one tube containing sterile water.
- 2) Using a permanent marker, label the top and sides of the tubes containing the beads. Label one with the name of your bacterial isolate and the second as a negative control.
- 3) Keep all tubes in the ice bucket as much as possible. To the tubes with a reaction bead in the bottom add, in this order:
 - 23 μl sterile dH_2O
 - 1 μl DNA preparation or sterile dH_2O (one tube gets DNA from the isolate; the negative control tube gets sterile dH_2O)
 - 1 μl forward primer (27F) (get this from your teacher)
 - 1 μl reverse primer (1522R) (get this from your teacher)
- 4) Gently vortex the contents of the tubes, and then spin briefly (30 sec) in a microcentrifuge. Keep the tubes in an ice bucket until it is time to load the thermal cycler.
- 5) Load the two tubes into the thermal cycler when the other groups are ready. The thermal cycler will cycle through three temperatures: 94°C to separate the double stranded DNA into single strands, 55°C to allow the primers to bind to complementary regions on the bacterial 16S rRNA gene, and 72° to allow *Taq* polymerase to fill in nucleotides between primer binding sites. After 35 temperature cycles (which takes about 3 hours), the machine will cool to 4°C .
- 6) When the reaction is complete, we will move the tubes to the refrigerator for storage until you return to run out the PCR reaction on an agarose gel.

Part IV

Objectives:

Students will be able to detect the presence of DNA in their samples by using gel electrophoresis.

Students will be able to make a standard curve using a marker.

Students will be able to calculate the number of base pairs in their DNA samples.

Materials:

Agarose gel

PCR reaction mixture

Blue loading dye

DNA marker

Automatic pipette

Parafilm

Gel box

Buffer

Power supply

UV lamp

UV-protective lab glasses

Before we start to describe the protocols for electrophoresis see this link

<http://learn.genetics.utah.edu/content/labs/gel/>

Electrophoresis:

- 1) An agarose gel has already been prepared, poured, and allowed to solidify (1% agarose gel: for 50 mL volume add 0.5 g agarose to 50 mL TAE buffer). A plastic “comb” was placed in the gel before it solidified to make small wells that can hold liquid.
- 2) Retrieve your two PCR tubes from the refrigerator. Cut a small square of parafilm and place it on the bench top. Place three small dots containing 1 μ l of blue loading dye about 2 cm apart on the parafilm (one dot for each PCR tube).
- 3) Remove 7 μ l of PCR reaction mixture from each tube and add to one of the spots of loading dye. Mix gently with the pipette tip. Use a separate pipette tip for each sample. Put the PCR tubes back in the refrigerator (you will need them later).
- 4) Pull up the loading dye/PCR mixture from one dot with the pipette. With a steady hand, load this liquid into a well in the agarose gel. Repeat for the other dots, using a clean pipette tip and loading each into a different well.
- 5) In a fourth well, load 5 μ l of DNA marker (labeled MWIII). This will provide a ladder of DNA fragments of known length, so you will be able to check the size of your PCR product.
- 6) When all the lanes are loaded, add buffer to the gel box and start the current.

- 7) After 40 minutes, a line of blue bands should be visible about a third of the way down the gel (this is the loading dye). Turn off the current. Wearing gloves, remove the gel from the gel box and place on the transilluminator (a light box containing a UV lamp). Close the plastic cover of the box and turn on the light. **Do not expose your eyes or skin to the UV lamp. Wear UV safety glasses and make sure the plastic cover is closed before looking at the transilluminator.**
- 8) Look at the bands in the gel. A compound added to the gel (GelRed) interacts with the DNA to make it visible (as light pink fluorescent bands) under UV illumination. Ideally, you should see one band in each of your bacterial DNA lanes with an approximate size of 1500 bp (Fig. 2) and no band in your negative control.

