Computational Biology: DNA Sequencing and Inherited Traits

Time
3-4 class periods

Level
Grade 11-12, regular or AP biology

Purpose
This lesson provides an authentic computational biology experience for students. The students use computational tools to support their STEM learning; in this case evaluating evolutionary relationships between species.

Overview
The students learn about computational biology careers and use professional computational biology tools to carry out DNA sequencing experiments.

Student Outcomes

Learner Objectives:
After this activity, students should be able to:
- Know that scientists gain understanding about genetic sequences by looking for similar sequences both across and within species.
- Understand the challenges that mutations pose in searching a genome.
- Use the BLAST tool to search for a genetic sequence or compare two sequences.

Computational Thinking in STEM Skills:
- Using computational tools to analyze data.
- Using computational tools to visualize data.
- Applying algorithmic solutions.
- Understanding the relationships within a system.

Next Generation Science Standards: (to be added soon)

Prerequisites
Students need to be familiar with the basics of DNA: they should know that DNA consists of 4 bases, A, T, C, and G, and that the arrangement of these bases codes for genes. Some understanding of algorithms would be helpful but is not required.
Background
Gene sequencing is all over the news these days. It wasn't too long ago that scientist sequenced the human genome for the first time. Now, there are many genomes that have been sequenced, and people are both excited and terrified by the scientific advances that this might bring. Many news reports on this subject are somewhat misleading, however, in that they imply that by knowing the gene sequence, we know everything about the organism. However, a genome sequence is just a collection of letters. Suppose you have a gene sequence like ATTAGGGCGACGCT. What could you say about the organism from which it came? In truth, biologists can’t say any more from that sequence than you can.

In the real world, a sequenced genome isn’t the end or the beginning of the end or even the end of the beginning. We need powerful computational techniques to assemble a genome, find the genes, and determine the functions of those genes.

One way to interpret a genetic sequence is to compare sequences. If we have, for example, a sequence of human DNA and we want to know its function, we might see if there is a similar sequence in an organism we have studied more thoroughly, such as yeast. If the two sequences are similar enough, then it is likely that they have similar function.

Comparing sequences is difficult, however, because DNA mutates. The comparison is done using an important computational biology program, the Basic Local Alignment Search Tool, or BLAST.

Teaching Notes
In this lesson, you help your students work with BLAST, notice the similarities and differences between gene sequences, use BLAST to identify a mystery animal, work with their peers to determine the similarities and differences in genetic sequences between animals in a given Class, and then conduct their own research project using the data and CT tools available. This is done in five phases.

Phase I. Intro

Start by showing the Computer Science and Medicine promotional video from CS Education Week: http://www.youtube.com/watch?v=88A3ZqJSs8
Phase II. Getting familiar with BLAST (through Sickle Cell Anemia)

Students are introduced to BLAST and one of its potential uses (determining the similarity between genes in different species). This is done in the context of scientists trying to find a cure for sickle cell anemia.

Before the Activity/Setup

1. Post on the board
   gk12northwestern.wikispaces.com/Computational+Biology+with+BLAST
   - Provides the links to the relevant websites and files needed for this activity.

With the Students

1. Go through the slides in preSickleCellWorksheet.ppt (intro material).
2. The students work individually or in small groups to complete the Sickle Cell handout.
3. The Extra Credit (question #5) could be used for homework, or as an extension for students who complete the activity faster than others.
4. Go through the slides in PostSickleCellWorksheet.ppt (career highlight)

Special note for Phase II.

1. If short on time, Part I can be done as a class or as homework the night before.

Phase III. Mystery Animal

Students use genetic sequences and BLAST to identify their mystery species. Then, with their 'scientific collaborators', they generate a hypothesis about the relatedness of the species within their group.

Before the Activity/Setup

1. Post on the board
   gk12northwestern.wikispaces.com/Mystery+Animal+DNA+Sequences
   - Provides the relevant files and link needed for this activity.

With the Students

1. The students work individually or in small groups to complete the Mystery Animal handout.
2. For Part I, you will need to assign each student a number between 1 and 30. Ideally everyone has their own number, but fine to repeat if you have more than 30 students.
3. For Part II
   a. Assign each ‘Class’ (e.g., Mammalia) to an area of the room. Place the page containing pictures of that ‘Class’ on a table in that area. Have the students find the ‘Class’ that contains their mystery animal and sit in that area.
   b. Remind the students that every student in the group must write down their hypothesis. It’s OK if they’re the same hypotheses, but they should all record it on their worksheet.

