DNA BARCODING

DNA Barcoding is a global program to establish a standardized DNA sequence reference library for all species. To give you some idea of the challenge this entails- there are approximately 2 million described species, though recent research indicates there may be as many as 8-9 million species on the planet (Table 1 modified from Mora et al. 2014).

Table 1.
Number of described, predicted and barcoded species

	Described	Predicted	Barcoded
Animal	953,000	7,770,000	144,000
Plant	216,000	298,000	54,000
Other	75,000	682,000	17,000
Total	**1,244,000**	**8,750,000**	**215,000**

DNA barcoding represents a powerful tool for understanding and conserving global biodiversity. DNA barcoding can be used alongside standard visual identification (ie. comparison of morphological characteristics) to improve species identification. More importantly, barcoding can (1) help scientists identify species that look alike (ie. cryptic species), (2) allow identification where traditional methods are not applicable (eg. *see Sushi Gate inset*), and (3) provide a reliable, relatively simple method for identifying species that does not require years of specialized training.

Sushi-gate: Using DNA to identify fish fraud

Students using DNA barcoding found 23% of fish sold in New York City sushi restaurants was mislabeled. http://www.nytimes.com/2008/08/22/world/americas/22iht-fish.1.15539112.html

A series of follow up studies found as much as 55% of fish sold nationwide is mislabeled.
http://www.huffingtonpost.com/2012/12/11/mislabeled-fish-new-york_n_2272570.html

New York: Fraud by Fish Type

	Mislabeled	Labeled Correctly
tuna	17	1
snapper	15	4
salmon	11	45
cod	4	12
sole	3	9
grouper	2	1
striped bass	2	2
halibut	1	3
blackfish	1	

HOW IT'S DONE

DNA barcoding is based on a relatively simple concept- short sections of DNA can be used to identify and discover species in the same way a supermarket scanner uses the UPC barcode to identify your purchase. While no single gene or gene region has been found to clearly identify all species, scientists have found that a roughly 700 **nucleotide** region of the mitochondrial **cytochrome oxidase subunit 1 (CO1)** gene can successfully identify most animals. In plants, the CO1 sequence evolves slowly (ie. fewer mutations are added each generation), making it less reliable when identifying unknown plant species. Instead, for plants scientists have identified two gene regions coding for chloroplast proteins (matK and rbcL) that can be successfully used to most identify species.

Gulf Islands Research & Education Center | NATIONAL PARK SERVICE | University of **West Florida**

To identify species using DNA barcodes requires (1) the specimen, (2) genetic analyses, and (3) a DNA reference library.

(1) DNA Sample

Tissue samples for DNA barcoding can be attained from living, preserved, and processed (eg. ground beef) specimens. Commonly, specimens are collected in the field or come from existing collections housed at natural history museums, zoos, and botanical gardens. Because as little as 5 mg of tissue is needed for DNA barcoding, in many cases live specimens can be sampled and returned un-harmed to the environment. Tissue from preserved and processed specimens can as well be used for DNA barcoding (*see Sushi Gate*) because DNA is a fairly robust molecule and multiple copies of the most common barcoding genes can be found in each cell. Once the tissue sample is collected it is either frozen or placed in a preservative (usually in 90% ethanol) for future genetic analyses.

(2) Genetic Analyses

Figure 1. Genetic analyses for DNA barcoding require isolation of the DNA, amplification of the target gene region (PCR), and sequencing of the amplified gene region.
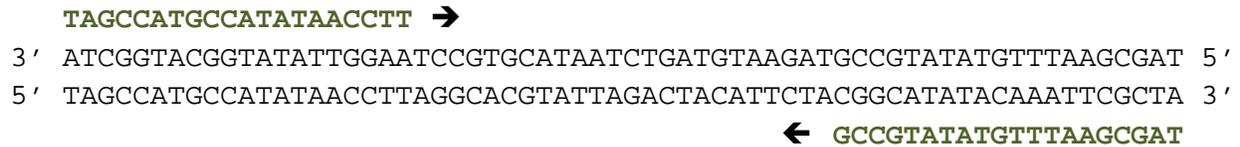

