

## Bioinformatics: Electronic PCR

### PURPOSE

In this activity students will use a computer program known as **BLAST** (**B**asic **L**ocal **A**lignment **S**earch **T**ool) to perform “electronic PCR” with the primer sequences that you will use in the GMO experiment. BLAST is a highly useful online tool for identifying and comparing DNA or protein sequences. Based on the results of the BLAST search, students will 1) identify the type of sequence to be amplified by their particular PCR primers and 2) predict the exact length of DNA produced in their PCR experiment.

### MATERIALS

Make sure each student has a copy of the “Bioinformatics: Electronic PCR” worksheet and has read the “Student Case Brief”, found in the student guide for the Genetically Modified Organisms experiment. If possible, also have students visit <http://www.dnai.org/b/index.html> on their own beforehand to go through a PCR animation (click on “techniques” then “amplifying” then “making many copies of DNA”). For the activity below, each student or pair of students will need access to a computer with internet.

### LESSON PLAN

#### Engagement (20 minutes)

Introduce students to the PCR laboratory by going over the scenario for the experiment. Define PCR and explain how PCR can be used to detect genetically modified organisms (e.g. PCR will be used to make many copies of a DNA sequence that is found in many genetically modified organisms etc.).

Discuss the following questions with the students:

1. What components or “ingredients” are necessary to perform PCR?
2. What are “primers” and why are they needed for PCR?
3. How do primers “know” where to attach to the template DNA?
4. How exactly will the PCR tell us if our soy samples are GMOs?

### Objectives

- Introduce students to bioinformatics
- Learn to use a genome search tool to identify DNA sequences and predict expected PCR product sizes
- Reinforce the basic concepts behind the polymerase chain reaction



### Technology Tip:

**If possible, you can use a SMART board or project the BLAST website for the whole class so as to provide guidance when needed. If you feel the activity is too challenging for individual students to complete, this activity can be done together as a class.**

### Go Online!

**For:** A great PCR tutorial and game for students....

**Visit:**

[http://nobelprize.org/educational\\_games/chemistry/pcr/](http://nobelprize.org/educational_games/chemistry/pcr/)

Next, introduce the bioinformatics activity to your students by posing the following questions: **What does it mean to do an experiment *in vitro*? *In vivo*? *In silico*?** Some students may know the difference between the first two types of experiments. *In vitro*, literally meaning “in glass”, refers to something outside the living organism, in an artificial environment. *In vivo*, literally meaning “in the living” refers to something occurring within the living organism. However, in the last several years, scientists have begun to use the phrase *in silico* to describe experiments occurring on a computer. Indeed, many molecular geneticists are using computer-based programs and databases to conduct their research and perform virtual experiments.

Explain to students that they will be conducting a **bioinformatics** “experiment” *in silico* with the help of a computer program known as **BLAST (Basic Local Alignment Search Tool)**. BLAST is a very popular online tool for geneticists and molecular biologists that utilizes a large database of known DNA sequences for searching out similar sequences. Before the actual PCR wet lab is performed, students will use the BLAST website to determine the type of DNA sequence that will be amplified by the reaction, as well as the predicted size of their PCR product, should their sample be genetically modified.

### Exploration (20+ minutes)

Direct students to complete the “Bioinformatics: Electronic PCR activity” on the following pages.

### Explanation and Assessment (10 minutes)

As a class, go over the answers to the worksheet questions, correcting any misunderstandings as needed. Come to a group consensus as to the expected size of the PCR product that may be obtained during the wet lab. Reinforce that *only samples that are genetically modified will have the foreign DNA sequence necessary for PCR amplification*. Some students will receive genetically modified soy flour, while others will not. Thus, not all samples will produce a PCR product!

### Fact Files

#### Did you know?

NCBI (National Center for Biotechnology Information) is the largest database for genetic and biotechnology information. The site maintains a wealth of literature and genetics databases as well as many bioinformatics search tools. By the end of 2005, NCBI’s GenBank database contained over 52 million published DNA and protein sequences!



