

2

DNA Barcoding and the Barcode of Life Database (BOLD)

Introduction

In this lesson, students will receive an “unknown” DNA sequence and use the bioinformatics tool Basic Local Alignment Search Tool (BLAST) to identify the species from which the sequence came. Students then visit the Barcode of Life Database (BOLD) to obtain **taxonomic** information about their species and form taxonomic groups for scientific collaboration. The lesson ends with each student generating a hypothesis about the relatedness of the species within each group. In *Lesson Two*, students also learn how **postdoctoral scientists in DNA and history** might use bioinformatics tools in their career.

Learning Objectives

At the end of this lesson, students will know that:

- Bioinformatics tools are used by people in many different career fields, including postdoctoral scientists in DNA and history.
- Bioinformatics tools such as BLAST and databases like the National Center for Biotechnology Information (NCBI) and the Barcode of Life Database (BOLD) can be used to identify unknown DNA sequences and obtain information about the species from which the sequences came.
- Scientific names for species, including the **genus** and **species** names, can be used to classify species based on evolutionary relatedness.
- Scientists from around the world compile their data into databases such as those at the NCBI and BOLD to encourage scientific collaboration and increase scientific knowledge.

At the end of this lesson, students will be able to:

- Perform a BLAST search to identify an unknown DNA sequence.
- Find a species in the Barcode of Life Database (BOLD) to determine whether that species has been barcoded and to obtain taxonomic information about that species.

Key Concepts

- The bioinformatics tool BLAST can be used to identify unknown DNA sequences.
- A particular gene or region of DNA called cytochrome c oxidase subunit 1 (*COI*) can be used to identify and catalog animals in a process called DNA barcoding.
- The exact sequence of the *COI* gene is unique, with very little variability within a species.
- Differences in the *COI* gene across species can be used to identify unknown samples and perform evolutionary analyses.
- Scientific research is a collaborative process and scientists from around the world make their data available in public databases for use by other scientists.

Class Time

1 class period of 50 minutes if Student Reading is assigned as homework; otherwise 2 class periods of 50 minutes each.

Prior Knowledge Needed

- DNA contains the genetic information that encodes traits.
- Basic knowledge of taxonomy (specifically the different categories used in taxonomy to classify organisms, and that the study of taxonomy seeks to reflect the evolutionary history of different organisms).
- Binomial nomenclature (i.e., the use of genus and species to refer to individual species).

Materials

Materials	Quantity
Copies of Student Handout— <i>Careers in the Spotlight</i> (handed out in <i>Lesson One</i>)	1 per student
Copies of Student Handout— <i>The Process of Genetic Research</i> (handed out in <i>Lesson One</i>)	1 per student
Class set of Student Handout— <i>Using BLAST and BOLD for Genetic Research Instructions</i>	1 per student (class set)
Copies of Student Handout— <i>Using BLAST and BOLD for Genetic Research Worksheet</i> [Note: This worksheet is for students' answers to lesson questions.]	1 per student
Copy of Student Group 1 Handout— <i>Class Mammalia</i> (ideally printed in color and placed in a plastic page protector)	1
Copy of Student Group 2 Handout— <i>Class Aves</i> (ideally printed in color and placed in a plastic page protector)	1
Copy of Student Group 3 Handout— <i>Class Osteichthyes</i> (or <i>Actinopterygii</i>) (ideally printed in color and placed in a plastic page protector)	1
Copy of Student Group 4 Handout— <i>Class Chondrichthyes</i> (or <i>Elasmobranchii</i>) (ideally printed in color and placed in a plastic page protector)	1
Copy of Student Group 5 Handout— <i>Class Reptilia</i> (ideally printed in color and placed in a plastic page protector)	1
Teacher Answer Key— <i>Using BLAST and BOLD for Genetic Research</i>	1
Plastic page protectors (optional)	5

Computer Equipment, Files, Software, and Media

Computer and projector to display PowerPoint slides.

Alternative: Print PowerPoint slides onto transparencies and display with overhead projector.

Lesson Two PowerPoint Slides—*DNA Barcoding and the Barcode of Life Database*.

Available for download at: <http://www.nwabr.org/curriculum/advanced-bioinformatics-genetic-research>.

One or more copies of the article "Scanning Life." Kunzig, Robert. Scanning Life. *National Geographic*. May 2010.

Available at: <http://ngm.nationalgeographic.com/big-idea/11/biodiversity>.

"Unknown DNA Sequences" 1–30. Available from the Bio-ITEST website under the "Resources" tab at:

<http://www.nwabr.org/curriculum/advanced-bioinformatics-genetic-research>.

A student version of lesson materials (minus Teacher Answer Keys) is available from NWABR's Student Resource Center at:

<http://www.nwabr.org/students/student-resource-center/instructional-materials/introductory-bioinformatics-genetic-research>.

Computer lab with internet access and a simple word processing program such as Microsoft® Notepad or TextEdit.

[Note: Use of Microsoft® Word is not recommended when performing bioinformatics analyses.]

Optional: Online video, "Barcode of Life: Global Biodiversity Challenge," produced by the International Barcode of Life Project Biodiversity Institute of Ontario. The video is 2 minutes and 44 seconds long. Available at: <http://www.youtube.com/watch?v=HU1ru9H332s>.

Teacher Preparation

- Load the classroom computer with the *Lesson Two* PowerPoint slides and the "Barcode of Life" video (optional).
- Make copies of the Student Handout—*Using BLAST and BOLD for Genetic Research Instructions*, one per student. This handout is designed to be re-used as a class set.

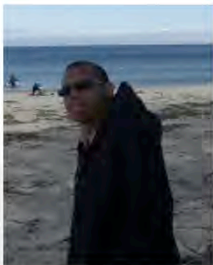
- Make copies of Student Handout—*Using BLAST and BOLD for Genetic Research Worksheet*, one per student. This worksheet is for students to write their answers to the lesson questions. Answers may also be written in students' lab notebooks.
- Make copies of the *Student Group Handouts*. If possible, print these in color; they can be re-used for multiple classes. Placing the *Student Group Handouts* in plastic page protectors will make them last longer.

Procedure

Warm Up

1. As students enter the classroom, display the PowerPoint slides for *Lesson Two*, beginning with **Slide #1**. This slide highlights postdoctoral scientist Krishna Veeramah, PhD, in the DNA and History program at the University of California, Los Angeles.

Postdoctoral Scientist, DNA and History
 KRISHNA VEERAMAH, PhD



Place of Employment:
University of California, Los Angeles (UCLA)

Type of Work:
Human population genetics, with emphasis on Sub-Saharan Africa

I always liked genetics, but it wasn't my first choice at school. I wanted to be a footballer (or soccer player), but... working with all these clever things like Y chromosomes possibly descended from the brother of Moses, I eventually started doing my own work and got wrapped into it.

DNA Barcoding and BOLD: **Slide #1**

2. Have students retrieve Student Handout—*Careers in the Spotlight*, which they were given during *Lesson One*.
3. Students should think about, and write down, the kind of work they think a postdoctoral fellow in DNA and History might do (*Postdoctoral Fellow in DNA and History Question #1*). This will be revisited at the end of the lesson, including how a postdoctoral fellow in DNA and history might use bioinformatics in his or her job.
4. Ask students to keep their *Careers in the Spotlight* Handout available for future lessons.

PART I: Barcoding Life—Understanding the Barcode of Life Project

5. Explain to students the **aims of this lesson**. Some teachers may find it useful to write the aims on the board.
 - a. **Lesson Aim:** Introduce students to DNA barcoding.
 - b. **Lesson Aim:** Understand the role of databases in organizing and searching for biological information, including identification of an unknown DNA sequence.

Voucher specimen: A specimen that serves as a basis of study and is retained as a reference.

Teachers may also wish to discuss the *Learning Objectives* of the lesson, which are listed at the beginning of this lesson plan.

6. Distribute copies of the “Scanning Life” article from *National Geographic* (see the *Materials* section for website information) or, using the classroom projector, read the article together as a class. Allow about ten minutes for students to read the article individually or as a class before discussing.

7. Lead a discussion with students about the Barcode of Life Project. Questions to ask include:

- What is the Barcode of Life Project?

Possible answer: *The Barcode of Life Project seeks to catalog all life on earth by sequencing a small region of DNA (or “barcode”) that can be used to identify species.*

- Why do scientists want to “barcode” all life on earth?

Possible answers:

- To help with conservation efforts (to know which species are found in which habitats).*
- To identify and stop illegal trading (such as ivory from protected species).*
- To make sure consumers are safe, and that foods are properly labeled, particularly when you only have a small part of the animal, and can’t identify it otherwise (such as fish or sushi served in a restaurant).*
- To identify life stages of animals that can be hard to identify otherwise (such as the larval stage of insects, amphibians, and fish).*

- What makes a gene a good choice for barcoding?

Possible answers: *It needs to have enough variation to be different between species; it needs to be found in all species.*

- What are some of the ways that scientists identify species now?

Possible answers: *By studying many genes (such as ribosomal RNA), morphological data, fossil records, and other information about habitat, ecological niche, and diet. Scientists often compare specimens to known voucher specimens in museums or universities.*

- Should barcoding be the **only** tool scientists use to identify species?