DNA Isolation   Polymerase Chain Reaction (PCR)   DNA Sequencing
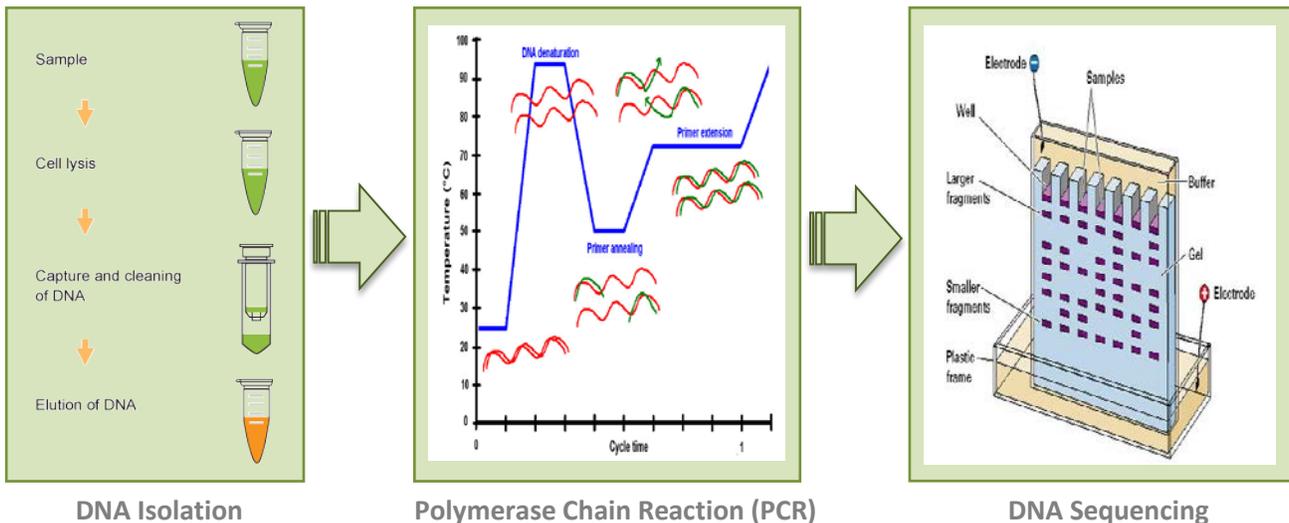
DNA Isolation: There are a number of kits that allow for rapid isolation of DNA from different tissue types (eg. muscle, blood, or hair). In all cases, a combination of physical and chemical processes is used to break open (lyse) the cells and to denature cellular proteins. Typically, researchers then use one of a variety of methods to isolate the DNA and rinse away the remaining cellular contents (eg. by binding the DNA to small glass beads or by centrifuging the sample to separate the DNA from denatured proteins

and other cellular molecules). The final concentrated solution of total cellular DNA can then be used as the template for PCR.

Polymerase Chain Reaction (PCR): To copy a target gene region (eg. CO1) researchers have figured out how to take advantage of the natural ability of **DNA polymerase** to synthesize a new DNA strand from a template strand (in this case the template strand comes from the DNA isolated from the tissue sample). Because DNA polymerase can only build a new strand of DNA onto an existing strand, researchers can specify the exact section of DNA that will be copied by using synthesized **DNA primers** (typically 20 nucleotides long) that bind to the beginning and end of the DNA region to be copied. The DNA primers are added to a buffered solution containing the template DNA, excess nucleotides (A, C, T, G), DNA polymerase, and polymerase co-factors (eg. MgCl), and then subjected to a series of alternating hot and cold cycles. This process produces million to billions of copies of the target gene region. (See Appendix 1: PCR)

Figure 2. Example double stranded DNA sequence with forward and reverse DNA primers shown above and below the template DNA. Arrows indicate the direction DNA polymerase will proceed when synthesizing a new strand from the template strand.

```
    TAGCCATGCCATATAACCTT  →
3'  ATCGGTACGGTATATTGGAATCCGTGCATAATCTGATGTAAGATGCCGTATATGTTTAAGCGAT  5'
5'  TAGCCATGCCATATAACCTTAGGCACGTATTAGACTACATTCTACGGCATATACAAATTCGCTA  3'
                                          ←  GCCGTATATGTTTAAGCGAT
```

DNA sequencing: DNA sequencing is very similar to PCR except that a small percentage of the nucleotides have been modified to terminate polymerase activity (**terminator nucleotides**). These nucleotides have also been labeled with a fluorescent dye so they can be detected by a camera. In the sequencing reaction, as DNA polymerase copies the template strand, by chance it will occasionally insert a terminator nucleotide instead of a normal nucleotide. This terminates the reaction and the polymerase moves on to start building a new strand. The result is that the last nucleotide in each of the copied DNA strands is a fluorescently labeled terminator nucleotide. After all cycles of the sequencing reaction are complete it is then run through a high percentage **polyacrylamide gel (PAGE)** and a camera records the identity of each fluorescent nucleotide as it moves through the gel. A software program then translates the observed fluorescent peaks into the appropriate DNA sequence (eg. Figure 2).

Figure 3. Observed fluorescent peaks and associated DNA sequences from three species of deep water sharks (genus: *Squallus*). The peaks in the graph reveal the presence of individual terminator nucleotides (A, C, T, or G). Each terminator nucleotide is represented by a different color: green for adenine, blue for cytosine, black for guanine, and red for thymine. DNA mutations occurring between species are highlighted.
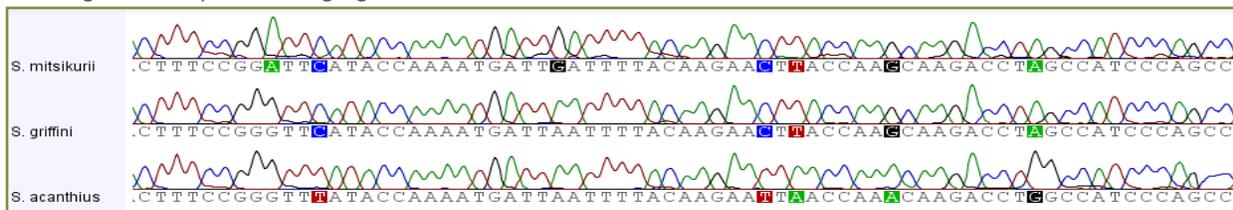
(3) DNA Reference Library

The most common DNA sequence databases are the **Barcode of Life Database (BOLD)** and the **GENBANK** database managed by the National Center for Biotechnology Information (NCBI). BOLD was specifically created to support genetic based species identification and is therefore the go to database for DNA barcoding. The BOLD database currently contains sequence data from more than 100,000 vouchered animal species (Figure 2); however, this is just a fraction of all the known and estimated species on the planet. Nonetheless, for some of the best studied groups (eg. fishes) BOLD contains sequence data from a majority of the known species.