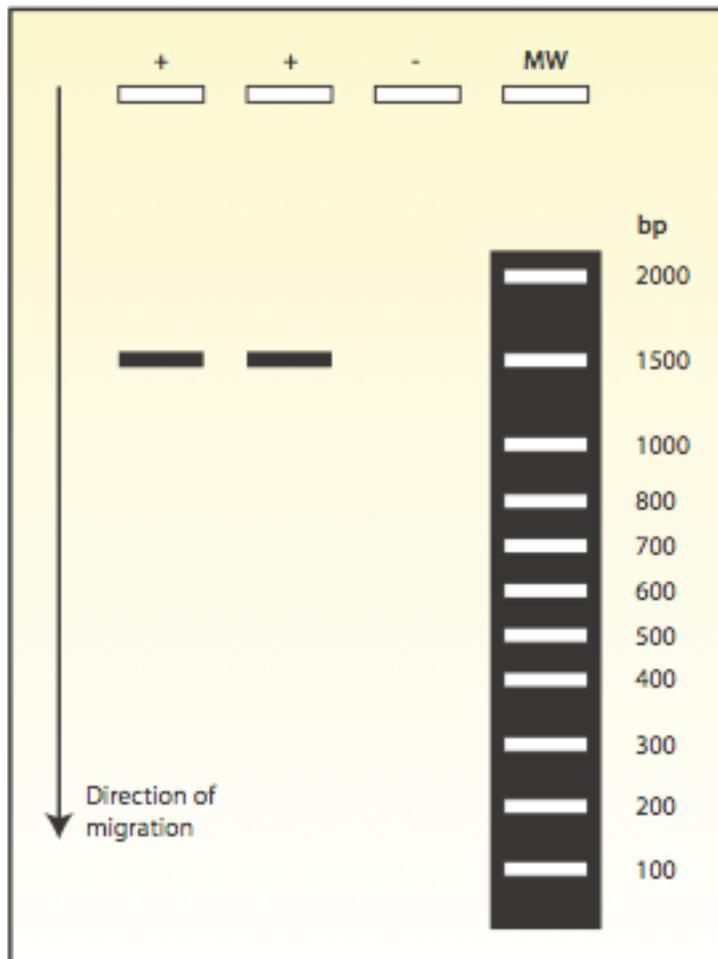


Figure 2. Agarose gel showing ~1500 bp PCR product from the bacterial 16S rRNA gene. The first two lanes (labeled '+') are PCR products from two bacterial isolates. The third lane (labeled '-') is the negative control. The fourth lane (labeled 'MWIII') is the molecular weight marker.

Using the molecular weight marker information prepare a standard curve and determine the fragment size of your PCR product.

Part V

Sequencing:

Choose the best PCR product for DNA sequencing based on its appearance in the gel (brightest band, no extra bands).

The sequencing will be done at UGA and the results will be used to perform the last part of this laboratory exercise. In the mean time we will be working on the [Bioinformatics unit](#) to gain an understanding of sequencing and perform gene annotation.

Part VI

In this part of the Biodiversity lab you will work with the 16S rRNA sequence of your bacterium using databases and programs available on the Internet. This exercise will tell you whether your isolate corresponds to a previously characterized species, or whether it is something new. If it is new, you can determine which organisms are its closest relatives, and perhaps predict some of the physiological properties of your unknown based on what is known about its relatives. For this part you will be assigned one or two sequences from the following file: [Marine Bacteria Sequences](#).

BLAST Search at GenBank:

GenBank[®] is the NIH genetic sequence database, a collection of all publicly available DNA sequences (*Nucleic Acids Research*, 2008 Jan;36(Database issue):D25-30). There were ~150,000,000 sequence records in GenBank as of August 2009.

Included among these many sequences are 16S rRNA gene sequences from Bacteria and Archaea. The current number of 16S rRNA sequences is approximately 12,000. Some of these are 'complete' (they cover the entire 1500+ bases of the 16S rRNA gene), while other are 'partial' (like yours will be) and cover only a portion of the gene (your sequence includes only ~400 bases near the beginning of the gene).

You will use the internet server at the National Center for Biotechnology Information (NCBI) for sequence searches. A sophisticated computer program called BLAST compares the sequence you submit (called the “query sequence”) to all of the sequences previously deposited in the GenBank database. BLAST finds sequences in the database that are most similar to the submitted sequence (called the “subject sequences”), and then returns the names and a brief description of these matching sequences in the output from the program.

Conducting a BLAST Search:

- 1) Open the text file that contains your sequence and trim off the beginning and end of the sequence to remove most of the "N"s. These Ns represent bases that could not be properly identified by the DNA sequencer, and are typically clustered at the ends of the sequence. By trimming the first and last ~25 bases, you can get rid of most of them. It's OK to have a few Ns in the rest of the sequence, but a large number of them indicate that your sequence is not very good.
- 2) Count the number of bases in your trimmed sequence and write it here: _____. You can use MS Word (the “word count” tool) or just count manually.
- 3) Using a web browser, go to the following URL: <http://www.ncbi.nlm.nih.gov/BLAST/>.
- 4) On the web page that opens, choose the link labeled “nucleotide blast”.
- 5) The next web page (shown in Fig. 1) provides a box into which you paste your trimmed 16S rRNA sequence. You may first need to click inside the box in order to activate it.
- 6) In the “Chose Search Set” section, use the pull down menu to change the database “Nucleotide Collection (nr/nt)”.
- 7) Click on the "BLAST" button to start the query.