Possible answer: *DNA barcoding is likely not a solution for all studies involving species identification and evolution. To get a more complete picture, multiple techniques should probably be used, but DNA barcoding is a very useful tool.*

8. **Optional:** Show students the online video “Barcode of Life: Global Biodiversity Challenge.” This 2:44 minute video introduces the Barcode of Life Project. Available at:

Barcode of Life: Global BioDiversity Challenge Video (2:44 minutes)

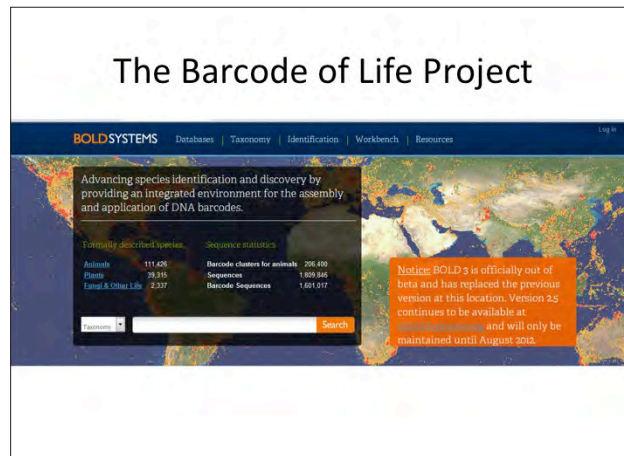
<http://www.youtube.com/watch?v=HU1ru9H332s>

9. Show **Slide #2**, which is a screen capture image of the Barcode of Life Data Systems homepage. Explain to students that the Barcode of Life project seeks to catalog all life on earth with a unique **DNA barcode**. Like a barcode on an item from the grocery store, a barcode is a **unique identifier** that can

distinguish one species from another. Among animals, the gene chosen for barcoding is the **cytochrome c oxidase subunit 1** gene (also called “*COI*”), which is found in the mitochondrial genome. The **mitochondrial genome** encodes some of the proteins needed to make adenosine triphosphate (ATP).

Mitochondrial genome: All of the genes found in the mitochondria (in contrast to the genes found in the nucleus).

DNA Barcoding and BOLD: **Slide #2**



10. Show **Slide #3**, which describes the cytochrome c oxidase subunit 1 (*COI*) gene that is used for DNA barcoding. Review the characteristics of *COI* that make it a good choice as a gene for DNA barcoding, including:

- It is present in all the species we want to study – so we can make comparisons among species, such as among dogs, wolves, coyotes, and foxes.
- It is variable, but not too variable – the ideal gene to study would be fairly similar **within a species** but would **contain DNA differences between species** so we can differentiate between species.
- It is standardized among scientists around the world – if many groups of scientists agree to study the same gene, it can make comparing results much easier.
- It can be used to explore questions about genetic relatedness and evolutionary relationships.

DNA Barcoding and BOLD: **Slide #3**

DNA Barcoding with Cytochrome c Oxidase subunit 1 (*COI*)

The Ideal Gene to Study is...

- Present in all species.
- Variable, but not too variable.
- Standardized among scientists around the world.

The *COI* Gene...


- All eukaryotes contain mitochondria; *COI* encodes a mitochondrial protein needed for cells to make ATP.
- *COI* is almost identical within a species but varies between different species.
- Agreement among scientists that the *COI* gene is used for animal barcoding.

PART II: Using the Bioinformatics Tool BLAST (Basic Local Alignment Search Tool)

11. Show **Slide #4**, which is a screen shot of the BLAST homepage, with some basic information about how BLAST can be used.

DNA Barcoding and BOLD: **Slide #4**

Basic Local Alignment Search Tool - BLAST



- Can be used to:
 - Compare 2 or more DNA sequences to each other.
 - Compare 2 or more protein sequences to each other.
 - Compare a single sequence to many sequences found in databases.
- You will perform a **nucleotide BLAST** (blastn) search to find the best possible match for your unknown sequence.

12. Share the following information with students in preparation for the activities they will perform on their own in Student Handout—*Using BLAST and BOLD for Genetic Research*:

• Collaboration:

Tell students that much of what is known about science is shared through the process of scientific collaboration, with many research scientists depositing their findings and data in central repositories like the Barcode of Life Database (BOLD) and the National Center for Biotechnology Information (NCBI). There are many species to study, and it is far too much work for one scientist to do alone!

• Species Identification:

Remind students that DNA barcoding is particularly useful in identifying species. One can search databases like the ones at the NCBI and BOLD to identify unknown DNA sequences.

• Phylogenetics:

Tell students that the lessons they will be studying in the next few days involve species identification and **phylogenetics** (“phylo” meaning “tribe” or “race” and “genetics” meaning “relative to birth”), which is the study of evolutionary relatedness among organisms.

• BLAST is a bioinformatics tool to compare sequences:

Remind students that one of the bioinformatics tools they used in the Bio-ITEST Introductory lessons, *Using Bioinformatics: Genetic Testing*, is called **BLAST** – **B**asic **L**ocal **A**lignment **S**earch **T**ool. BLAST can be used to compare the sequences of two or more proteins or nucleic acid molecules, or to compare a single sequence to a collection of sequences in a database.

Phylogenetics: The study of evolutionary relationships among organisms.

[**Note:** There are additional programs in the BLAST family (blastx, tblastn, and tblastx) but these are beyond the scope of the curricula and will not be addressed. Additional information about BLAST can be found in the *Appendix*.]

• **Nucleotide BLAST:**

Students will be performing a Nucleotide BLAST (blastn) search, to identify an unknown DNA barcode sequence. BLAST will use their unknown sequence (called the **query**) to search through the entire collection of **nucleotide** sequences at the NCBI to find the best possible matches (called **subjects**). Searches can also be performed with protein sequences, using Protein BLAST (blastp).

• **Alignments:**

The results of a BLAST search are in the form of alignments, which note regions of similarity and regions of difference, as well as scores, which assign numerical values to the alignments and other statistics that describe characteristics of the alignments and look at the probability of a random match.

13. Assign students a number (1–30) and tell them to visit the Bio-ITEST website to retrieve the unknown DNA sequence corresponding to their assigned number. The URL for the Bio-ITEST website is the same as that noted above in the *Materials* section: (<http://www.nwabr.org/curriculum/advanced-bioinformatics-genetic-research>).

Tell students their assigned number, but not the name of the corresponding animal species; it is their task to discover the name of their animal species.

The numbers and their corresponding common names are listed below:

1. Gorilla	11. Chicken	21. Bennett Stingray
2. Chimpanzee	12. Chukar Partridge	22. Great White Shark
3. Orangutan	13. Coho Salmon	23. Mackerel Shark
4. Dusky Leaf Monkey	14. Alfonsino	24. Roughnose Stingray
5. Rhesus Macaque	15. Chilipepper or Rockfish	25. American Alligator
6. Spider Monkey	16. Japanese Soldierfish	26. Iguana
7. Raven	17. Star Snapper	27. Komodo Dragon
8. Nutcracker	18. Japanese Jack Mackerel	28. Nile Crocodile
9. Little Penguin	19. Whitecheek Shark	29. African Chameleon
10. Tricolor Heron	20. Dusky Shark	30. Shield-tailed Agama

14. Pass out Student Handout—*Using BLAST and BOLD for Genetic Research* and have students work through the handout individually or in small groups, at computers.
15. Tips and trouble-shooting for BLAST and BOLD searches:

- **Problem: Errors when performing BLAST searches.** When performing BLAST searches, be sure students have selected **Nucleotide BLAST (blastn)**, not Protein BLAST, blastx, tblastn, or tblastx, as shown on Student Handout—*Using BLAST and BOLD for Genetic Research*, **Figure 1**. Protein BLAST and tblastn cannot be performed with nucleotide sequences and will result in errors. Blastx and tblastx will return protein sequence results, not nucleotide results, as the search functions translate DNA sequences into protein sequences.

Query: When searching databases like those at the NCBI, your “query” is the sequence you are searching with and trying to match. In this case, your query is your unknown sequence.

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 group.

Subject: When searching the databases at the NCBI, the subject sequences are sequences from the database that match the query. In this case, if a subject sequence is identical to your query, your query sequence probably came from the same creature that contributed the subject sequence.

Taxonomic: Taxonomy is the science of classifying organisms into groups based on similarities in their physical characteristics. The goal of taxonomy is to organize species in groups based on their evolution and common ancestors. The word comes from the Greek “taxis” which means “order” or “arrangement.”

- **Problem: Only human sequence results are returned.** Be sure students select the “Nucleotide collection (nt/nr)” option under “**Other Databases**” from the BLAST search page. The default setting is for human and mouse sequences.
- **Problem: Could not find species in BOLD.** Be sure to use the scientific name (genus and species) for each species when searching the Barcode of Life Database (BOLD). Common names are unlikely to yield results.
- **Problem: Species did not match Teacher Answer Key.** Be sure that students selected all of the DNA sequence(s), either by going to the **Edit** menu and choosing **Select All** and then **Copy**, or using the corresponding keyboard shortcut commands.