Figure 3. Location and number of BOLD barcoded specimens as of 10/01/2014.
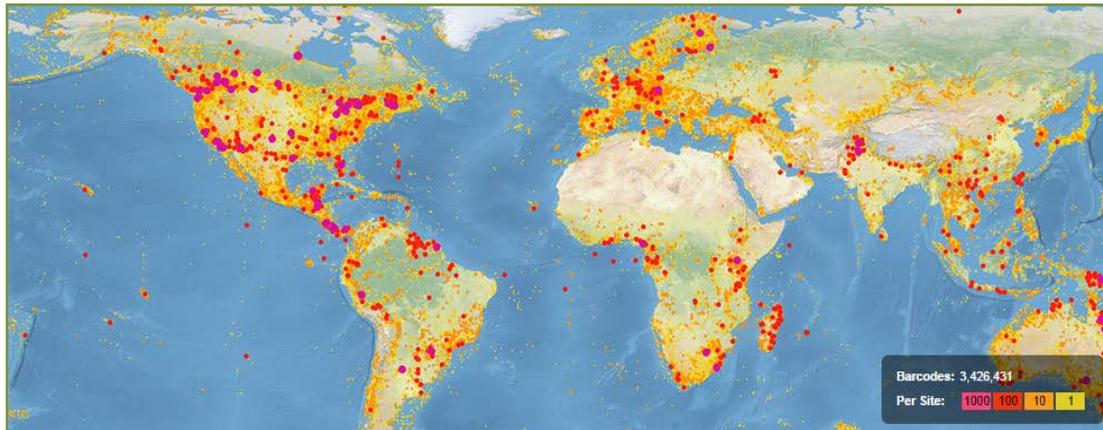
REFERENCES

C. Mora, D. Tittensor, S. Adl, A. Simpson, B. Worm (2011) How many species are there on earth and in the ocean? PLOS Biology, 9 (8): e1001127.

Gulf Islands Research & Education Center | University of West Florida

Lionfish Dissection | Teacher's Guide

Modified from Green, S., Akins, J. L., Morris, J. A. 2012. Lionfish Dissection: Techniques and Applications. NOAA Technical Memoranda. 24pp

PRECAUTIONS

- *Soak lionfish in an ice-seawater slurry or freeze overnight to denature venomous spines.*
  - Handle fish carefully to avoid puncture wounds from the needle sharp dorsal and anal spines (wear puncture resistant gloves if available).

- *First aid for envenomation*
  - Immediately immerse area in near-scalding hot water until pain subsides. Secondary medical attention may be needed if symptoms persist or worsen.

BEFORE CLASS

1) If frozen, be sure to thaw lionfish in the refrigerator 1-2 days before starting the dissection

2) Gather equipment and supplies (below)
   i) If possible, to save class time and improve data collection pre-label lionfish and prey sample collection tubes (eg. table 1)

SUPPLIES AND EQUIPMENT

Class

- ○ Specimen scale (with minimum 0.1g resolution and 1500g capacity)
- ○ Specimen weighing tray or paper towels
- ○ Lionfish data sheet (see attached)
- ○ Alcohol resistant permanent marker

Dissecting Teams

- ○ Thawed lionfish
- ○ Plastic weighing boats (or similar to hold sorted prey items)
- ○ 2.0ml screw top tubes with >90% ethanol
- ○ Tube Racks
- ○ Dissecting tray
- ○ Forceps
- ○ Disposable razors
- ○ Paper towels
- ○ Small beaker with 10% bleach
- ○ Specimen sample labels ('rite in the rain' paper cut into ~2" x ½" pieces)

Gulf Islands Research & Education Center | University of **West Florida**

Table 1. Example Sample Identification numbers indicating the research group (e.g. Pensacola High School) and linking prey items to the lionfish from which they were dissected.

Specimen	Eg. Label
Dissected lionfish #1:	PHS-LF01
Prey item #1:	PHS-LF01-A
Prey item #2:	PHS-LF01-B
Dissected lionfish #2	PHS-LF02
Prey item #1	PHS-LF02-A

DURING CLASS

1) Assign each fish a unique specimen ID if one has not yet been provided

2) Have students record lionfish weight and total length on the **Lionfish Datasheet**

3) To ensure accuracy of data, have students…

   i) Repeat measurements if necessary to ensure accuracy
   ii) Have a second person record the data and verbally confirm the information
   iii) Record all units of measurement (eg. mm, cm, g)

4) To reduce cross contamination between lionfish and prey, or between prey items…

   i) Remind students to use **clean** forceps and new razors when working with prey items

AFTER CLASS

1) Collect team **Lionfish Prey Datasheets**

2) Ensure each sample is labeled and contains sufficient preservative (~5x sample volume)

3) Discard fish and any dissected spines in a robust, puncture-proof, lidded container to avoid later injuries

Gulf Islands Research & Education Center | University of West Florida

LIONFISH DATASHEET

Lionfish ID	Collection Site	Date Collected	TL (mm)	Weight (g)	# of Prey Items	Student Team
Eg. PHS-LF01	USS Massachusetts	10/15/16	245	410	6	Billy, Joe, Bob

