Biodiversity Research in Undergraduate Lab Courses

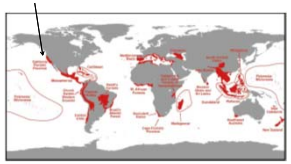
Stephanie Mel, Heather Henter, and Mandy Butler
Division of Biological Sciences, University of California, San Diego, La Jolla, CA 92093-0355

Abstract

As part of our effort to bring research into the undergraduate Biology labs at UCSD, we have initiated a project that seeks to document biodiversity in San Diego and to use this data to pose a variety of research questions. In lab classes, students collect specimens from the UCSD Reserves and the San Diego area, purify DNA from the specimens and then use PCR to generate a partial sequence of the mitochondrial cytochrome c oxidase (CO1) gene. This region of the CO1 gene has been designated as the standard barcode sequence by the Consortium for the Barcode of Life (CBOL). Students then evaluate and compare the DNA sequences in order to answer specific questions. This year we generated data for two different organisms, honeybees (*Apis mellifera)* and marine bristle worms (*Polychaete* species). The power of this general approach is that students can barcode a variety of organisms and then ask many different research questions. Students will contribute their sequences to biodiversity databases to create an inventory for the San Diego area.

Introduction

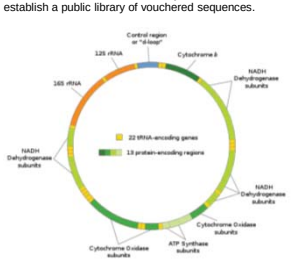
San Diego area is a biodiversity hotspot!



Earth's biodiversity remains incompletely documented and scientific estimates of the total number of species differ widely. Our lack of knowledge about the species around us is particularly troubling at a time when the human footprint is impacting virtually every organism on the planet. An inventory of the biota in the region surrounding UCSD is particularly urgent since the California Floristic Province, which covers most of California, is considered one of the 25 most threatened biodiversity hotspots in the world (Myers et al., 2000). More specifically, the county of San Diego has more rare and endangered species than anywhere in the continental US (Dobson et al., 1997). Yet much of this diversity remains poorly documented and has not been studied.

What is barcoding?

DNA barcoding is a method that uses a specific region within the cytochrome c oxidase (CO1) gene on mitochondrial DNA in animals as a genetic marker. The barcode sequence can be used to verify the identity of known species and to identify potentially new species. The Consortium for the Barcode of Life (CBOL) is an international collaboration whose mission is to compile DNA barcodes of known and newly discovered taxa and establish a public library of vouchered sequences.



General Strategy

Ecology students

Collect specimens in the field

Document collection location, date, zip code, and tentatively identify genus and species if possible

Preserve samples for DNA isolation

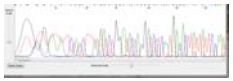
Recombinant DNA students

Purify DNA from collected specimens
We used Qiagen DNeasy kit

Set up CO1 gene PCR
We used validated primer sets from literature or BOLD

Run gel to verify PCR products and purity

Clean up PCR product and send for Sanger sequencing



Use bioinformatics tools to analyze sequences

Facts about *Apis mellifera*

- There are subspecies of the honeybee *Apis mellifera* that have been separated geographically until recently.
- Human-assisted introductions have caused the mixing of large populations of African and European subspecies in South and Central America, and there is evidence that the subspecies are cross breeding.
- Bees in the African subspecies are more aggressive, and there is evidence that the Africanized bees are moving into San Diego.
- There is a known assay that distinguishes Africanized bees from non-Africanized using the cytochrome B gene.
- However, no one has ever identified a marker for Africanized in the CO1 gene.

A CO1 marker for Africanized bees?

- Students performed a cleaved amplified polymorphic sequence assay (CAPS) on *Apis mellifera* cytochrome B PCR products to determine Africanization as described by Crozier et al. (1991).
- Students then examined chromatograms of their CO1 barcode sequences and assessed the quality of the sequences.
- If quality was acceptable, students trimmed ends and exported Fasta files of Africanized and non-Africanized bee sequences into ClustalW.
- Students then aligned CO1 barcode sequences from Africanized and non-Africanized bees and looked for single nucleotide polymorphisms (SNPs) in the CO1 gene that correlated with Africanization

Example of Clustal W alignment of *A. mellifera*

Note that all local non-Africanized bees have a T at the same position in the barcode sequence, while all Africanized bees have a C at that locus.

Facts about *Thoracophelia mucronata*

- Polychaetes, or bristle worms, are a common and diverse class of primarily marine worms with over 10,000 species described so far.
- *Thoracophelia mucronata* is a polychaete species found in abundance on San Diego beaches.
- Within this worm population, significant morphological differences have been observed.
- Furthermore, *T. mucronata* can be found in different locations within the intertidal zone.
- These differences call into question whether more than one species exists within the putative *T. mucronata* population.

Species diversity at the beach?

- Students examined chromatograms of their CO1 barcode sequences and assessed the quality of the sequences
- If quality was acceptable, students trimmed ends and exported Fasta files of different size *Thoracophelia muraconata* specimens from different parts of the intertidal zone.
- Students then aligned the CO1 barcode sequences using ClustalW and looked for SNPs to assess intraspecies diversity.

Example of Clustal W alignment of *T. mucronata*

Over 40 sequences were compared and all appeared to be members of the same species as shown by the low frequency of polymorphisms.

Project Summaries

Apis mellifera project

- DNA was isolated from legs from 60 local bees
- There were 55 positive CAPS results which demonstrated that 69% of local bees are Africanized
- Students generated 50 high quality CO1 sequences
- Students were able to identify a SNP within the CO1 barcode sequence that consistently correlated with Africanization.

Thoracophelia mucronata project

- DNA was isolated from 110 putative *T. mucronata* specimens from the Scripps Coastal Reserve and students generated 100 high quality CO1 sequences
- Students determined that all specimens, regardless of size or intertidal location, were members of the same species based on the very low frequency of CO1 polymorphisms
- Students also discovered that the CO1 sequence for *T. mucronata* is not in the Genbank or BOLD databases

Conclusions

- Our goal is to involve students enrolled in large undergraduate laboratory courses at UCSD in authentic research.
- The methodology needed to generate barcodes is relatively inexpensive and straightforward, and requires only basic molecular lab equipment.
- Barcoding can be used to address many different types of research questions and generate novel, useful, and potentially publishable scientific data.

Learning outcomes

By doing these labs, students are able to:

- Explain how species are identified and differentiated from one another using molecular biology techniques and bioinformatics.
- Explain why and how the cytochrome C oxidase gene sequence is used to identify animals in barcoding studies.
- Give examples of polymorphisms that can occur in DNA sequences.
- Isolate DNA and set up a PCR reaction, including designing primers.
- Explain the theoretical basis of Sanger sequencing and assess the quality of a sequencing run.
- Use a variety of bioinformatics databases and tools to analyze and compare DNA and amino acid sequences, including Blast and Clustal W.
- Communicate their research results in a written report..

References

http://www.barcodeoflife.org

Crozier et al. (1991) An improved test for Africanized honeybee mitochondrial DNA *Experientia* 47: 968-969

Dobson, et all (1997) Geographic distribution of endangered species in the United States. *Science* 275: 550–53.

Myers et al. (2000) Biodiversity hotspots for conservation priorities *Nature* 40: 853-858

Acknowledgements

This work is being supported in part by an NSF TUES grant #1140640