Phase IV. Using Phylogenetic Trees to Evaluate Evolutionary Relationships

Students create a phylogenetic tree for their group’s species to evaluate their evolutionary relationships.

*Before the Activity/Setup*

1. Post on the board
   [gk12northwestern.wikispaces.com/Correctly+Labelled+DNA+Sequences](gk12northwestern.wikispaces.com/Correctly+Labelled+DNA+Sequences)
   - Provides the relevant files and link needed for this activity.

*With the Students*

1. Show slides from PrePhylogeneticWorksheet.ppt (intro material).
2. The students work individually or in small groups to complete the Phylogenetic Trees handout.
3. Once the students complete the worksheet, have them draw their phylogenetic tree on the blackboard.
4. Ask students to share with the class how their prediction(s) compare to what they see in their phylogenetic tree.
5. Show slides from PostPhylogeneticWorksheet.ppt (career highlight).

*Summarize previous phases:* students have learned how to compare DNA sequences to one another by performing multiple sequence alignments, and they have used those alignments to construct phylogenetic trees to infer the evolutionary relatedness among the organisms in their group. They have also used these analyses to answer their research question, and to confirm or reject their hypothesis. They also learned the importance of scientific collaboration, by pooling or sharing their DNA data with other group members.
Phase V. Computational Biology Mini Research Project

Before the Activity/Setup
1. Have the different phylogenetic trees on the blackboard.

With the Students
1. The students work individually or in small groups to complete the research project handout.

Pre-class Preparation
The teacher should be familiar with and able to use the BLAST program available on the NIH website. If the teacher is able to run through the unit before giving it to the students, this should be sufficient.

Materials and Tools
Technology including software and files
Suggested resources or equipment needed
Supplementary documents or handouts

Assessment
The student handouts and the final presentation are used to assess student understanding.

Additional Information
These materials are adapted from curricular materials from the eCSite and from the Bio-ITEST program by the Northwest Association for Biomedical Research.

Handouts begin on following page.
Sickle Cell Anemia
Using BLAST to Determine Gene Similarity

This activity will introduce you to BLAST (Basic Local Alignment Search Tool), an algorithm used by computational biologists to identify common gene sequences across species.

How can we use BLAST in genetic research?

• To Identify a Sample. For example, perhaps you work for the National Wildlife Federation and were onsite during the 2010 Gulf of Mexico oil spill. There were many fish, birds, and marine mammals killed in the spill. Identifying all of these animals when they are coated in oil or are in their juvenile stage can be difficult. Luckily, you can analyze their genes with BLAST to determine the species of each animal.

• To Find Related Species. For example, perhaps you work for the Centers for Disease Control and Prevention (CDC), and your boss just identified a patient with what looks like avian influenza (bird flu). Luckily, you can determine whether the patient really has influenza, and where they may have gotten infected, using BLAST.

Part I. What is Sickle Cell Anemia?

2. Click on the link for “Blood and Lymph Diseases”, then the link for “Anemia, sickle cell.” Answer the following questions:

   a. What causes sickle cell anemia?

   b. What are some of the symptoms?

   c. Which gene is mutated and on which chromosome is that gene?

   d. Are there any beneficial effects to that mutation?
Part II. Let’s learn more about the similarity and differences between humans and rats!

Download the files “humanHBB.txt” and “ratHBB.txt” from http://gk12northwestern.wikispaces.com/HBB+Sequences. Then, follow the instructions below to compare the human and rat sequences:

2. Under the second subject heading “Basic BLAST”, click on “nucleotide blast” to compare DNA sequences.
3. Click on the option that says “Align Two or More Sequences.” This should cause a second window to appear.
4. Under where it says “Enter Query Sequence,” copy and paste the sequence from “ratHBB” into the window.
5. Under where it says “Enter Subject Sequence,” copy and paste the sequence from “humanHBB” into the window.
6. Under where it says ‘Program Selection’, choose the option “Optimize for Somewhat Similar Sequences (blastn).”
7. Press the “BLAST” button at the bottom of the page to run the algorithm.
8. Once the algorithm program finishes, the new page will have information on how well the sequences align (or match up).