Gulf Islands Research & Education Center | University of West Florida

DNA Extractions	Teacher's Guide

SUPPLIES AND EQUIPMENT

Equipment

- ○ Incubator (95 deg. C)
- ○ Centrifuge (6,000 x g)
- ○ Pipettes and tips

For each dissecting team

- ○ Prey tissue sample
- ○ 500ul of 5% Chelex-100 in 1.5ml tube for each prey item
- ○ Microcentrifuge tube (1.5ml)
- ○ Plastic pestle
- ○ Fine-tipped Sharpie
- ○ Sterile scissors and tweezers (if sample was not prepared prior to class)
- ○ 5% bleach in individual cups (if sample was not prepared prior to class)

BEFORE CLASS

- If 5% chelex tubes were not provided, for each prey item to be extracted, add 500ul of 5% chelex to a 1.5ml tube.
- Bring the incubator or water bath to 95 deg. C
- *Optional*: If time is a concern you may want to prepare the tissue samples prior to class. If this is the case, for each prey item to be analyzed, dissect a very small ($\leq$ 2mm) tissue sample and place it in new, labeled 1.m ml tube.
    - o To avoid cross contamination between prey items, be sure to use new scissors and tweezers for each sample OR briefly (but vigorously) rinse scissors and tweezers in 5% bleach between samples.
    - o Keep the samples frozen until class.

DURING CLASS (NOTES AND CAUTIONS)

- Maintain sterile methods, including wearing clean gloves during every step, and using clean scissors and tweezers for each sample.
- Use only a small piece (<2 mm) of muscle tissue for the DNA extraction.
  Note: It's easy to use too much tissue so make sure your sample is about the size of this box 🟧
- 'burp' the tubes after about 5 minutes in the incubator to avoid have the lids pop open.
- Make sure students do not transfer any chelex pellets to the new tube as they will cause the PCR to fail (figure 2).

Figure 1. Proper centrifuge loading. If you are not loading an even number of samples use an appropriate balance tube (balance tube must be the same weight as tube in opposite well).
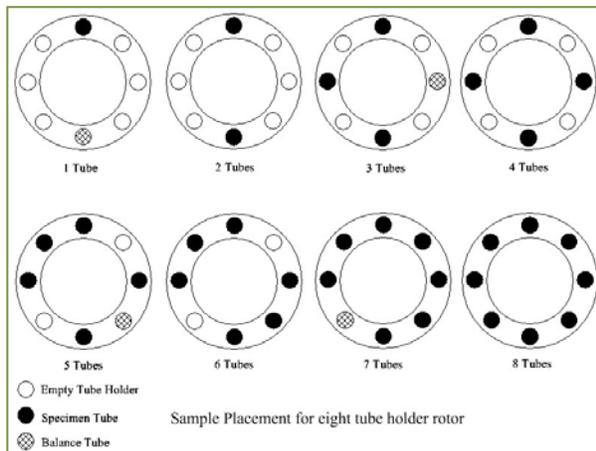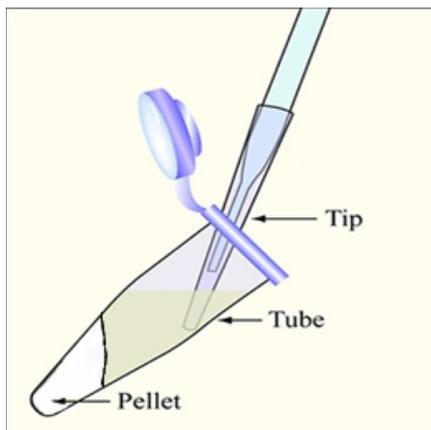

Figure 2. Proper technique for pipetting the DNA solution without disturbing the protein pellet.

| PCR	Teacher's Guide

INTRODUCTION TO PCR

PCR (Polymerase Chain Reaction) uses **DNA polymerase**, a pair of **DNA primers**, and the **four DNA nucleotides** (adenine, thymine, cytosine, and guanine) to create millions to billions of copies of a particular gene or gene segment from the provided template DNA (eg. tissue sample from an unknown fish).

Because DNA polymerase can <u>only</u> add new nucleotides to the end of an existing DNA strand, a pair of DNA primers is needed to allow DNA polymerase to start synthesizing the new DNA strands. The DNA primers also determine the specific region the polymerase will copy (see "How It Works" below). This makes it possible to define a specific region of template DNA to be amplified (eg. CO1 gene region).

After PCR, the copied gene region can be viewed using **agarose gel electrophoresis (AGE)** or "read" using a combination of **DNA sequencing** and **polyacrylamide gel electrophoresis (PAGE)**.
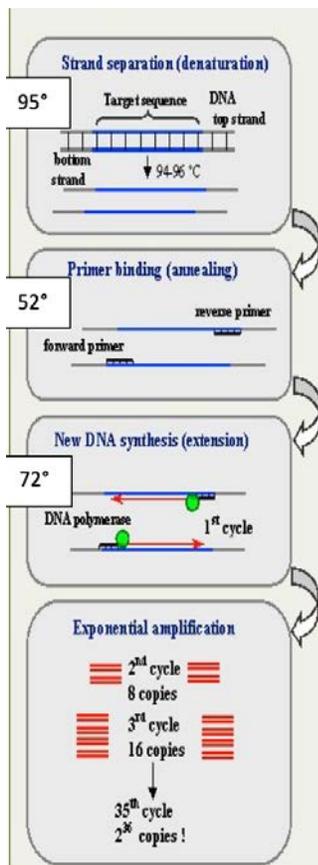
The PCR reaction requires the following ingredients:

DNA Template – DNA that contains the desired target sequence. The beginning of the reaction applies a high temperature to the DNA in order to separate the strands from each other. This is known as **Denaturation.**

DNA Polymerase – This is an enzyme that synthesizes new strands of DNA complementary to the target sequence. The enzyme used in this lab is the Taq DNA Polymerase because it is one of enzymes that is able to generate new strands of DNA and is heat resistant.

