

UNIT 4. Analyzing DNA Barcodes

Overview: Analysis of DNA barcodes uses bioinformatics programs to compare the sequence data. This unit introduces users to two systems that are uniquely designed for DNA barcoding, Cold Spring Harbor Laboratory's DNA Subway and The Barcode of Life Data Systems - Student Data Portal.

1. *Introduction to Bioinformatics:* In order to analyze the DNA sequences, students must be familiar with the basics of bioinformatics. Bioinformatics uses computer programs to analyze large amounts of data (like DNA sequences). Genbank is the online database that stores all publically available DNA sequences. This first lesson (4.1a) introduces students to bioinformatics through an activity that investigates the phylogenetic relationship of different species of bears based on DNA sequences.
2. *Using The BOLD Student Data Portal:* The Barcode of Life Data Systems Student Data Portal is a powerful tool to store information about your sample. Sample information including location, collector, notes, images, and a map, are all organized on one page. BOLD also allows users to clean up their sequence data and input clean sequences and trace files onto sample pages. BOLD has great [video tutorials](#) and a [user-friendly guidelines](#) for students and educators.

Specimen - The Biodiversity Center of Belize-Masters-DNA Barcoding W... Print

IDENTIFIERS

Sample ID: 13-BIOB-Kc-12
 Process ID: SDP96014-13
 Institution: *unvouchered
 Storing:
 Field ID: 13-BIOB-Kc-12
 Museum ID:
 Collection:
 Code:


TAXONOMY

Identification: Chordata
 Rank: Phylum
 Identifier: Student
 Identification Method: Barcode
 Identifier:
 Institution:
 Identifier:
 Email:
 Taxonomy Note:
 Rank: Current Record (13-BIOB-Kc-12)
 Phylum: Chordata
 Class:
 Order:
 Family:
 Subfamily:
 Genus:
 Species:

SPECIMEN DETAILS

Voucher
 Status:
 Tissue
 Descriptor:
 Sex:
 Reproduction: Sexual
 Life Stage: Adult
 Extra Info:
 Note: The seller guaranteed that the fish is red snapper.
 Associated Taxa:
 Associated Specimens:
 Reference Link:
 Contributors: Specimen: Kenny Chan
 Image: Kenny Chan

PHOTOGRAPHS



License: Creative Commons Attribution ShareAlike (2013)
 License: Unspecified, The Biodiversity Center of Belize
 Holder:
 Comments: No Tags
 Associated Tags:

GEOGRAPHY

Country: Belize
 Province/State:
 Region/County: San Ignacio
 Sector:
 Exact Site: Savannah Street Fruits and Vegetable
 Lat/Lon: 17.1598, -89.0698

COLLECTION DETAILS

Collectors: Kenny Chan
 Collection:
 Event ID:
 Date Collected: 13-Jul-2013
 Date:
 Accuracy:
 Time Collected:
 Site Code:
 Habitat:
 Sampling Protocol:

3. *DNA Subway:* [DNA Subway](#) is a bioinformatics pipeline that was developed by Cold Spring Harbor Laboratories. It allows users to create new projects, upload sequence data and trace files, clean up sequences, and BLAST sequences to identify species. DNA Subway can also build phylogenetic trees

using Neighbor Joining or Maximum Likelihood methods. [DNA Subway's manual](#) is very helpful in moving students through the interface.

FAST TRACK TO GENE ANNOTATION AND GENOME ANALYSIS

Username:

Password:

[Forgot Password?](#) [Register](#)

D N A

S U B W A Y

NOW FOR
PLANTS
& ANIMALS

This site ties together key bioinformatics tools and databases to assemble gene models, investigate genomes, work with phylogenetic trees and analyze DNA barcodes. Roll over the "stations" on the subway map to find out more about the analysis steps. Analyze your own data or sample data provided. To start a project, select one of the "lines" (red, yellow, blue). Register and login to be able to save and share your results.

[▲ DNA Subway Training](#) [▲ DNA Barcoding 101](#)

[● Background](#) [● Manual](#) [● Tour](#)

[● About](#) [● Credits](#) [● Resources](#) [● Contact Us](#)

iPlant Collaborative

CSH Cold Spring Harbor Laboratory

4.1a. MOLECULAR PHYLOGENETICS

STEP 1: OBTAIN THE FOLOWING SEQUENCES FROM GENBANK

American black bear	Y08520
American brown bear	L21889
Spectacled bear	L21883
Asiatic black bear	L21890
Polar bear	L22164
Giant panda	Y08521

- (1) Enter Genbank: www.ncbi.nlm.nih.gov
- (2) Select Nucleotide on the pull-down search menu.
- (3) Enter accession number (e.g. Y08520) and select Go
- (4) Select FASTA to view the data
- (5) View the nucleotide sequence for this species in FASTA format
- (6) Highlight and copy the entire sequence including the ">" and paste into phylogeny.fr (Step 11)
 - Do this for each Accession number and any other sequences you want to add (THEY MUST BE THE SAME GENE!!!!)

STEP2: GET THE NUCLEOTIDE DATA INTO PHYLOGENY.FR SO THAT WE CAN ANALYZE IT

- (7) Open Phylogony.fr in a new tab (<http://www.phylogeny.fr/>)
- (8) Go to Phylogeny Analysis and click "Advanced"
- (9) Switch the choice from all at once to step by step
- (10) Click Create Workflow
- (11) Paste the FASTA sequence you copied into the textbox
- (12) When done inputting data click submit
- (13) Scroll through the Alignment Results to see where in the genetic sequences match and where there are differences
- (14) Go down to the bottom of the page and click next step
- (15) Do not change any of the Curation settings, just click submit
- (16) Scroll through the Curation Results to see where the sequences matched. How is it different from the alignment?
- (17) Go down to the bottom of the page and click next step
- (18) Leave the settings alone, and click submit
- (19) Look at your Phylogeny Results
- (20) Scroll down and click submit
- (21) Click submit again
- (22) What does the Phylogenetic Tree tell you? What can you infer about the evolutionary relationship between the species you have viewed?