Answer the following questions about the alignment between rats and humans HBB gene:

1. Your BLAST search result will include a Pairwise Comparison for each result. Each alignment is a comparison between two sequences: your query (Rat) sequence and the subject (Human) sequence. Look through your Pairwise Comparison, which looks similar to the following:

   The HBB gene sequence you’re working with has over 1,000 nucleotides. Your Query will begin at nucleotide #1, but your Subject result might not. In the example above, the Subject begins with nucleotide #5326 and ends with #46622.

   At what nucleotide does your Subject sequence start? ____________

   At what nucleotide does your Subject sequence end? ______________
2. In the table in the middle of the page, there are 2 columns that are useful to us here:
   
i. **Query coverage:** This is the percent of the query sequence BLAST finds to be a match with the subject sequence. For example, if 75% of the query matched something in the subject with a few mutations, but the other 25% couldn’t be matched up at all, we would have a query coverage of 75%. Higher query coverage indicates a better alignment.
   
ii. **Max ident.:** This is the highest percent of base-pairs that match within a segment of the query sequence that matches up to the subject sequence. For example, say we have the 75% of the query matched as in the previous example, and within that 75%, we have a segment where 90% of the bases match. In this case, our max ident. would be 90%.

What percentage of the human and rat sequences match for this gene (i.e., the query coverage)?

What is the largest percentage of the sequence that matches (i.e., the max ident.)?

3. Scroll down to where the sequences are written out. Write 8 characters of the alignment between the query (i.e., rat) and the subject (i.e., human) that includes at least one base that has mutated into another. (These mutations are called Single Nucleotide Polymorphisms; SNPs):

4. Write 8 characters of the alignment that includes at least one base that has either been deleted from the genome. (These are called indels):

5. By finding treatments that relieve the pain and other symptoms of sickle cell anemia in rats, scientists have made great strides in developing treatments for sickle cell anemia in humans. However, as you saw in your answer to #1, human and rat sequences for the gene that codes for hemoglobin beta (HBB) are not 100% alike. Furthermore, no cure for sickle cell anemia has yet been found. Significant work remains to be done.

*Extra Credit:* As a scientist, what steps would you take to make progress in finding a cure for sickle cell anemia?
Sickle Cell Anemia: Looking at a gene using BLAST
Teacher Answer Key

Answer the following questions:
   a. What causes sickle cell anemia?
      A point mutation in the HBB gene (some students may give the answer a sickle-
      shape in the red blood cells, this is also acceptable).
   b. What are some of the symptoms?
      Pain, anemia, infections.
   c. Which gene is mutated and on which chromosome is that gene?
      Gene HBB on Chromosome 11.
   d. Are there any beneficial effects to that mutation?
      Sickle-cell carriers have a higher resistance to malaria.

Answer the following questions about the alignment between rats and humans HBB gene:
   1. What percentage of the human and rat sequences match for this gene?________
      42% [Note that this is the query coverage column on the BLAST program]
   2. What is the largest percentage of the sequence that matches?_______
      100% [Note that this is the max ident. column on the BLAST program].
   3. Write 8 characters of the alignment that includes at least one base that has
      mutated into another. (These mutations are called Single Nucleotide
      Polymorphisms; SNPs.):
         CACAACTG
         | | |        |  |
         CATAGTTG
   4. Write 8 characters of the alignment that includes at least one base that has either
      been added or deleted from the genome. (These are called indels.):
         CA - - AACA
         | |    | | | | | |
         CAGAAACA

Extra credit: As a scientist, what steps would you take to make progress in finding a cure for
sickle cell anemia?

Students may discuss comparing the Human HBB gene with other species, to see if there is a
closer match. Then using this other species as a better test case. Others may discuss the ethics of
animal testing and discuss the potential for creating complex simulations to address this question.
Others may discuss conducting controlled tests with patients who have sickle cell anemia. Others
may discuss developing computer simulations to better understand how sickle cell anemia works.
There are no wrong answers, as long as they’re on topic.
Mystery Animal
Using Genetic Sequences and BLAST to Identify Species

DNA barcoding is the use of a DNA sequence as a means to identify new species, identify unknown samples, and compare the relatedness and evolution among different species. Certain genes can be used in this manner because some regions of these genes are conserved—that is, they show a very slow rate of evolution and show very little change in their DNA sequence. Others evolve more rapidly and show more changes in their DNA.