Primers – These are short pieces of single stranded DNA that are complementary to the target sequence. This designates the region of the DNA that will be amplified and copied. The polymerase attaches to the primer once the primer has attached itself to the DNA sequence.

Nucleotides (dNTPS) – These are single units of the bases Adenine, Thymine, Guanine, and Cytosine which are the "building blocks" for new DNA strands. The bases always pair up as Adenine and Thymine and Guanine and Cytosine. (A-T & G-C)

Before Class (the day of)

Set up PCR cocktail

1. Calculate the volume of each PCR reagent (Table 1) by multiplying the "per sample" volume times the number of samples to be run. For every eight samples add 1 additional sample to account for reagents lots to pipette tips and tubes.

    a. Eg. For 16 samples, calculate volumes as if you were running 18 samples

2. In a sterile 1.5ml tube, add the appropriate volume of each PCR reagent (Table 1)

3. To mix the reagents, vortex the cocktail for 2 seconds (or gently flick the bottom of the tube) and forcefully tap the tube on the bench to clear any liquid in the cap

4. Add 18.3ul of the cocktail to a labeled 1.2ul PCR tube (for simplicity label tubes 1, 2, 3… and note the sample ID on the attached data sheet).

5. Store the prepared PCR cocktails in the freezer until the students are ready to load their samples.

Table 1: Mytaq 4 primer 'universal' fish PCR cocktail (from Ivanova et al. 2007) and (B) thermal cycler profile

Reagent	per sample (uL)	# of samples	Cocktail (uL)
2x MyTaq DNA polymerase	10		
Primer VF2-t1 (10uM)	0.2		
Primer FR1D-t1 (10uM)	0.2		
Primer FishF2-t1 (10uM)	0.2		
Primer FishR2-t1 (10uM)	0.2		
Nanopure H_2O	7.5		
Total volume per PCR	**18.3ul**		

During Class

Adding DNA to PCR cocktail

1. Have students add **1.7ul of DNA extract** to each prepared PCR cocktail (total volume = 20ul)
2. Place prepared PCRs in the freezer until ready to run the thermalcycler

Loading and running the thermalcycler

1. Follow instructions on thermalcycler to load "4 primer my taq" protocol (Table 2)

Table 2: Mytaq 4 primer 'universal' fish thermal cycler profile

Step	Time
Step 1- 95 deg C	3 min
Step 2- 95 deg C	15 sec
Step 3- 52 deg C	15 sec
Step 4- 72 deg C	10 sec
Step 5- Repeat 2-4 35x	NA
Step 6- 15 deg C	Infinity

Gulf Islands Research & Education Center | University of West Florida

Date: _____   **School:** _____   **Class:** _____

PCR Tube #	Sample #	Team
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		

Gulf Islands Research & Education Center | University of West Florida

Gel Electrophoresis	Teacher's Guide

PRECAUTIONS

- Syber Green is light and temperature sensitive. Keep it wrapped in foil and chilled (in the refrigerator or on ice) before and after use.
- Syber Green is also a potential carcinogen. Always wear gloves when preparing or handling syber green stained gels; however, used gels and tips do not require special disposal.
- If using a UV illuminator, always make sure the UV shield is in place before turning on the illuminator. Serious eye damage can result from short exposure to UV light.

BEFORE CLASS

- Prepare the gel (see below 'preparing the gel' instructions). For best results, prepare the gel just before class; however, the gel can be prepared up to 12 hours before you plan to use it as long is it is stored chilled in TBE buffer with very limited light exposure.
- Print out Table 1 gel loading template.

SUPPLIES AND EQUIPMENT

Class

- ○ Gel rig, combs and power station
- ○ 1x TBE buffer
- ○ Syber Green
- ○ Agarose
- ○ DNA Ladder
- ○ UV or blue light imager
- ○ Camera
- ○ 10 ul pipette with tips

Teams

- ○ PCR Sample
- ○ Gloves

Gulf Islands Research & Education Center | University of West Florida

PREPARING THE GEL

Figure 1. Gel casting tray with comb

1. Tightly tape both open ends of gel casting tray.

2. Calculate grams of agarose powder needed to make a **100mL, 1.0% Agarose Gel**.
    - *% agarose = (grams agarose / mL of buffer) x 100*
    - *Eg. 100ml of a 1.0% gel requires 1.0g of agarose*

4. Use graduated cylinder to transfer **100mL 1.0xTBE** into your Erlenmeyer Flask.

5. Place weight-boat on balance and tare the balance (balance will again read zero)

6. Weigh correct grams of agarose powder on balance in weigh-boat.

7. Pour agarose powder from weigh-boat into the flask.

8. Heat agarose/TBE in microwave for 45 seconds, or until barely boiling.

9. Reheat the solution for ~10s intervals until the solution is **COMPLETELY CLEAR.**

10. Use hot glove or folded paper towel around neck of the flask to return the heated (**VERY HOT!**) flask to your work station.