PART III: Forming Collaborative Groups and Developing a Research Hypothesis

In this portion of the lesson, students will work in small groups to answer one of the five scientific questions (related to their **taxonomic** groups or classes) for which the Bio-ITEST curriculum has previewed the available data.

16. Once students have identified their DNA sequences using BLAST and have obtained a barcode sequence and taxonomic information using BOLD, help them find their scientific collaborators (i.e., other students who are studying related species). Write the following information on the board:

- **Group 1:** Class Mammalia (Mammals)
- **Group 2:** Class Aves (Birds)
- **Group 3:** Class Osteichthyes or Actinopterygii (Bony Fishes)
- **Group 4:** Class Chondrichthyes or Elasmobranchii (Cartilaginous Fishes)
- **Group 5:** Class Reptilia (Reptiles)

17. Post *Student Group Handouts 1–5* around the room. Have students gather at the handout that represents their group to find their collaborators. Once the group has formed, students should take the handout and sit with their group. Each handout contains pictures of the species within each group to help students make predictions about evolutionary relatedness.

18. Within their groups, tell students to fill out their group name, their collaborators’ names, and the common names of the species they are studying in *Part IV* of Student Handout—*Using BLAST and BOLD for Genetic Research*.

Key to Student Groups:

Group 1: Class Mammalia (Mammals)

Species: Chimpanzee, Gorilla, Orangutan, Dusky Leaf Monkey, Rhesus Macaque, Spider Monkey

Group 2: Class Aves (Birds)

Species: Raven, Nutcracker, Little Penguin, Tricolor Heron, Chicken, Chukar Partridge

Group 3: Class Osteichthyes or Actinopterygii (Bony Fishes)

Species: Coho Salmon, Alfonsino, Chilipepper (or Rockfish), Japanese Soldierfish, Star Snapper, Japanese Jack Mackerel

Group 4: Class Chondrichthyes or Elasmobranchii (Cartilaginous Fishes)

Species: Whitecheek Shark, Dusky Shark, Bennett Stingray, Great White Shark, Mackerel Shark, Roughnose Stingray

Group 5: Class Reptilia (Reptiles)

Species: American Alligator, Iguana, Komodo Dragon, Nile Crocodile, African Chameleon, Shield-tailed Agama

19. On the board, beside each group name, write the **Research Questions** below, or show **Slide #5**.

Which of the following animals do you predict to be most closely related?

Group 1: Which species of primates are most closely related?

Group 2: Which species of birds are most closely related?

Group 3: Which species of fish are most closely related?

Group 4: Which species of sharks and rays are most closely related?

Group 5: Which species of reptiles are most closely related?

Predicting Evolutionary Relationships

Which of the following animals do you predict to be most closely related and why?

Group 1: Which species of primates are most closely related?

Group 2: Which species of birds are most closely related?

Group 3: Which species of fish are most closely related?

Group 4: Which species of sharks and rays are most closely related?

Group 5: Which species of reptiles are most closely related?

DNA Barcoding and BOLD: **Slide #5**

20. Ask students what these questions have in common. Teachers are encouraged to draw parallels (as appropriate) to other lessons students have studied. Possible answers include:

- One group of organisms is being compared to other groups of organisms.
- Each group of organisms has some physical characteristics in common (e.g., sharks and rays have gills and swim in the ocean).
- Each question deals with the evolution of a group of organisms (i.e., genetic relatedness implies common ancestry).
- These are all questions that can be answered with genetic research.

21. Ask students to think about how the species in their group may be related, and to write their **hypothesis** (prediction) about these relationships in the space provided on *Part IV* of Student Handout—*Using BLAST and BOLD for Genetic Research*. Students are encouraged to use the handouts with pictures of the organisms within their group to inform their predictions, and to discuss these organisms within their group. Teachers may wish to have students choose which two species they think are most closely related, or include multiple or all

six species in their hypothesis. For example, Group 1 students may hypothesize that all primates with tails are more closely related to each other than to the primates without tails. Emphasize that each group member should have his own hypothesis, and that these hypotheses may be very similar to one another, or may be different or contradict one another. That is part of the scientific process!

Closure

22. Summarize today's lesson:

- Students have learned how to use the bioinformatics tool BLAST to identify an unknown DNA sequence and find the scientific name of the species from which the sequence came. They have also learned how to locate species in the Barcode of Life Database (BOLD), and download and format DNA barcode sequences for bioinformatics analysis and for scientific collaborations with other student researchers studying related species.
- In the next lessons, they will pool their DNA sequence data into a data set and study the evolutionary relationships of the animals in their group. Then they will use that analysis to address their research hypotheses.

23. Ask students to fill out the section about *Lesson Two* in Student Handout—*The Process of Genetic Research*, which was handed out during *Lesson One*. Students could also answer these questions in their lab notebooks:

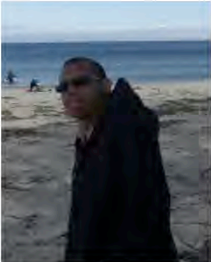
- What **did you do** in this lesson?
- **Methods:** What bioinformatics tool(s) and/or database(s) did you use?
- **Results & Conclusions:** What did you find? What could you conclude from your analysis?
- What **skills** did you learn or practice?
- What is your **Research Hypothesis** about the relatedness of the species within your group?

24. Lastly, return to the picture of the postdoctoral scientist in DNA and history, shown in **Slide #6**.

DNA Barcoding and BOLD: **Slide #6**

Postdoctoral Scientist, DNA and History

KRISHNA VEERAMAH, PhD



Place of Employment:
University of California, Los Angeles (UCLA)

Type of Work:
Human population genetics, with emphasis on Sub-Saharan Africa

I always liked genetics, but it wasn't my first choice at school. I wanted to be a footballer (or soccer player), but... working with all these clever things like Y chromosomes possibly descended from the brother of Moses, I eventually started doing my own work and got wrapped into it.

25. Show **Slide #7**, which provides job information for a postdoctoral scientist. Review this information with students.

CAREERS IN SPOTLIGHT:
Postdoctoral Scientist, DNA and History

What do they do?
Postdoctoral Scientists are people with a PhD who desire more training. The DNA and History program at UCLA is unique – a result of the dramatic increase in DNA data and biotechnology. The increase in DNA data influences research beyond biology, into the humanities and social sciences, like the history of human populations.

What kind of training is involved?
Dr. Veeramah has a PhD and is doing additional training and teaching at UCLA.

What is a typical salary for a Postdoctoral Scientist, DNA and History?
The minimum salary for a Postdoctoral Scientist funded by the National Institutes of Health (NIH): Starting salary: \$37,740 (\$18/hour). 5 years of experience: \$47,940 (\$24/hour).

Source: National Institute of Health

26. Ask students, “What more do we know about studying DNA and history after today’s lesson?” Point out that people studying DNA and history ask questions similar to those students have asked today, including:
- Which human populations alive today are most closely related to one another?
 - What can we learn about ancient civilizations based on genetic data?
 - How can knowledge gained from DNA and genetics inform what we already know about populations and the different languages they speak?
27. Ask students to answer Postdoctoral Scientist in DNA and History *Question #2* on their *Careers in the Spotlight* Handout, which has students explain how this lesson has changed their understanding of the kind of work a postdoctoral scientist in DNA and history does.
28. Ask students to also answer Postdoctoral Scientist in DNA and History *Question #3* on their *Careers in the Spotlight* Handout, which has students explain how a postdoctoral scientist in DNA and history might use bioinformatics in his or her work.
29. Tell students to keep their *Careers in the Spotlight* Handout available for future lessons.

Homework

- A. As homework, ask students to write about the activities they learned in *Lesson Two* in their lab notebooks, on another sheet of paper, or in a word processing program like Microsoft® Notepad or Word which they then provide to the teacher as a printout or via email. This can serve as an entry ticket for the following class. Have them complete these prompts:
- Today I learned that...
 - An important idea to think about is...
 - Something that I don’t completely understand yet is...
 - Something that I’m really confident that I understand is....
- B. The *Lesson Two* section of Student Handout—*The Process of Genetic Research* could also be assigned as homework.

Glossary

Accession number: A unique identifier or code assigned to every entry in the National Center for Biotechnology Information (NCBI) databases. This unique code can be used to search the databases to find your gene or protein of interest.

Conserved: DNA or protein sequences are said to be “conserved” if the sequences are the same or very similar.

Mitochondrial genome: All of the genes found in the mitochondria (in contrast to the genes found in the nucleus).

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 group.

Phylogenetics: The study of evolutionary relationships among organisms.

Query: When searching databases like those at the NCBI, your “query” is the sequence you are searching with and trying to match. In this case, your query is your unknown sequence.

Subject: When searching the databases at the NCBI, the subject sequences are sequences from the database that match the query. In this case, if a subject sequence is identical to your query, your query sequence probably came from the same creature that contributed the subject sequence.

Taxonomic: Taxonomy is the science of classifying organisms into groups based on similarities in their physical characteristics. The goal of taxonomy is to organize species in groups based on their evolution and common ancestors. The word comes from the Greek “taxis” which means “order” or “arrangement.”