A common gene that scientists use to compare animal species is: cytochrome oxidase subunit 1 gene, or COI.

Today, your job as a researcher is to:
1. Use the bioinformatics tool BLAST to identify an unknown DNA barcode sequence.
2. Find your scientific collaborators and generate a hypothesis about the relatedness of the species within your group.

Part I: Identify Your Mystery Animal

Your teacher will assign you a mystery animal number. Write that number here______.

Download the file that corresponds to that number, at http://gk12northwestern.wikispaces.com/Mystery+Animal+DNA+Sequences. The file contains your mystery animal’s COI gene sequence.

Follow the instructions below to identify your mystery animal:

2. Under the second subject heading “Basic BLAST”, click on “nucleotide blast” to compare DNA sequences.
3. Unclick ‘Align Two or More Sequences’.
4. Under where it says “Enter Query Sequence,” copy and paste the sequence from your mystery animal into the window.
7. Click the BLAST button to start your search.
8. Select your first search result, either by:
   a. clicking on the first/top red line
   b. scrolling down and clicking on the first Max Score result.

Nucleotide: The basic building blocks of DNA and RNA: guanine (G), cytosine (C), adenine (A), thymine (T) and uracil (U). Each nucleotide contains a nitrogenous base, a five- carbon sugar, and a phosphate.
9. What is the Description of your first (best) BLAST result? These will be the two or three words next to the green letter ‘D’ (they’re in Latin). Be sure to include the genus, species, and subspecies name (if applicable).

For example, the scientific name for the domestic dog is Canis lupus familiaris.

The Description (scientific name) for YOUR mystery animal is:

_________________________________________________________________

10. What is the common name for your animal (i.e., domestic dog is the common name for Canis lupus familiaris). Hint: If you don’t know Latin, google the scientific name.

_____________________________________________________

11. Look through your Pairwise Comparison, which looks similar to the following:

The COI gene is encoded by the mitochondrial genome, which is over 13,000 nucleotides long. Your Query will begin at nucleotide #1, but your Subject result might not. In the example above, the Subject begins with nucleotide #5326 and ends with #46622.

At what nucleotide does your Subject sequence start? ____________

At what nucleotide does your Subject sequence end? ____________
Part II. Finding your Collaborators and Generating a Hypothesis about Species Relatedness

Scientists often specialize in studying one particular Class, such as reptiles, birds, or mammals. They form collaborations with other scientists who study the same or related species. Your teacher will help you find the other student researchers who are studying animals in your same Class.

Your teacher will also provide you with a handout that contains pictures of your group’s species. You can use these pictures to help you come up with your research hypothesis about how your species are related to one another.

1. Your Group Name (must include your Class; e.g., mammalia):

2. List the names of all of the collaborators in your group and the species they are studying below. Be sure to list yourself as well!

<table>
<thead>
<tr>
<th>Scientist’s Name</th>
<th>Species Studied</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collaborator #1:</td>
<td></td>
</tr>
<tr>
<td>Collaborator #2:</td>
<td></td>
</tr>
<tr>
<td>Collaborator #3:</td>
<td></td>
</tr>
<tr>
<td>Collaborator #4:</td>
<td></td>
</tr>
<tr>
<td>Collaborator #5:</td>
<td></td>
</tr>
<tr>
<td>Collaborator #6:</td>
<td></td>
</tr>
</tbody>
</table>

3. What is your research hypothesis about the relatedness of the species within your group? Specifically, which species are most closely related within your group? Which species are most distantly related?
**Mystery Species Answer Key**