11. Run the flask under cool water until just hot to the touch.
    - *Be sure to not get any water in the flask to avoid diluting the gel*

12. Pipette the appropriate volume of syber green fluorescent gel stain into the cooled agarose/TBE solution (1.5ul syber green / 100ul TBE).
    - Wait until the flask is cool enough to touch before adding syber green
    - Because syber green is light and heat sensitive, make sure to return it to the fridge and store it in light proof container

13. Place the well-comb into the taped casting tray.

14. Pour agarose/TBE solution into casting tray and push any bubbles to the side of the tray.

15. Do not disturb tray again until gel is cool and has solidified (~20 minutes).

16. Once the gel has cooled carefully remove the comb and tape, then place the gel in the electrophoresis rig.

    - *Remember DNA "runs to red", so make sure your gel is aligned properly*

DURING CLASS: Running the gel

1) Place the gel in the electrophoresis rig with the wells closest to the black (negative) pole (DNA 'runs to red')

2) Add enough new or used 1x TBE to cover the gel

3) Load 2ul of the DNA ladder into the first well

4) Have each team load 5ul of their PCR sample in the remaining wells
    i) Remember to have each team fill out the gel map

5) Place cover on the gel rig and plug the electrodes into the power station
    i) Match red to red and black to black
    ii) Ensure the DNA Samples will 'run to red'

6) Run the gel at 120 volts for 20 minutes

7) Turn off the gel and unplug the electrodes

8) Carefully remove the gel
    i) The gel will be slippery and can easily slide off the gel tray

9) Slide the gel off of the gel tray onto the UV or blue light illuminator
    i) Turn out most/all of the overhead lights and turn on the illuminator
    ii) Photograph the gel (if using a blue light illuminator or a shielded UV illuminator, students can also photograph the gel)

ANALYZING YOUR GEL

Figure 1. Thermo Scientific Low Range DNA ladder

Use the DNA ladder in the first lane as a guide you can interpret the bands in your sample lanes to determine if your experiment (DNA extraction or PCR) worked as expected. See Figure 2 for the size standard banding pattern for the **Thermo Scientific Low Range DNA ladder**.

You should see a single, bright band that aligns with the appropriate location in the DNA ladder. The location of your band will depend on the size DNA fragment you amplified. No bands or multiple bands indicate the PCR did not work. Usually this is the result of (1) experimental error, (2) poor match between PCR primers and template DNA, or (3) too much DNA was added to the PCR.

Gulf Islands Research & Education Center | University of **West Florida**

Table 1. Gel map template

Well	Well 1	Well 2	Well 3	Well 4	Well 5	Well 6	Well 7	Well 8
Sample #								
Initials								

Well	Well 9	Well 10	Well 11	Well 12	Well 13	Well 14	Well 15	Well 16
Sample #								
Initials								

Note: Table should be modified depending on the number of samples you plan to run and rows you have set in the gel.

Introduction to DNA Sequencing using the Sanger method

DNA sequencing is the process of 'reading' the order of nucleotides in a provided DNA sample. Typically, the sample must contain concentrated, purified DNA (eg. CO1 PCR product), with all DNA in the sample being identical (eg. PCR product); however, recent technological advances now allow for simultaneous sequencing of mixed samples (generally referred to as Next-Generation Sequencing), though this process is not yet appropriate for smaller sample sets and as well requires considerable computational expertise.

Here we use traditional Sanger sequencing, which does not work on mixed samples (samples containing DNA from more than one gene or individual); however, the benefits of Sanger sequencing is that it is relatively affordable and does not require much in the way of specialized equipment or training. Sanger sequencing is, in fact, very similar to PCR with a few important exceptions (Table 1). Like PCR, DNA sequencing copies template DNA strands over and over using DNA polymerase, DNA primers, and DNA nucleotides in a buffered solution; however, DNA sequencing requires just a single DNA primer (rather than the 2 or more used in PCR), and requires using 'terminator' nucleotides in addition to the standard nucleotides.

Table 1. Comparison of the reagents required for PCR and Sanger sequencing

	PCR	Sanger Sequencing
Template DNA	Genomic DNA	Purified PCR product
DNA polymerase	Yes	Yes
DNA primers	2	1
Nucleotides	A, T, C, G	A, T, C, G
Terminator nucleotides	No	Yes

Terminator nucleotides

Differ from standard nucleotides in two ways; (1) a unique fluorescent marker is attached to each of the four nucleotides to allow for visual identification of the nucleotide once it has been incorporated into a new DNA strand, and (2) the hydroxyl (OH) has been removed from the 5-carbon sugar. Because the presence of the hydroxyl is required to allow DNA polymerase to continue extending the new DNA sequence, when a terminator nucleotide is added to the growing DNA sequence extension is stopped, and DNA replication is terminated.

### Why just 1 sequencing primer?

Unlike PCR, DNA sequencing only copies and then 'reads' one of the two complimentary strands of the template DNA. In essence, half of the DNA with the sample is ignored. This is because only one strand can be read at a time; however, since the complimentary strands carry the same information, the DNA sequence can be reconstructed from just one strand without losing any information.