Voucher specimen: A specimen that serves as a basis of study and is retained as a reference.

Resources

For more information about the **Barcode of Life Project**, visit the Barcode of Life project website at: <http://www.barcodeoflife.org/>.

Additional readings about **DNA barcoding** are freely available online from The Scientist. Free registration may be required.

Grant, B. Cataloging life: Can a single barcode of DNA record biodiversity and keep us safe from poisons? *The Scientist*. 2007; 21: 36. <http://www.the-scientist.com/2007/12/1/36/1/>

Weir, K. Shark Rx. *The Scientist*. 2007; 21: 24. <http://classic.the-scientist.com/article/display/53521/>

The Food and Drug Administration (FDA) has recently approved DNA barcoding to identify both **domestically processed and imported seafood**.

US Food and Drug Administration. (October 28, 2011). DNA-based Seafood Identification. <http://www.fda.gov/Food/FoodSafety/Product-SpecificInformation/Seafood/DNAspeciation/default.htm>

McGuirk, R. “Restaurants plan DNA-certified premium seafood.” *Associated Press*. November 27, 2011. <http://news.yahoo.com/restaurants-plan-dna-certified-premium-seafood-121222097.html>

Schwartz, J. “Fish tale has DNA hook: Students find bad labels.” *The New York Times*. August 21, 2008. <http://www.nytimes.com/2008/08/22/science/22fish.html>

Credit

Kunzig, Robert. “Scanning Life.” *National Geographic*. May 2010. Available at: <http://ngm.nationalgeographic.com/big-idea/11/biodiversity>

Veeramah, Krishna. Personal Interview. 19 March 2010.

The authors wish to thank Wikimedia Commons, NCBI BLAST, and Barcode of Life Datasystems for many of the images found in the PowerPoint slides and handouts associated with this lesson.

2 Using BLAST and BOLD for Genetic Research Instructions

Student Researcher Background:

DNA Barcoding, BLAST, and BOLD

DNA barcoding is the use of a standardized DNA sequence as a means to identify new species, identify unknown samples, and compare 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 very little change in their DNA sequence—unlike other regions that evolve more rapidly and show more changes in their DNA. The gene that scientists in the Barcode of Life community have decided to use for animal species is the cytochrome c oxidase subunit 1 gene, or *COI*.

DNA barcodes, if they are known, can be accessed from several databases, including the National Center for Biotechnology Information (NCBI) and the Barcode of Life Database (BOLD).

Conserved: DNA or protein sequences are said to be “conserved” if the sequences are the same or very similar.

Aim: Today, your job as a researcher is to:

1. Use the bioinformatics tool **BLAST** to identify the source of an unknown DNA barcode sequence.
2. Find important **taxonomic information** for your species from **BOLD**.
3. Find your **scientific collaborators** to discuss your findings and then **generate a hypothesis** about the relatedness of the species within your group.



Instructions: Write the answers to your questions on the *Student Worksheet*, in your lab notebook, or on a separate sheet of paper, as instructed by your teacher.

How can we use BLAST in genetic research?

- **To identify a sample.** For example, many fish, birds, and marine mammals were killed in the 2010 Gulf of Mexico oil spill. Identifying all of these animals when they are coated in oil or are in their juvenile stage can be difficult. Or perhaps you believe that the packaged fish sold by a particular market is not the wild caught Alaskan Coho salmon they are advertising, but actually farmed Atlantic salmon.
- **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). It is your job to determine whether the infection really is influenza, and where the patient may have gotten infected.

PART I: Performing a BLAST Search

Your instructor will provide information on how to obtain an “unknown DNA sequence.” Copy this sequence using the **Copy** command from the **Edit** menu of your web browser. Alternatively, you can save it as a text file (.txt) to your computer or memory stick.

1. Once you have your unknown DNA sequence, go to the BLAST homepage: <http://blast.ncbi.nlm.nih.gov/Blast.cgi>.
2. Choose a BLAST program to run, as shown in **Figure 1**.

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 group.

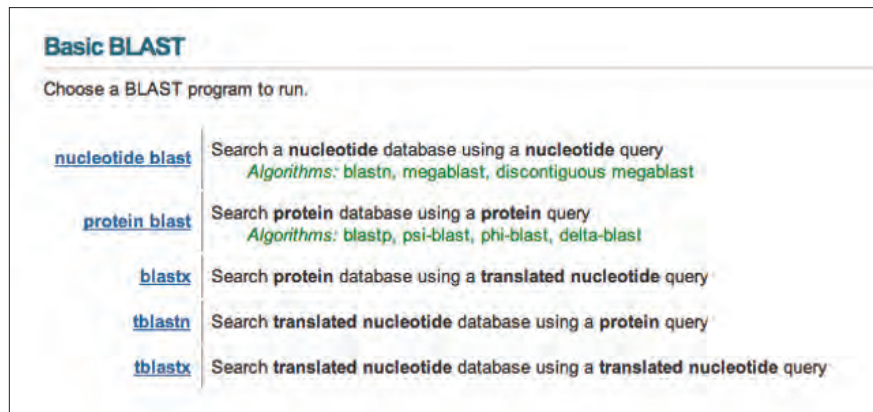


Figure 1: Choosing a BLAST Program. Source: NCBI BLAST.



You will be using your DNA sequence to search for other DNA sequences. Which BLAST program will you run?

3. Paste your DNA sequence into the search box, **Enter Query Sequence** (green box, **Figure 2**), or upload a file containing the sequence (.txt files).
4. Select the database to search, as shown in **Figure 2**. To identify an unknown sequence, or to find related sequences, select the **Nucleotide Collection (nr/nt)**.
5. Make your **Program Selection**, based on the goal of your search. In this exercise, you are identifying a sample.
 - **Highly similar sequences (megablast).** Example use: identifying a sample.
 - **More dissimilar sequences (discontinuous megablast).** Example use: find closely related species.
 - **Somewhat similar sequences (blastn);** best for shorter searches (fewer than 20 bases).

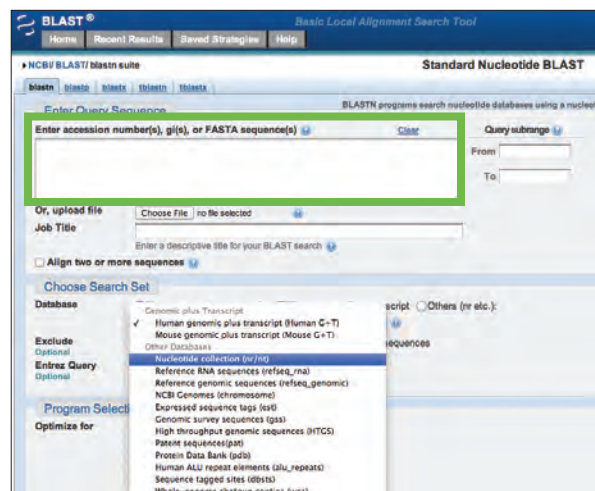


Figure 2: Selecting the Nucleotide Collection. Source: NCBI BLAST.

6. Click the **BLAST** button to start your search, as shown in **Figure 3**.

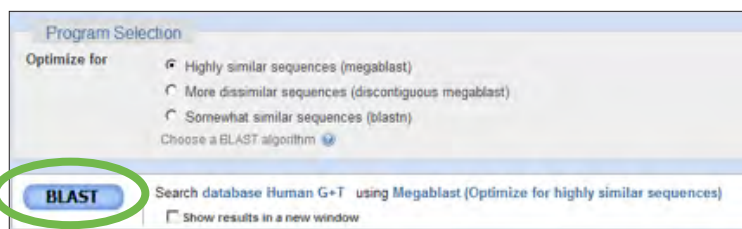


Figure 3: Starting the BLAST Search. Source: NCBI BLAST.

PART II: Understanding BLAST Search Results and Pairwise Comparisons

7. Each BLAST result is a pairwise comparison between your DNA sequence and the DNA sequences in an NCBI database. **Figure 4** provides information on the different parts of the BLAST search result.

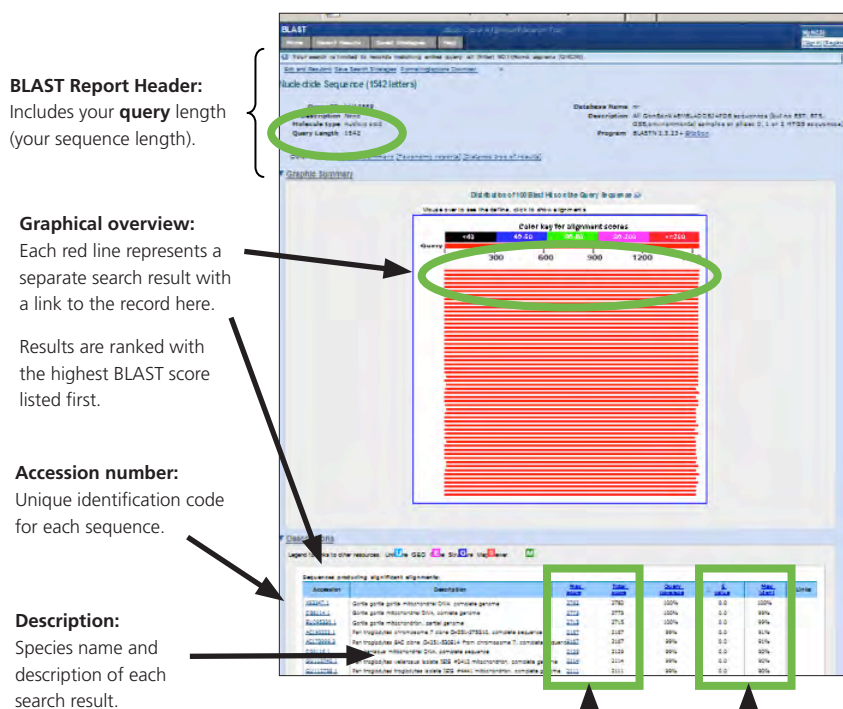


Figure 4: Parts of a BLAST Result. Source: NCBI BLAST.