*(The numbers correspond to the mystery number assigned to each student.)*

**Scientific and common name for the species:**

<table>
<thead>
<tr>
<th>Group 1: Class Mammalia (Primates)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Pan troglodytes</td>
<td>Chimpanzee</td>
</tr>
<tr>
<td>2. Gorilla gorilla</td>
<td>Gorilla</td>
</tr>
<tr>
<td>3. Pongo pygmaeus</td>
<td>Orangutan</td>
</tr>
<tr>
<td>4. Trachypithecus obscurus</td>
<td>Dusky Leaf Monkey</td>
</tr>
<tr>
<td>5. Macaca mulatta</td>
<td>Rhesus Monkey</td>
</tr>
<tr>
<td>6. Ateles Geoffroyi</td>
<td>Spider Monkey</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group 2: Class Aves (Birds)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>7. Corvus corax</td>
<td>Raven</td>
</tr>
<tr>
<td>8. Nucifraga caryocatactes</td>
<td>Nutcracker</td>
</tr>
<tr>
<td>9. Eudyptula minor</td>
<td>Little Penguin</td>
</tr>
<tr>
<td>10. Egretta tricolor</td>
<td>Tricolor Egret</td>
</tr>
<tr>
<td>11. Gallus gallus</td>
<td>Chicken</td>
</tr>
<tr>
<td>12. Alectoris chukar</td>
<td>Chukar Partridge</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group 3: Class Osteichthyes (Bony Fishes)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>13. Oncorhynchus kisutch</td>
<td>Coho Salmon</td>
</tr>
<tr>
<td>14. Beryx decadactylus</td>
<td>Splendid Alfosino</td>
</tr>
<tr>
<td>15. Sebastes goodie</td>
<td>Chilipepper or Rockfish</td>
</tr>
<tr>
<td>16. Ostichthys japonicas</td>
<td>Golden Perch</td>
</tr>
<tr>
<td>17. Lutjanus stellatus</td>
<td>Star Snapper</td>
</tr>
<tr>
<td>18. Trachurus japonicas</td>
<td>Japanese Jack Mackerel</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group 4: Class Chondrichthyes (Cartilaginous Fishes)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>19. Carcharhinus dussumieri</td>
<td>Whitecheek Shark</td>
</tr>
<tr>
<td>20. Carcharhinus obscurae</td>
<td>Dusty Shark</td>
</tr>
<tr>
<td>21. Dasyatis bennetti</td>
<td>Bennett’s Stingray</td>
</tr>
<tr>
<td>22. Carcharodon carcharias</td>
<td>Great White Shark</td>
</tr>
<tr>
<td>23. Lamna nasus</td>
<td>Mackerel Shark</td>
</tr>
<tr>
<td>24. Pastinachus solocirostris</td>
<td>Roughnose Stingray</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group 5: Class Reptilia (Reptiles)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>25. Alligator mississippiensis</td>
<td>American Alligator</td>
</tr>
<tr>
<td>26. Iguana iguana</td>
<td>Iguana</td>
</tr>
<tr>
<td>27. Varanus komodoensis</td>
<td>Komodo Dragon</td>
</tr>
<tr>
<td>28. Crocodylus niloticus</td>
<td>Nile Crocodile</td>
</tr>
<tr>
<td>29. Chamaeleo africanus</td>
<td>African Chameleon</td>
</tr>
<tr>
<td>30. Xenagama taylori</td>
<td>Shield-tailed Lizard</td>
</tr>
<tr>
<td>----------------</td>
<td>-------------</td>
</tr>
<tr>
<td><strong>Group 1: Class Mammalia</strong></td>
<td></td>
</tr>
<tr>
<td>1. <em>Pan troglodytes</em></td>
<td>gb</td>
</tr>
<tr>
<td>2. <em>Gorilla gorilla</em></td>
<td>X93347.1</td>
</tr>
<tr>
<td>3. <em>Pongo pygmaeus</em></td>
<td>D38115.1</td>
</tr>
<tr>
<td>4. <em>Trachypithecus obscurus</em></td>
<td>AY863425.1</td>
</tr>
<tr>
<td>5. <em>Macaca mulatta</em></td>
<td>AY612638.1</td>
</tr>
<tr>
<td>6. <em>Ateles geoffroyi</em></td>
<td>ARO16730.1</td>
</tr>
<tr>
<td><strong>Group 2: Class Aves</strong></td>
<td></td>
</tr>
<tr>
<td>7. <em>Corvus corax</em></td>
<td>AYS27236.1</td>
</tr>
<tr>
<td>8. <em>Nucifraga caryocatactes</em></td>
<td>GQ482243.1</td>
</tr>
<tr>
<td>9. <em>Eudyptula minor</em></td>
<td>EUS25395.1</td>
</tr>
<tr>
<td>10. <em>Egretta tricolor</em></td>
<td>DQ437903.1</td>
</tr>
<tr>
<td>11. <em>Gallus gallus</em></td>
<td>AP003323.1</td>
</tr>
<tr>
<td>12. <em>Alectris chukar</em></td>
<td>FJ52425.1</td>
</tr>
<tr>
<td><strong>Group 3: Class Osteichthyes</strong></td>
<td></td>
</tr>
<tr>
<td>13. <em>Oncorhynchus kisutch</em></td>
<td>EF126369.1</td>
</tr>
<tr>
<td>14. <em>Beryx decadactylus</em></td>
<td>AP004430.1</td>
</tr>
<tr>
<td>15. <em>Sebastes goodiei</em></td>
<td>GQ396749.1</td>
</tr>
<tr>
<td>16. <em>Ostichthys japonicas</em></td>
<td>AP004431.1</td>
</tr>
<tr>
<td>17. <em>Lutjanus stellatus</em></td>
<td>EU502687.1</td>
</tr>
<tr>
<td>18. <em>Trachurus japonicas</em></td>
<td>AP003092.1</td>
</tr>
<tr>
<td><strong>Group 4: Class Chondrichthyes (or Class Elasmobranchii)</strong></td>
<td></td>
</tr>
<tr>
<td>19. <em>Carcharhinus dussumieri</em></td>
<td>DQ108305.1</td>
</tr>
<tr>
<td>20. <em>Carcharhinus obscurus</em></td>
<td>FJ518938.1</td>
</tr>
<tr>
<td>21. <em>Dasyatis bennetti</em></td>
<td>EU595098.1</td>
</tr>
<tr>
<td>22. <em>Carcharodon carcharias</em></td>
<td>FJ518943.1</td>
</tr>
<tr>
<td>23. <em>Lamna nasus</em></td>
<td>FJ519722.1</td>
</tr>
<tr>
<td>24. <em>Pastinachus solocirostris</em></td>
<td>EU398977.1</td>
</tr>
<tr>
<td><strong>Group 5: Class Reptilia</strong></td>
<td></td>
</tr>
<tr>
<td>25. <em>Alligator mississippiensis</em></td>
<td>Y13113.1</td>
</tr>
<tr>
<td>26. <em>Iguana iguana</em></td>
<td>AJ278511.1</td>
</tr>
<tr>
<td>27. <em>Varanus komodoensis</em></td>
<td>AB080275.1</td>
</tr>
<tr>
<td>28. <em>Crocodylus niloticus</em></td>
<td>DQ273697.1</td>
</tr>
<tr>
<td>29. <em>Chamaeleo africanus</em></td>
<td>EF222196.1</td>
</tr>
<tr>
<td>30. <em>Xenagama taylori</em></td>
<td>DQ008215.1</td>
</tr>
</tbody>
</table>