How It's Done

The DNA sequencing reaction proceeds just like PCR (~35 cycles of denature, anneal, and extend); however, each time a terminator nucleotide is added to a growing DNA strand, replication is terminated for that strand because the missing hydroxyl makes it impossible for DNA polymerase to add any additional nucleotides. This means that by the end of the PCR reaction every fragment that has been created ends with a dye-labeled terminator nucleotide (Figure 1). This mixed pool of DNA fragments is then run through a polyacrylamide gel, which is very similar to an agarose gel except the gel pore size is considerably smaller, making the gel able to distinguish between fragments that differ in size by just one nucleotide. Like with an agarose gel, the DNA fragments are sorted by size as they move through the polyacrylamide gel (smaller fragments migrate more quickly than larger fragments).

As the fragments move through the gel they encounter a laser, which stimulates the fluorescent dye attached to each terminator nucleotide. A camera then records the observed fluorescent color, which corresponds to one of the four possible terminator nucleotides. The fluorescent signals are then displayed in the order they were observed (known as a chromatogram; Figure 1) and a computer program then 'reads' the colored peaks to determine the proper nucleotide sequence of the submitted sample.

Figure 2. Example fragments constructed during the sequencing reaction (left) with capillary polyacrylamide gel (right) and associated chromatogram (bottom). (figure credits: www.ambgood.com/marketing/knowledge-base)

Bioinformatics	Teacher's Guide
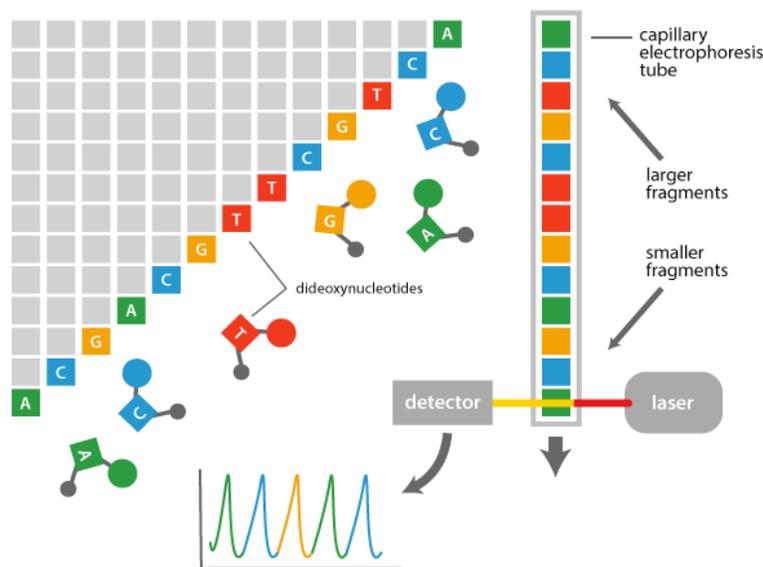
BEFORE CLASS

1. Make sure "Finch TV" is installed on each of the student's computers (http://www.geospiza.com/Products/finchtv.shtml)
2. Load sequence files (.ab1 files) onto each computer or onto several USB drives for students to upload at the start of class

DURING CLASS

- Be prepared to walk students through an example sequence clean-up and identification (see student's 'Bioinformatics handout' for detailed instructions)
- In general, students will need to..
    - I. Delete low quality DNA sequence from the beginning and end of their sequence (Figure 1 below)
    - II. Correct any sequencing errors in the middle of their sequence (see student handout)
    - III. Copy and paste their cleaned DNA sequence into the Barcode of Life Species Identification website to identify their unknown sample (see student handout)

ADDITIONAL RESOURCES

- Appendix 1: DNA Sequencing offers an introduction to the methodologies of DNA sequencing, including a comparison of the similarities and differences between DNA sequencing and PCR.
- For additional information on DNA barcoding… http://www.barcodeoflife.org/content/about/what-dna-barcoding
- If your students have access to Mac computers, you may find the program 4Peaks more user friendly than Finch TV (http://nucleobytes.com/4peaks/)

Figure 1. Example low quality region at the beginning of a DNA sequence. In the below example from the Finch TV website, the region from 0-45 base pairs should be deleted before submitting the sample for identification.
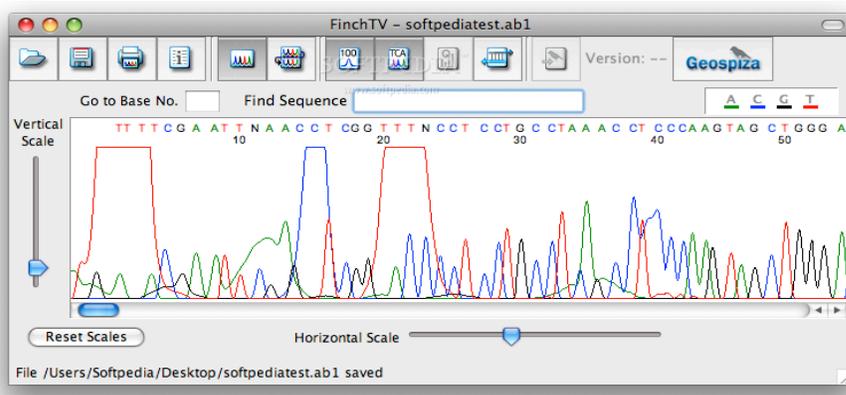
Discovering Biodiversity Case Study
A new deep-water shark from the Gulf of Mexico