BLAST Scores and Statistics: How to Rank Your Results

Max Score/Total Score: Score of the pairwise comparison between your DNA sequence and the DNA sequence in the NCBI database (EX: +2 for match; -1 for mismatch, -2 for a gap). Higher scores mean better alignments.

```

A T G G T T A C T
A - G G A A T C T
+2 -2 +2 +2 -1 -1 -1 +2 +2   Total Score = 5

```

E-value: Would this result be "expected" by chance alone? Lower E values show results that are less likely to be obtained by chance. Lower E- value numbers are better.

- Example 1:** A hit with an E-value of 0.0 is less likely to be due to chance than a hit with an E-value of 5.4
- Example 2:** A hit with an E-value of $2e-8$ (2.0×10^{-8} or 0.00000002) is less likely to be due to chance than a hit with an E- value of 7.5

LESSON 2

CLASS SET

Query: When searching databases like those at the NCBI, your "query" is the sequence you are searching with and trying to match. In this case, your query is your unknown sequence.

Subject: When searching the databases at the NCBI, the subject sequences are sequences from the database that match the query. In this case, if a subject sequence is identical to your query, your query sequence probably came from the same type of organism that contributed the subject sequence.

Accession number: A unique identifier or code assigned to every entry in the National Center for Biotechnology Information (NCBI) databases. This unique code can be used to search the databases to find your gene or protein of interest.

8. Select your first search result, either by clicking on the first/top red line, or scrolling down and clicking on the **Max Score** to the right of the **Description**.

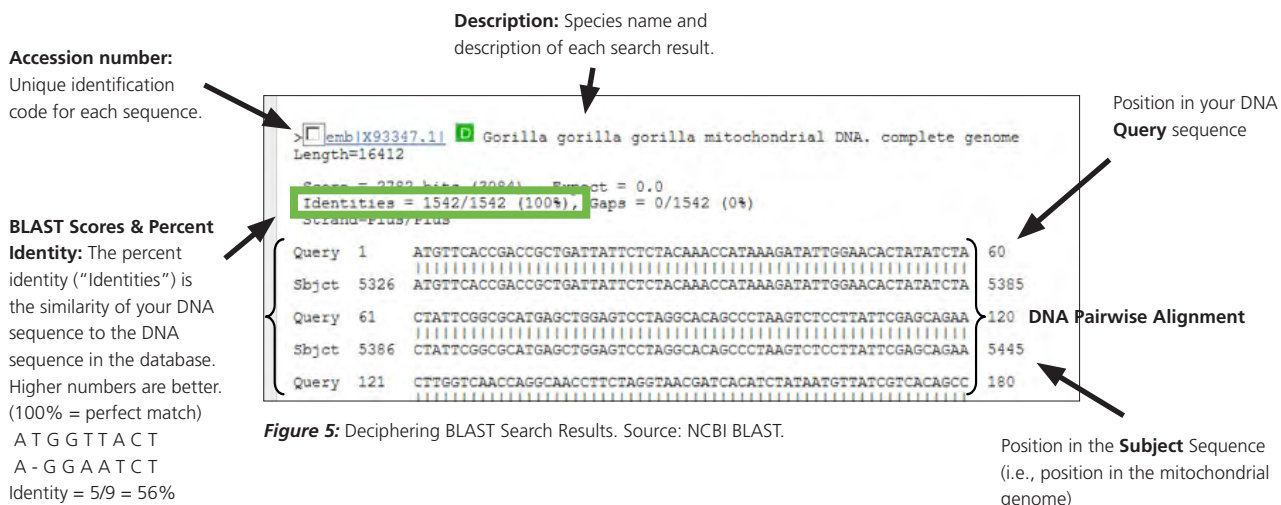
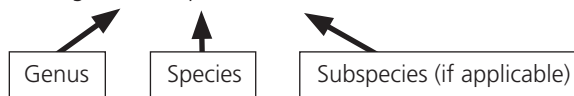


Figure 5: Deciphering BLAST Search Results. Source: NCBI BLAST.

9. What is the description of your first (best) BLAST result? (See **Figure 5**). 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*.



What is scientific name for **your** species? Be sure to include the Genus, species and subspecies (if applicable).



10. What is the accession number of your search result?



11. Do you feel confident that you have correctly identified the species from which your sequence was taken? Justify your answer in terms of the BLAST scores (percent identity and E-value).



12. Your BLAST search result will include a pairwise comparison for each result. Each alignment is a comparison between two sequences: your **query** (unknown) sequence and a **subject** (result) sequence. Look through your pairwise comparison data. Do you see any differences between the query sequence and the best matching subject sequence in the pairwise alignment?



13. Your query will likely begin at nucleotide #1, but the best matching subject sequence might not (see “DNA Pairwise Alignment” in **Figure 5**). The *COI* gene is encoded by the mitochondrial genome, which is over 13,000 nucleotides long. Look for the nucleotide number of your subject sequence (abbreviated “Sbjct”) just below where your query starts (at #1), as shown in **Figure 6**. This is the nucleotide position where your subject sequence starts.

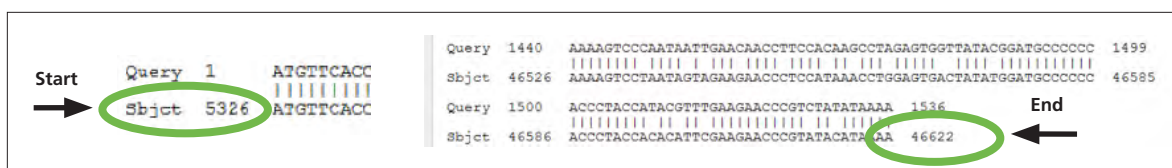


Figure 6: Locating the Subject Sequence. Source: NCBI BLAST.

- At what nucleotide position does your subject sequence start?
- At what nucleotide position does your subject sequence end?

PART III: Finding Taxonomic Information in the Barcode of Life Database (BOLD)

The Barcode of Life Database (BOLD) contains a great deal of information. Like the NCBI, it is a collection of databases that allows scientists around the world to collaborate by having a single location where researchers can submit or retrieve scientific data—specifically, DNA barcode information.

At BOLD, you will:

- Determine** whether your species has been barcoded.
- Learn important **taxonomic** information about your species, and use that information to inform your hypothesis and to choose which other researchers in your class you will **collaborate with**.

Taxonomic: Taxonomy is the science of classifying organisms into groups based on similarities in their physical characteristics. The goal of taxonomy is to organize species in groups based on their evolution and common ancestors. The word comes from the Greek “taxis” which means “order” or “arrangement.”

Understanding Taxonomy: What’s the Point?

Scientists categorize all life on planet Earth into a hierarchy called a **taxonomy**—Kingdom: phylum: class: order: family: genus: species. The taxonomic hierarchy is derived from evolutionary relationships, and **makes it possible for scientists around the world to be sure they are all talking about the same species** by providing a standard format for identification. The last two levels of the taxonomic hierarchy, the *Genus* and *species*, constitute the scientific name for a species, which you found in your BLAST results. All the species you will be studying are animals, which are part of Kingdom Animalia.

Determine Whether Your Species Has Been Barcoded

Barcoding species involves matching the DNA barcode sequence from an identified, known physical specimen (such as in a zoo, aquarium, or museum) with your DNA barcode sequences. You can use BOLD to identify unknown sequences, similar to what you did with your BLAST search. You can also search BOLD using the scientific name of your species, which you learned from your BLAST search. Now that you know the scientific name of your species, it is time to find the DNA barcode sequence for your species at the BOLD.

14. Go to the BOLD database: <http://www.barcodinglife.org>.

15. Type the scientific name of your species in the search box in the lower left portion of the screen, as shown in **Figure 7**. If your search does not return any results, try searching with just the Genus name.

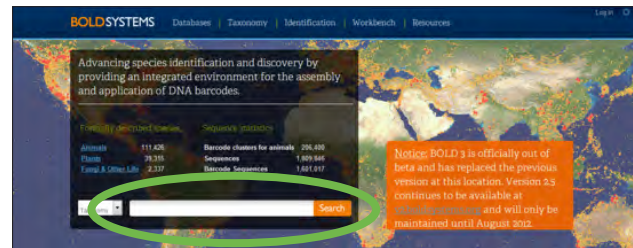


Figure 7: Searching for a Species in BOLD.
Source: Barcode of Life Data Systems.