### Bioinformatics

(bi-O-in-f&r-'ma-tiks) – The analysis of biological information using computers

## Bioinformatics: Electronic PCR

In this activity you will use a computer program known as **BLAST** (**B**asic **L**ocal **A**lignment **S**earch **T**ool) to perform “electronic PCR” with the primer sequences that you will use in the GMO experiment. BLAST is a highly useful online tool for identifying and comparing DNA or protein sequences. Based on the results of the BLAST search, you will 1) identify the type of sequence to be amplified by their particular PCR primers and 2) predict the exact length of DNA produced in their PCR experiment.

Directions:

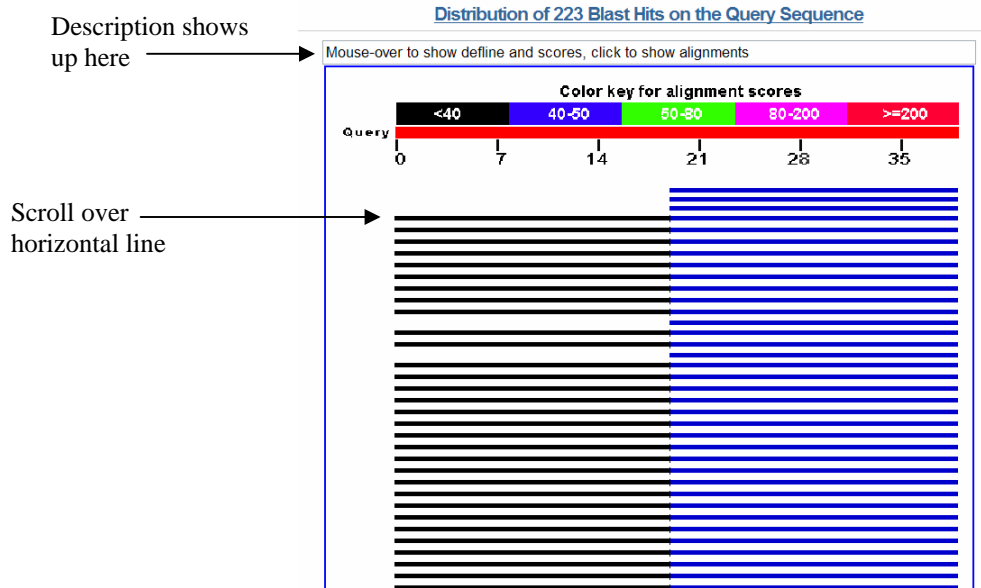
1. Navigate to the NCBI home page at <http://www.ncbi.nlm.nih.gov/>
2. Click on “**BLAST**” at the top of the page.
3. Scroll down and click on “**nucleotide blast**”.
4. Enter the following primer sequences into the box at the top of the page under “**Enter accession number, gi, or FASTA sequence**”. You can enter the two primer sequences one after another, or leave space between the two. **Enter DNA sequences only, no words.**

Forward Primer: GCT CCT ACA AAT GCC ATC A

Reverse Primer: GAT AGT GGG ATT GTG CGT CA

5. Under “Database” choose “**Nucleotide collection (nr/nt)**”.
6. Under “Optimize for” choose “**Somewhat similar sequences (blastn)**”.
7. Click “**BLAST**” on the bottom of the page. Your request will then be sent to the server. This may take a few seconds. The page will continue to refresh itself until the results are displayed.
8. Once your results are displayed, scroll down to “**Distribution of BLAST Hits on the Query Sequence**” (the **query** is the sequence you enter in the search, in this case, the query is our primer sequences). The **hits** represent all the matches in the database that are similar in sequence to your query, arranged from highest to lowest similarity.

9. As you move your mouse over the hits (represented by horizontal lines) you will see a written description of the hit sequence at the top of the box, as well as an **E-value** (Expect value). The E-value represents the number of hits in the database that are expected to match *just by chance*. Therefore, the *smaller* the E-value, the smaller the chance that your match is just random, and therefore, the *more likely* the two sequences are related to each other (i.e. smaller E-value = more similar).