**Your Group Name (includes the Class of your species):**

Class Mammalia, Class Aves, Class Osteichthyes, Class Chondrichthyes (or Class Elasmobranchii), or Class Reptilia, with something catchy afterwards, if the students get into it.

**Your research hypothesis about the relatedness of the species within your group:**
Student research hypotheses will vary but should include a prediction about the relatedness between or among at least two species within their group. You may wish to instruct students to include all six species in their hypothesis.
Class Mammalia
Class Aves

Raven  *Corvus corax*

Little Penguin  *Eudyptula minor*

Chicken  *Gallus gallus*

Tricolored Egret  *Egretta tricolor*

Nutcracker  *Nucifraga caryocatactes*

Chukar Partridge  *Alectoris chukar*
Class Osteichthyes (The Bony Fishes)
Class Chondrichthyes (The Cartilaginous Fishes)
Class Reptilia

- Shied-tailed Lizard *Xenagama taylori*
- Iguana *Iguana iguana*
- Komodo Dragon *Varanus komodoensis*
- American Alligator *Alligator mississippiensis*
- African Chameleon *Chamaeleo africanus*
- Nile Crocodile *Crocodylus niloticus*
Evaluating Evolutionary Relationships Using Phylogenetic Trees

Computational biologists use BLAST to create phylogenetic trees (flowcharts that help us visualize the genetic difference between species). In this activity, you will create a phylogenetic tree for your group’s species to evaluate their evolutionary relationships.

1. Copy below your hypothesis for which species are most closely related within your group and which species are most distantly related.