The screenshot shows the NCBI BLAST web interface. At the top, there are navigation links: Home, Recent Results, Saved Strategies, and Help. The main heading is "BLAST Basic Local Alignment Search Tool". Below this, there's a section for "Enter Query Sequence" with a large text input area and a "Clear" button. To the right of the input area are "Query subrange" fields for "From" and "To". Below the input area is an "Or, upload file" section with a "Choose File" button and a "Job Title" input field. The "Choose Search Set" section includes a "Database" dropdown menu set to "Human genomic plus transcript (Human G+T)", an "Exclude" section with checkboxes for "Models (X/M/X/P)" and "Uncultured/environmental sample sequences", and an "Entrez Query" input field. The "Program Selection" section has radio buttons for "Highly similar sequences (megablast)", "More dissimilar sequences (discontiguous megablast)", and "Somewhat similar sequences (blastn)". At the bottom, there is a blue "BLAST" button and a summary of the search parameters: "Search database Human G+T using Megablast (Optimize for highly similar sequences)" and a checkbox for "Show results in a new window".

Figure 3. View of NCBI BLAST page. Paste your sequence into the top box labeled 'Enter Query Sequence', chose the correct database (“Nucleotide collection”), and click the blue 'BLAST' button.

8) Typically, it takes less than 2 minutes for the query to be completed; in the meantime, you will see a “Job Title” page that automatically refreshes. Once the results of your database search are returned you will see that they are presented in three different formats:

A. **The Graphic Summary.** The first format (under the heading "Distribution of Blast Hits on the Query Sequence") is a graphical representation of the results, showing the matches to your input sequence as a stack of colored bars with the best matches at the top. The color of the bar indicates the match score (i.e., degree of similarity) between your sequence and the database matches. Red indicates the best matches; black indicates the poorest. The length of the bar shows the length of the subject sequence that was similar to your query sequence.

B. **The Descriptions.** Below the graphical output is a list of the top 100 sequences that matched your query, giving their reference number in the database, a very brief description, a score (the higher the score, the more similar the sequences), and an E value (the odds that the similarity between the query sequence and the match might have occurred purely by chance; the lower the score, the more similar the sequences).

C. **The Alignments.** Finally, a base-to-base alignment of the query sequence with each of its matches is shown, starting from the "best match" and progressing down the list. This format repeats the alignment score and E value given above, but also has a measurement called "Identities", which give the number of exact matches between your sequence and the sequence from GenBank and is presented as a fraction and as a percent.

9) Scroll down to the alignments section and look at the identities of the best match (the first one listed). Did you get a 100% match? Look down the alignment list and record the names of the 5 best matches, including the number of positions compared and the percent match.

10) As a “rule of thumb” sequences that are $\geq 97\%$ similar are in the same species. Sequences that are $\geq 95\%$ similar are in the same genus. Did you find a match in an existing genus or species?

11) Follow the link to the description of the top 5 matches by clicking on the accession number (listed to the left of the organism name). This page shows you the GenBank entries for the matches, and provides information on who entered those sequences and where they came from. Record information ([on this table](#)) about where each sequence was found, whether it came from an isolate or a clone (i.e., a sequence that was obtained without isolating the organism), whether or not it is marine, whether anything is known about its physiology, and anything else that looks interesting. Because each researcher enters his/her own sequences into GenBank, there is a lot of variation in the type and amount of detail given. In some cases, there may be a reference to a published paper that will give you more information.

12) Many of the entries in GenBank are not described species. Instead, they are listed as an 'unidentified bacterium' or a 'uncultured bacterium'. Scroll down through the list of matches to find the ten closest relatives that actually have species names (you may already have up to 5 from the step above). As above, follow the link to the actual GenBank entry and collect information on where and how these related bacteria were isolated as well as anything that is known about their physiology. Record the "identities" of the sequences (i.e. number of matching positions and percent match).