16. Has your species been barcoded?

17. If your species has been barcoded, you will find a page for your species in BOLD. Take a moment to look at the information available, as shown in **Figure 8**.

- How many specimens are available with Barcodes? [**Note:** "Barcodes" are complete DNA sequences, while "Sequences" may include only fragments.]
- How many of these records are Public Records?

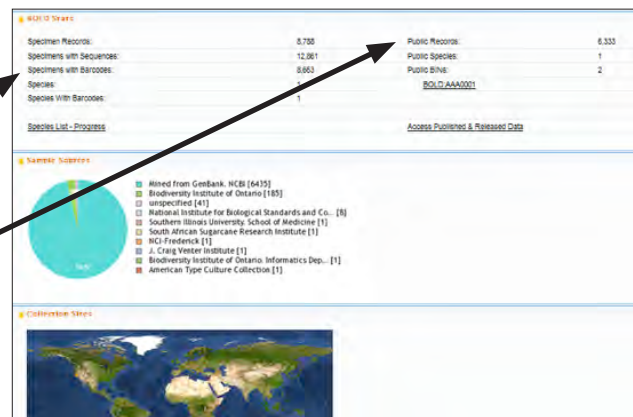
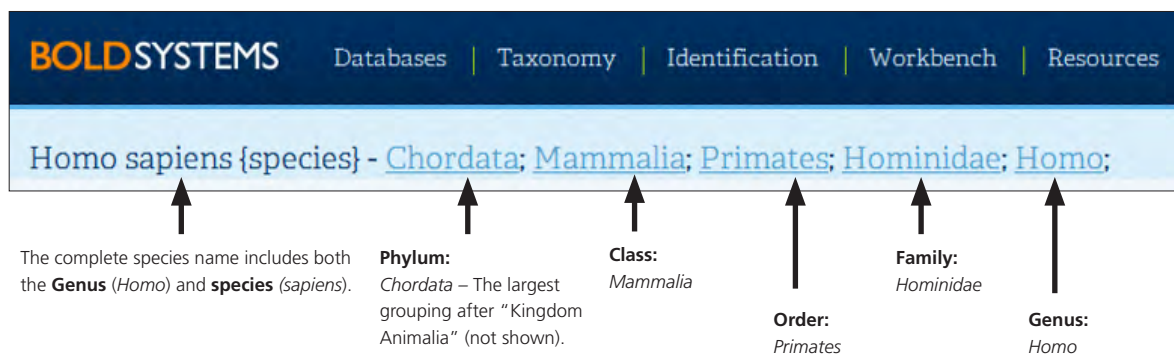


Figure 8: Finding information on a BOLD Species Page.
Source: Barcode of Life Data Systems.

18. Find the **taxonomic information** about the species across the top left side of the web page. Other taxonomic categories are also listed on the BOLD species page. Each is a subcategory of the one before it, from more general (larger) categories to more specific (smaller) categories. In the example shown in **Figure 9**, taxonomic information for the species *Homo sapiens* is displayed.

Figure 9: Decoding the Taxonomic Information. Source: Barcode of Life Data Systems.



19. What is the Family for your species?



20. What is the Order for your species?



21. What is the Class for your species?

PART IV: Finding Your Collaborators and Generating a Hypothesis about Species Relatedness

Scientists often specialize in studying one particular group or type of species, such as canines, fish, or marine mammals. They form collaborations with other scientists who study the same or related species. **Your teacher will help you find the other student researchers in your class who are studying species in the same taxonomic class as your species. Let your teacher know when you are ready to find your collaborators.**

Your teacher will also provide you with a handout that contains pictures of your group's species. Have a discussion with your collaborators about the characteristics of the species within your group. You can use these pictures and your discussion to help you come up with your research hypothesis about how your species are related to one another. Each group member should write down her own hypothesis. It is all right if your hypothesis is the same as or similar to the hypotheses of other members in your group; your hypothesis may also be completely different from or even contradict others in your group.



22. What is your Group Name (includes the class of your species)?



23. List the names of all the collaborators in your group and the species they are studying in the table on your *Worksheet*, or draw a similar table in your lab notebook or on your answer sheet. Be sure to list yourself as well!

Collaborators	Scientist's Name	Species Studied
Collaborator #1		
Collaborator #2		
Collaborator #3		
Collaborator #4		
Collaborator #5		
Collaborator #6		



24. What is your hypothesis about the relatedness of the species within your group?

LESSON 2

HANDOUT

Name _____ Date _____ Period _____

2

Using BLAST and BOLD for Genetic Research Worksheet

Aim: Today, your job as a researcher is to:

1. Use the bioinformatics tool **BLAST** to identify the source of an unknown DNA barcode sequence.
2. Find important **taxonomic information** for your species from **BOLD**.
3. Find your **scientific collaborators** to discuss your findings and then **generate a hypothesis** about the relatedness of the species within your group.



Instructions: Write the answers to your questions on the *Student Worksheet*, in your lab notebook, or on a separate sheet of paper, as instructed by your teacher.

2. You will be using your DNA sequence to search for other DNA sequences. Which BLAST program will you run?
9. What is the scientific name for **your** species?
10. What is the accession number of your search result?
11. Do you feel confident that you have correctly identified the species from which your sequence was taken? Justify your answer in terms of the BLAST scores (percent identity and E-value).
12. Your BLAST search result will include a pairwise comparison for each result. Each alignment is a comparison between two sequences: your query (unknown) sequence and a subject (result) sequence. Look through your pairwise comparison data. Do you see any differences between the query sequence and the best matching subject sequence in the pairwise alignment?
- 13a. At what nucleotide position does your subject sequence start? _____
- 13b. At what nucleotide position does your subject sequence end? _____

LESSON 2

HANDOUT

16. Has your species been barcoded?

17a. How many specimens are available with Barcodes? **[Note: "Barcodes" are complete DNA sequences, while "Sequences" may include only fragments.]**

17b. How many of these records are Public Records?

19. What is the Family for your species?

20. What is the Order for your species?

21. What is the Class for your species?

22. Name your student group and include the Class of your species in your group's name (for example, the Marvelous Mammals Group).

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

Collaborators	Scientist's Name	Species Studied
Collaborator #1		
Collaborator #2		
Collaborator #3		
Collaborator #4		
Collaborator #5		
Collaborator #6		

24. What is your hypothesis about the relatedness of the species within your group?

2

Group 1:
Class Mammalia



Orangutan (*Pongo pygmaeus*)

Source: Eleifert, Wikimedia Commons; http://commons.wikimedia.org/wiki/File:Orang_Utan,_Semenggok_Forest_Reserve,_Sarawak,_Borneo,_Malaysia.JPG



Dusky Leaf Monkey (*Trachypithecus obscurus*)

Source: Robertpollai, Wikimedia Commons; <http://commons.wikimedia.org/wiki/File:DuskyLeavenonkey1.jpg>



Chimpanzee (*Pan troglodytes*)

Source: Thomas Lersch, Wikimedia Commons; http://commons.wikimedia.org/wiki/File:Schimpanse_zoo-leipzig.jpg



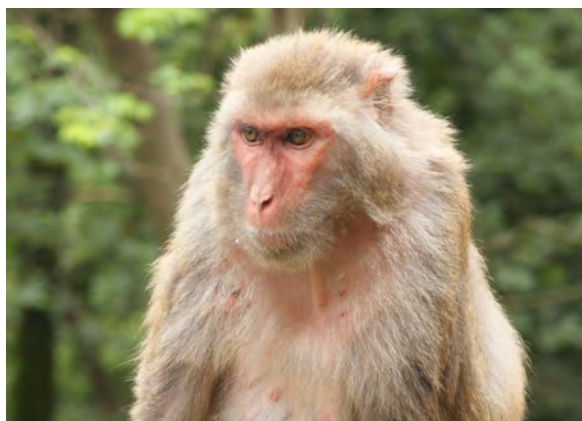
Gorilla (*Gorilla gorilla*)

Source: Ltshears, Wikimedia Commons; http://commons.wikimedia.org/wiki/File:Gorilla_498.jpg



Spider Monkey (*Ateles geoffroyi*)

Source: Patrick Gijsbers, Wikimedia Commons; <http://commons.wikimedia.org/wiki/File:Zwarthandslingeraap-12.jpg>



Rhesus Macaque (*Macaca mulatta*)

Source: Einar Fredriksen, Wikimedia Commons; http://commons.wikimedia.org/wiki/File:Macaca_mulatta_in_Guiyang.jpg

2

Group 2:
Class Aves



Raven (*Corvus corax*)

Source: Public Domain, Wikimedia Commons, [http://commons.wikimedia.org/wiki/File:Corvus_corax_\(NPS\).jpg](http://commons.wikimedia.org/wiki/File:Corvus_corax_(NPS).jpg)



Little Penguin (*Eudyptula minor*)

Source: Aaron Jacobs from Los Angeles, CA, Wikimedia Commons, http://commons.wikimedia.org/wiki/File:Little_Penguin-Sydney.jpg