2. Go to gk12northwestern.wikispaces.com/Correctly+Labelled+DNA+Sequences and download the file labeled with the ‘Class’ name for your group.

3. Follow the instructions below to determine how closely the species within your group are related:
   b. Under the second subject heading “Basic BLAST”, click on “nucleotide blast” to compare DNA sequences.
   c. Click on the option that says “Align Two or More Sequences.” This should cause a second window to appear.
   d. Under where it says “Enter Query Sequence,” copy and paste the sequence for the first species in your file (i.e., >chimpanzee) into the window. This will be a species that is NOT in your Class. It serves as a base for comparison.
      NOTE: Be sure that you include the text (>NameOfSpecies) at the beginning.
   e. Under where it says “Enter Subject Sequence,” copy and paste the sequences from all the other species in your file into the window.
      NOTE: Be sure that you include the text (>NameOfSpecies) at the beginning of each sequence.
   f. Under where it says “Program Selection” choose the option “Optimize for Somewhat similar sequences (blastn).”
g. Click the blue plus sign beside the “Algorithm Parameters” menu to adjust the BLAST settings.

h. In the Algorithm Parameters menu, make the following changes:
   i. Uncheck the “Short queries” box.
   ii. Change the “Word size” to 7 in the drop down menu.
   iii. Uncheck the Filter box for “Low complexity regions.”
   iv. Uncheck the Mask box for “Mask for lookup table only.”

i. Press the “BLAST” button at the bottom of the page to run the algorithm.

j. Once the results are posted, near the top of the page, click the link for the “Distance tree of results” next to ‘Other Reports’ near the top.

2. In the space below, draw your phylogenetic tree (the figure that appears after you click on ‘Distance tree of results’). Be sure to write the names of the species beside the appropriate branches.

3. Based on your phylogenetic tree, which species appear to be most closely related to one another (which species cluster closest together)?

4. Which species are most distantly related?

5. Do these data support or contradict your hypothesis?
Using Phylogenetic Trees, Answer Key

Phylogenetic tree results for each group: See figures at: http://gk12northwestern.wikispaces.com/Phylogenetic+Tree+Answers

Based on your phylogenetic tree, which species appear to be most closely related to one another (which species cluster closest together)?

*Group 1:* Gorillas and Chimpanzees are the most closely related, followed by Orangutans.

*Group 2:* Chickens and Partridges appear closely related, as do Nutcrackers and Ravens.

*Group 3:* Chilipepper and Snapper appear to be the most closely related, followed by Perch, Splendid Alfonsino, and Mackerels.

*Group 4:* Both the Dusky Shark and White Cheek Sharks are closely related, as are the Great White and Mackerel Sharks. Both of the Stingrays are closely related to each other, but appear more distantly related to the Sharks.

*Group 5:* The Nile Crocodile and American Alligator are closely related, followed by the Iguana and Komodo Dragon.

Which species are most distantly related?

*Group 1:* The Rhesus Macaque, the Leaf Monkey and the Spider Monkey are most distantly related.

*Group 2:* Egrets and Penguins are the most distantly related.

*Group 3:* Coho Salmon are the most distantly related.

*Group 4:* Both of the Stingrays are closely related to each other, but appear more distantly related to the Sharks.

*Group 5:* Chameleons and Lizards appear to be the most distantly related.

Do these data support or contradict your hypothesis?

*In this question, students refer back to their original research question and hypothesis, and then evaluate their hypothesis in the context of their data analysis above. For instance, if their hypothesis was that Chimpanzees and Spider Monkeys were more closely related than Chimpanzees and Gorillas, they would reject their hypothesis.*
Name:_________________________________

Computational Biology Mini Research Project:
Formulate a Question, Pursue Evidence, and Justify Your Conclusions

1. Look at the phylogenetic trees on the board. You have access to the genetic sequences for all of those species at the website your teacher made available.

You also have access to the tools available through BLAST, at http://blast.ncbi.nlm.nih.gov/Blast.cgi

2. Your task is to design an answerable research question using these data, propose a plan to pursue evidence, collect data using the tools described above, and create and evidence-based conclusion about evolutionary relationships.

3. The end goal is to create a poster with the following:
   a. Specific research question
   b. Step-by-step procedure, with sketches if needed, to collect evidence
   c. Figures/tables/visualizations of your results
   d. Evidence-based conclusion statement