Figure 1. The Cuban dogfish (*Squalus cubensis*; bottom) with the new, as yet unnamed, Gulf of Mexico *Squalus* species (formerly identified as *Squalus mitsukurii*; top)

Problem: The shortspine spurdog shark (*Squalus mitsukurii*) is a deep-water shark that is currently thought to occur globally (circumglobal). However, sharks of the genus *Squalus* are easily misidentified due to how similar they look (morphology) and recent research on this species from the Pacific has indicated that *S. mitsukurii* may actually be a group of many closely related species (species complex). To investigate whether or not *S. mitsukurii* is really more than one species we have compared the DNA of sharks collected from the Gulf of Mexico to specimens collected at sites throughout the Pacific Ocean.

Methods: We sequenced the mitochondrial (mtDNA) cytochrome c oxidase 1 gene (CO1) as well as another mtDNA gene (ND2). DNA sequences were generated for all sampled specimens to determine how genetically similar individual sharks were to each other (genetic distance). We as well collected morphological data on all specimens to see whether there were previously unidentified differences in shape, size, or structure between *S. mitsukurii* collected at different sites.

Findings: We found the Gulf of Mexio *S. mitsukurii* were on average nearly 3% genetically different from sharks collected in the Pacific (ie. 1 mutation for every 100 nucleotides examined). In addition, significant differences were found in morphological measurements between Gulf of Mexico specimens and Pacific specimens. These results reveal that there is a previously unidentified species of deep-water shark in the Gulf of Mexico.

Gulf Islands Research & Education Center | University of West Florida

Implications: While finding a new species is exciting; there is cause for concern, as this species is being caught as bycatch in the Gulf of Mexico fisheries. Species within this genus are not able to handle the pressures of heavy fishing pressure, and declines in populations have been documented in both Hawaiian and Australian dogfishes. New management strategies will need to be developed to reduce impacts on this new, and vulnerable, Gulf of Mexico deep-water shark.

Meet the scientist

Mariah Pfleger is originally from Detroit, Michigan and attended Florida State University for her undergraduate degree. She is currently at the **University of West Florida** seeking her Master's degree in Dr. Toby Daly-Engel's Shark Lab. She is most interested in Elasmobranch biology and conservation.

Discovering Biodiversity Case Study
Genetic Analysis of Hagfishes in the Gulf of Mexico

Figure 1. (A) Unknown Gulf of Mexico hagfish (family Myxinidae) with secreted 'slime' and (B) close-up of hagfish teeth which are currently one of the only ways to identify different hagfish species.

Problem: Hagfish are a seldom-studied fish that lives at the bottom of the ocean (up to 1200 meters, or 3600 feet, deep!). For many fish, telling species apart is based on physical appearance and structure (morphology), and there are easy to recognize characteristics to help (eg. forked tail or blue stripe). Hagfish all pretty much look the same: worm-like and slimy! The only way to tell them apart is actually tiny differences in their teeth, which don't always give a clear answer. Moreover, unlike most fish, where morphology of the internal organs (dissection) can give us information about how that fish reproduces, the insides of a hagfish are so strange that very little is known about how they reproduce!

Therefor the research questions I'm investigating are:
1. Can genetics help tell identify existing and new species of hagfishes?
2. Can we learn anything about hagfish reproduction from their genetics alone?

Methods: Hagfish samples are being collected from the Gulf of Mexico on research ships, and so far we have more than 1000 different fish sampled! Fin clips are taken and DNA is extracted. To tell apart species, I use mitochondrial markers Cytochrome Oxidase I (COI) and NADH Dehydrogenase II (NDII). The resulting DNA sequence data is then used to construct phylogenetic trees, which can be used to identify samples within the data set that represent different species (DNA barcoding). Results from DNA analyses are then combined with published descriptions of hagfish morphology to see if the species I've identified match those previously described, and if not, whether there exists one or more previously undescribed species.

Findings: So far I have determined that there are three genetically distinct species of hagfish in the Gulf of Mexico, but the work is ongoing and indications are that we will identify additional species. The genetic data also shows that hagfish visit sites throughout the northern Gulf of Mexico, which is pretty impressive for a fish that doesn't swim very well.

Implications: Hagfishes are detritivores, meaning they eat all the large dead marine life that reaches the sea floor. As detritivores, they play a critical role in nutrient cycling in the ocean. And as gross as it sounds, hagfish are actually eaten (hagfish sushi; not kidding). Even more are caught for materials: their skin is used as "eelskin" leather. In recent decades, many hagfish fisheries have been established and some of those fisheries have already collapsed (fished so many hagfish out that almost none remain). This means we may be losing species to extinction even before we've identified them. To limit this loss of critical biodiversity, it is important to understand how many species there are, where those species occur, and how those species reproduce, so that we can protect these fish in the future. They aren't pretty, but they are pretty important.

Meet the scientist
Rebecca Varney is a Biology Masters student at the **University of West Florida** in Dr. Toby Daly-Engel's Shark Lab. She plans to continue on to a PhD program after completing her Master's degree. She majored in Genetics as an undergraduate and is enjoying learning how to use genetic data to understand and protect marine biodiversity.

Gulf Islands Research & Education Center | University of West Florida