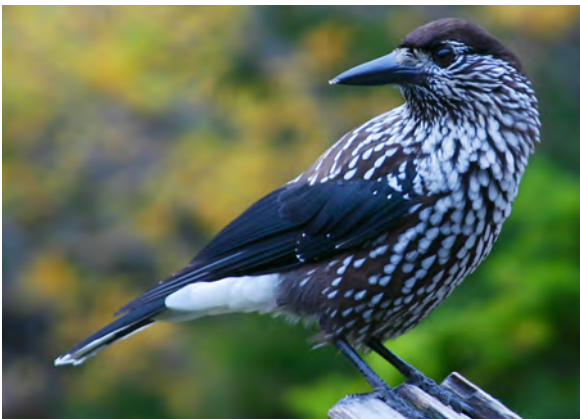
Chicken (*Gallus gallus*)

Source: Dina Kovarik, NWABR



Tricolored Heron (*Egretta tricolor*)

Source: Dick Daniels (<http://carolinabirds.org/>), Wikimedia Commons, [http://commons.wikimedia.org/wiki/File:Tricolored_Heron_\(Egretta_tricolor\)_RWD2.jpg](http://commons.wikimedia.org/wiki/File:Tricolored_Heron_(Egretta_tricolor)_RWD2.jpg)



Nutcracker (*Nucifraga caryocatactes*)

Source: Released to Public Domain by MurrayBHenson, Wikimedia Commons, http://commons.wikimedia.org/wiki/File:Spotted_Nutcracker.jpg



Chukar Partridge (*Alectoris chukar*)

Source: Mdf, Wikimedia Commons, <http://commons.wikimedia.org/wiki/File:Alectoris-chukar-001.jpg>

2

Group 3: Class Osteichthyes or Actinopterygii
(The Bony Fishes)



Star Snapper (*Lutjanus stellatus*)

Source: Randall, John E. From Fishbase



Japanese Jack Mackerel (*Trachurus japonicus*)

Source: Izuzuki, Wikimedia Commons, <http://commons.wikimedia.org/wiki/File:MaAji.jpg>



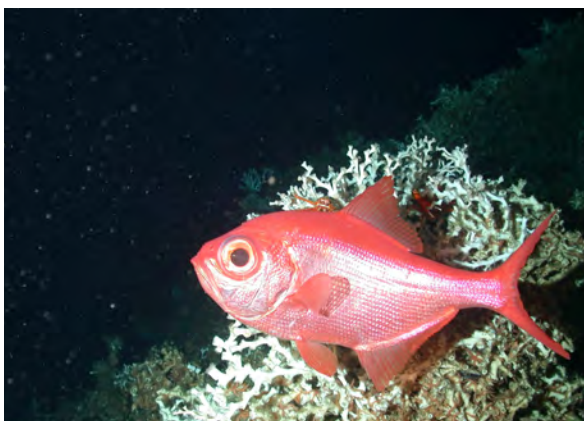
Japanese Soldierfish (*Ostichthys japonicus*)

Source: Randall, John E. From Fishbase



Coho Salmon (*Oncorhynchus kisutch*)

Source: Thomas Quinn, School of Aquatic and Fishery Sciences, University of Washington



Alfonsino (*Beryx decadactylus*)

Source: Life on the Edge 2004 Expedition: NOAA Office of Ocean Exploration, Wikimedia Commons, http://commons.wikimedia.org/wiki/File:Beryx_decadactylus_4.jpg



Chilipepper or Rockfish (*Sebastes goodei*)

Source: Gotshall, Daniel W. From FishBase

2

**Group 4: Class Chondrichthyes
or Elasmobranchii (The Cartilaginous Fishes)**



Roughnose Stingray (*Pastinachus solocirostris*)

Source: CSIRO Marine and Atmospheric Research, Tasmania, Australia.



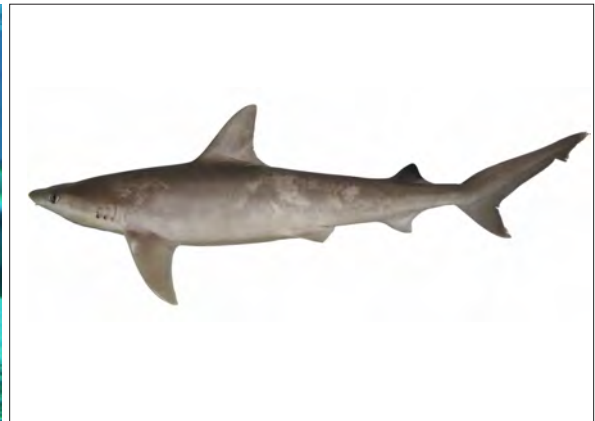
Mackerel Shark (*Lamna nasus*)

Source: Wikimedia Commons, http://en.wikipedia.org/wiki/File:Lamna_nasus_noaa.jpg



Dusky Shark (*Carcharhinus obscurus*)

Source: shishihenge, Wikimedia Commons, http://commons.wikimedia.org/wiki/File:Dusky_shark_seaworld.jpg



Whitecheek Shark (*Carcharhinus dussumieri*)

Image from paper, "A redescription of *Carcharhinus dussumieri* and *C. sealei*, with resurrection of *C. coatesi* and *C. tjutjot* as valid species (Chondrichthyes: Carcharhinidae)". Figure 5C. Source: W. White, CSIRO.



Great White Shark (*Carcharhinus carcharias*)

Source: Hermanus Backpackers, Wikimedia Commons, http://commons.wikimedia.org/wiki/File:Great_white_shark_south_africa.jpg



Bennett's Stingray (*Dasyatis bennetti*)

Source: BIO Photo Group, Biodiversity Institute of Ontario.

2

Group 5:
Class Reptilia



Shield-tailed Agama (*Xenagama taylori*)

Source: H. Zell, Wikimedia Commons; http://commons.wikimedia.org/wiki/File:Xenagama_taylori_01.JPG



Iguana (*Iguana iguana*)

Source: Cary Bass, Wikimedia Commons; http://commons.wikimedia.org/wiki/File:Cuban_Iguana_full.jpg



American Alligator (*Alligator mississippiensis*)

Source: Postdlf, Wikimedia Commons; http://commons.wikimedia.org/wiki/File:American_Alligator.jpg



Komodo Dragon (*Varanus komodoensis*)

Source: Dezidor, Wikimedia Commons; http://commons.wikimedia.org/wiki/File:Varan_komodsk%C3%BD.jpeg



African Chameleon (*Chamaeleo africanus*)

Source: Benny Trapp, Wikimedia Commons; http://commons.wikimedia.org/wiki/File:Benny_Trapp_Chamaeleo_africanus_kletterndes_Weibchen.jpg



Nile Crocodile (*Crocodylus niloticus*)

Source: Wikimedia Commons; <http://commons.wikimedia.org/wiki/File:NileCrocodile-SRG001c.jpg>

LESSON 2

CLASS SET

2

Using BLAST and BOLD for Genetic Research

Teacher Answer Key

[Note: The suggested total point value for this worksheet is **30 points**, or approximately 2 points per question.]

PART I: Performing a BLAST Search

2. Which BLAST program will you run?

Nucleotide blast.

PART II: Understanding BLAST Search Results and the Pairwise Comparison

9. Description	10. Accession	11. Percent Identity	11. E-value	13. Subject sequence Start	13. Subject sequence End
Group 1: Class Mammalia					
1. <i>Pan troglodytes</i>	gb GU112742.1	100%	0.0	5326	6867
2. <i>Gorilla gorilla</i>	X93347.1	100%	0.0	5326	6867
3. <i>Pongo pygmaeus</i>	D38115.1	100%	0.0	5332	6867
4. <i>Trachypithecus obscurus</i>	AY863425.1	100%	0.0	5348	6854
5. <i>Macaca mulatta</i>	AY612638.1	100%	0.0	5850	7389
6. <i>Ateles geoffroyi</i>	AB016730.1	100%	0.0	1	1540
Group 2: Class Aves					
7. <i>Corvus corax</i>	AY527236.1	99%	0.0	1	658
8. <i>Nucifraga caryocatactes</i>	GQ482243.1	100%	0.0	1	694
9. <i>Eudypitula minor</i>	EU525395.1	100%	0.0	1	841
10. <i>Egretta tricolor</i>	DQ432903.1	100%	0.0	652	1
11. <i>Gallus gallus</i>	AP003323.1	100%	0.0	6663	7406
12. <i>Alectoris chukar</i>	FJ752426.1	100%	0.0	6566	7309
Group 3: Class Osteichthyes					
13. <i>Oncorhynchus kisutch</i>	EF126369.1	100%	0.0	6480	7362
14. <i>Beryx splendens</i>	AP004430.1	100%	0.0	5486	6368
15. <i>Sebastes goodei</i>	GQ396749.1	100%	0.0	1	1098
16. <i>Ostichthys japonicus</i>	AP004431.1	100%	0.0	5488	6368
17. <i>Lutjanus stellatus</i>	EU502687.1	100%	0.0	1	1551
18. <i>Trachurus japonicas</i>	AP003092.1	100%	0.0	5521	6401