10. Focus on the first match that is a **complete line** (see arrow above for example). If you continue to scroll down the page you will see a written list of all the hits that match your query. You can then click on the **accession number** for that match for a more detailed description of the sequence. This match indicates the *type* of DNA sequence that will be amplified by the primers during the PCR reaction. You'll notice that many sequences are a perfect match to your primers. Click on a few more accession numbers to get a better idea of the type of sequence that is amplified by your primers.

11. Based on the written descriptions of the matches, answer the following questions:

1. What *type* of DNA sequence do you think will be amplified by your primers?

---



---



---

2. Does the amplified sequence represent DNA found in the plant, or does it represent foreign DNA (i.e. DNA not naturally present in the plant genome)? What word(s) in the sequence descriptions make you think this?

---



---



---

3. Suppose your soy plant is *not* genetically modified. Would you expect a product in the PCR reaction? Why or why not?

---



---



---

12. If you clicked on an accession number, hit the back button to return to the main results page. Mouse over and click on the **first full horizontal line** under the hit distribution list. This will show you the exact alignment of your primer sequences (query) with the matching DNA results (subject). The lowest and highest numbers in the subject represent the borders of the DNA sequence that will be amplified with the particular primers.

13. To calculate the exact length of the expected PCR product, subtract the lowest from the highest number and add one. For example:

```
Score = 48.1 bits (24), Expect = 0.002
Identities = 24/24 (100%), Gaps = 0/24 (0%)
Strand=Plus/Minus
```

```
Query 25 GGAACCCACATCACCACGGTACAT 48
      |||
Sbjct 1097 GGAACCCACATCACCACGGTACAT 1074
```

Highest number

```
Score = 46.1 bits (23), Expect = 0.008
Identities = 23/23 (100%), Gaps = 0/23 (0%)
Strand=Plus/Plus
```

```
Query 2 GGATCCACTTCATGCTTTCGTCC 24
      |||
Sbjct 913 GGATCCACTTCATGCTTTCGTCC 935
```

Lowest number

$$1097 - 913 + 1 = 185$$

Based on your match, what is the predicted length of the PCR product we will generate in our experiment?

Predicted length is \_\_\_\_\_ base pairs

## Teacher Answer Key

### 1. What *type* of DNA sequence do you think will be amplified by your primers?

The highest matching hits are those of plant cloning and expression vectors, also known transformation vectors. Vectors are the means by which foreign DNA is delivered into cells. Because plasmids are often used as vectors, their name will often be preceded by a small letter ‘p’ (for example, pYW3310).

In their PCR lab, students will amplify a sequence from the vector that was used to create the genetically modified soybeans. The particular section of the vector that is amplified is the **35s promoter**, a sequence that is very commonly found in vectors used to transform or genetically modify organisms.

### 2. Does the amplified sequence represent DNA found in the plant, or does it represent foreign DNA (i.e. DNA not naturally present in the plant genome)? What word(s) in the sequence descriptions make you think this?

Foreign DNA. Expression or cloning vectors are delivered or “transformed” into plant cells by various methods, and carry with them foreign DNA. They are therefore not natural to the plant. The particular foreign DNA sequence that your students will amplify is the 35s promoter, a sequence that originates from a virus.

Key words: cloning, vector, transformation, plasmid.

### 3. Suppose your soy plant is *not* genetically modified. Would you expect a product in the PCR reaction? Why or why not?

No, you would not expect to see any PCR product if your plant was non-GMO. The primer sequences attach to a *foreign DNA sequence* delivered by a plant vector, and therefore is present only in genetically modified plants. If there is no foreign DNA sequence, there is nowhere for the primers to attach, and therefore, no DNA amplification occurs.

**Predicted length of PCR product should be 195 base pairs.** Student answers may vary here. Have the class come to a consensus, and demonstrate where and how to calculate the length if needed.