9. Description	10. Accession	11. Percent Identity	11. E-value	13. Subject sequence Start	13. Subject sequence End
Group 4: Class Chondrichthyes (or Class Elasmobranchii)					
19. <i>Carcharhinus dussumieri</i>	DQ108305.1	100%	0.0	1	655
20. <i>Carcharhinus obscurus</i>	FJ518938.1	100%	0.0	1	652
21. <i>Dasyatis bennetti</i>	EU595098.1	100%	0.0	2	652
22. <i>Carcharodon carcharias</i>	FJ518943.1	100%	0.0	2	628
23. <i>Lamna nasus</i>	FJ519722.1	100%	0.0	1	652
24. <i>Pastinachus solocistrostris</i>	EU398977.1	100%	0.0	1	655
Group 5: Class Reptilia					
25. <i>Alligator mississippiensis</i>	Y13113.1	99%	0.0	5323	6205
26. <i>Iguana iguana</i>	AJ278511.2	100%	0.0	5336	6218
27. <i>Varanus komodoensis</i>	AB080275.1	100%	0.0	5426	6308
28. <i>Crocodylus niloticus</i>	DQ273697.1	100%	0.0	5331	6213
29. <i>Chamaeleo africanus</i>	EF222196.1	100%	0.0	5243	6125
30. <i>Xenagama taylori</i>	DQ008215.1	100%	0.0	5173	6052

12. Your BLAST search will result will include a **Pairwise Comparison** for each result. This is the alignment between your **query** (unknown) sequence and the **subject** (result) sequence. Look through your Pairwise Comparison. Do you see any differences between the query sequence and the best matching subject sequence in the pairwise alignment?

There should not be any differences between the Query and Subject. This is expected given the high Maximum identity score and low E-value above (Question #4).

PART III: Finding Taxonomic Information in the Barcode of Life Database (BOLD)

16. Has your species been barcoded?

All of the species pre-selected by the Bio-ITEST program have been barcoded.

17a. How many specimens are available with Barcodes? **[Note: "Barcodes" are complete DNA sequences, while "Sequences" may include only fragments.]**

Answers vary by species and are constantly updated in the BOLD. All specimens have at least one barcode sequence.

17b. How many of these records are Public Records?

Answers vary by species and are constantly updated in the BOLD. All Specimens have at least one Barcode Sequence.

19. What is the Family for your species?

20. What is the Order for your species?

21. What is the Class for your species?

Species Name	Family	Order	Class
Group 1: Class Mammalia			
<i>Pan troglodytes</i>	Hominidae	Primates	Mammalia
<i>Gorilla gorilla</i>	Hominidae	Primates	Mammalia
<i>Pongo pygmaeus</i>	Hominidae	Primates	Mammalia
<i>Trachypithecus obscurus</i>	Cercopithecidae	Primates	Mammalia
<i>Macaca mulatta</i>	Cercopithecidae	Primates	Mammalia
<i>Ateles geoffroyi</i>	Atelidae	Primates	Mammalia
Group 2: Class Aves			
<i>Corvus corax</i>	Corvidae	Passeriformes	Aves
<i>Nucifraga caryocatactes</i>	Corvidae	Passeriformes	Aves
<i>Eudiptula minor</i>	Spheniscidae	Sphenisciformes	Aves
<i>Egretta tricolor</i>	Ardeidae	Ciconiiformes	Aves
<i>Gallus gallus</i>	Phasianidae	Galliformes	Aves
<i>Alectoris chukar</i>	Phasianidae	Galliformes	Aves
Group 3: Class Osteichthyes			
<i>Oncorhynchus kisutch</i>	Salmonidae	Salmoniformes	Osteichthyes (or Actinopterygii)
<i>Beryx decadactylus</i>	Berycidae	Beryciformes	Osteichthyes (or Actinopterygii)
<i>Sebastes goodei</i>	Sebastidae	Scorpaeniformes	Osteichthyes (or Actinopterygii)
<i>Ostichthys japonicus</i>	Holocentridae	Beryciformes	Osteichthyes (or Actinopterygii)
<i>Lutjanus stellatus</i>	Lutjanidae	Perciformes	Osteichthyes (or Actinopterygii)
<i>Trachurus japonicus</i>	Carangidae	Perciformes	Osteichthyes (or Actinopterygii)
Group 4: Class Chondrichthyes (or Elasmobranchii)			
<i>Carcharhinus dussumieri</i>	Carcharhinidae	Carcharhiniformes	Chondrichthyes (or Elasmobranchii)
<i>Carcharhinus obscurus</i>	Carcharhinidae	Carcharhiniformes	Chondrichthyes (or Elasmobranchii)
<i>Dasyatis bennetti</i>	Dasyatidae	Rajiformes	Chondrichthyes (or Elasmobranchii)
<i>Carcharodon carcharias</i>	Lamnidae	Lamniformes	Chondrichthyes (or Elasmobranchii)
<i>Lamna nasus</i>	Lamnidae	Lamniformes	Chondrichthyes (or Elasmobranchii)
<i>Pastinachus solocirostris</i>	Dasyatidae	Rajiformes	Chondrichthyes (or Elasmobranchii)
Group 5: Class Reptilia			
<i>Alligator mississippiensis</i>	Crocodylidae	Crocodylia/Crocodilia	Reptilia
<i>Iguana iguana</i>	Iguanidae	Squamata	Reptilia
<i>Varanus komodoensis</i>	Varanidae	Squamata	Reptilia
<i>Crocodylus niloticus</i>	Crocodylidae	Crocodylia/ Crocodilia	Reptilia
<i>Chamaeleo africanus</i>	Chamaeleonidae	Squamata	Reptilia
<i>Xenagama taylori</i>	Agamidae	Squamata	Reptilia

PART IV: Finding Your Collaborators and Generating a Hypothesis About Species Relatedness

Scientists often specialize in studying one particular group or type of species, such as canines, fish, or marine mammals. They form collaborations with other scientists who study the same or related species. **Your teacher will help you find the other student researchers in your class who are studying species in the same taxonomic class as your species. Let your teacher know when you are ready to find your collaborators.**

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.

22. Name your student group and include the Class of your species in your group's name (for example, the Marvelous Mammals Group).

Class Mammalia, Class Aves, Class Osteichthyes, Class Chondrichthyes (or Class Elasmobranchii), or Class Reptilia

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

Collaborators	Scientist's Name	Species Studied
Collaborator #1	Student Name #1	See list below
Collaborator #2	Student Name #2	See list below
Collaborator #3	Student Name #3	See list below
Collaborator #4	Student Name #4	See list below
Collaborator #5	Student Name #5	See list below
Collaborator #6	Student Name #6	See list below

Group 1: Class Mammalia (Primates)

1. <i>Pan troglodytes</i>	Chimpanzee
2. <i>Gorilla gorilla</i>	Gorilla
3. <i>Pongo pygmaeus</i>	Orangutan
4. <i>Trachypithecus obscurus</i>	Dusky Leaf Monkey
5. <i>Macaca mulatta</i>	Rhesus Macaque
6. <i>Ateles geoffroyi</i>	Spider Monkey

Group 2: Class Aves (Birds)

7. <i>Corvus corax</i>	Raven
8. <i>Nucifraga caryocatactes</i>	Nutcracker
9. <i>Eudyptula minor</i>	Little Penguin
10. <i>Egretta tricolor</i>	Tricolor Heron
11. <i>Gallus gallus</i>	Chicken
12. <i>Alectoris chukar</i>	Chukar Partridge

Group 3: Class Osteichthyes or Actinopterygii (Bony Fishes)

- | | |
|---------------------------------|-------------------------|
| 13. <i>Oncorhynchus kisutch</i> | Coho Salmon |
| 14. <i>Beryx decadactylus</i> | Alfonsino |
| 15. <i>Sebastes goodei</i> | Chilipepper or Rockfish |
| 16. <i>Ostichthys japonicus</i> | Japanese Soldierfish |
| 17. <i>Lutjanus stellatus</i> | Star Snapper |
| 18. <i>Trachurus japonicus</i> | Japanese Jack Mackerel |

Group 4: Class Chondrichthyes or Elasmobranchii (Cartilaginous Fishes)

- | | |
|--------------------------------------|--------------------|
| 19. <i>Carcharhinus dussumieri</i> | Whitecheek Shark |
| 20. <i>Carcharhinus obscurus</i> | Dusky Shark |
| 21. <i>Dasyatis bennetti</i> | Bennett's Stingray |
| 22. <i>Carcharodon carcharias</i> | Great White Shark |
| 23. <i>Lamna nasus</i> | Mackerel Shark |
| 24. <i>Pastinachus solocirostris</i> | Roughnose Stingray |

Group 5: Class Reptilia (Reptiles)

- | | |
|---------------------------------------|---------------------|
| 25. <i>Alligator mississippiensis</i> | American Alligator |
| 26. <i>Iguana iguana</i> | Iguana |
| 27. <i>Varanus komodoensis</i> | Komodo Dragon |
| 28. <i>Crocodylus niloticus</i> | Nile Crocodile |
| 29. <i>Chamaeleo africanus</i> | African Chameleon |
| 30. <i>Xenagama taylori</i> | Shield-tailed Agama |

24. What is your 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.

