

The Use of DNA Barcoding to Teach Students the Importance of Classifying Biodiversity

Antonia Florio¹, Noemi Rivera², and Kathleen Nolan²

¹American Museum of Natural History, Central Park West & 79th St, New York, NY 10024

²St. Francis College, Department of Biology, Health Promotion, and Health Care Management, 180 Remsen St., Brooklyn, NY 11201

(antonia.m.florio@gmail.com, nrivera633@sfc.edu, knolan@sfc.edu)

Students can develop an awareness and concern for biodiversity loss by learning to identify the organisms around their neighborhoods or campuses. In response to this, we have implemented a DNA barcoding curriculum at St. Francis College (SFC) to identify species in New York City (NYC). Classifying species based on morphology can prove challenging for students because they often become frustrated by species descriptions and taxonomic keys. The use of DNA barcoding has been championed as a way to overcome this, while also providing an inquiry-based approach for student-driven research. Using the DNA barcoding method, students propose projects, collect samples, extract whole genomic DNA, and use PCR to amplify the appropriate gene for their taxonomic group (plants: chloroplast genes *rbcL* and *matk*; animals: mitochondrial COI; fungi: nuclear ITS; and bacteria: 16S rRNA). Successful PCRs (confirmed by gel electrophoresis) can then either be sequenced in-house or sent away to a company for Sanger sequencing. Students clean their data using any sequence editing program, and perform BLAST searches through GenBank to identify their samples. Students can also learn how to resolve evolutionary relationships by generating multiple sequence alignments and phylogenetic trees. This poster will summarize how we have used DNA barcoding at SFC to introduce high school and undergraduate students to classifying NYC biodiversity. The pros and cons of the method are presented, and a framework for implementing the project over a semester is proposed.

Keywords: DNA barcoding, DNA analysis, biodiversity

Introduction

Students can develop an awareness and concern for biodiversity loss by learning to identify the organisms around their neighborhoods or campuses. Previous studies have suggested that this can be accomplished by inspiring a connection with nature (Novacek, 2008) and by allowing students to develop a relationship with the species that surround them (Lindemann-Matthies, 2002; Tosh et al., 2016).

In response to this, we implemented a DNA barcoding curriculum at St. Francis College (SFC) to identify species in New York City (NYC). Classifying species based on morphology can prove challenging for students because they often become frustrated by species descriptions and taxonomic keys. The use of DNA barcoding has been championed as a way to overcome this, while also providing an inquiry-based approach for student-driven research. DNA barcoding, or sequence-

based specimen identification, was developed by Paul Hebert (2003) to identify a broad range of taxa by sequencing a standardized short DNA fragment, the “DNA barcode”.

Using the DNA barcoding method, students propose projects, collect samples, extract whole genomic DNA, and use PCR to amplify the appropriate gene for their taxonomic group (plants: chloroplast genes *rbcL* and *matk*; animals: mitochondrial COI; fungi: nuclear ITS; and bacteria: 16S rRNA). Successful PCRs (confirmed by gel electrophoresis) can then either be sequenced in-house or sent away to a company for Sanger sequencing. Students perform BLAST searches through GenBank to identify their samples and clean their data using any sequence editing program. Students can also learn how to resolve evolutionary relationships by generating multiple sequence alignments and phylogenetic trees.

This poster will summarize how we have used DNA barcoding at SFC to introduce high school and undergraduate students to classifying NYC biodiversity.

We focused on how DNA barcoding was used to identify plant biodiversity in two NYC wetland localities: Brooklyn Bridge Park and the Jamaica Bay Wildlife Refuge. Our study can be easily expanded to be included in many

undergraduate laboratory courses such as genetics, ecology, evolutionary biology, and conservation biology.

Student Outline

Objectives

- Formulate a project
- Collect biological specimens outdoors
- Isolate and analyze DNA from these specimens
- Use computer programs and online pictures and keys to identify specimens based on DNA sequences
- Generate phylogenetic trees with sequences
- Learn about the biodiversity in the NYC area

Introduction

You will first think of a biodiversity project you would like to explore. You will then collect samples from either a place of your choosing or from a place indicated by your instructor. You will either use a key or photos from the Internet in an attempt to identify your specimens. Your instructor might help you as well. You will then isolate DNA from your specimen using the instructions that come with a DNA barcoding kit. A DNA barcode is a unique signature that is either a piece of a chloroplast gene (*rbcL*) or a mitochondrial gene of an animal (cytochrome c oxidase subunit 1 or *CO1*, which are the ones that will be used in this exercise. You will follow the protocol given to you that comes with the kit; what you will do is mince and macerate tissue with a plastic pestle and a lysis buffer, incubate to further break down the cells, do a series of centrifugations--first to remove a pellet of “unwanted” material, then, after adding a silica resin that binds to the DNA, keep the pellet, wash, and, at the last step, elute the DNA from the pellet. Once you have isolated the DNA, you will add the appropriate primers to amplify the DNA, perform the amplification step, and then conduct gel electrophoresis with your samples and a marker to determine if you have the proper sized fragment (usually around 600 base pairs). If you do, then your instructor will send the remaining PCR product to a company (in this case, GENEWIZ) for Sanger sequencing. The company will send the sequences back to you electronically, and you will use these for a Basic Local Alignment Tool or BLAST to see with what genus or genera (or species) your sequence aligns. Once you have multiple species (say, plants, or insects) you can go to DNA subway, put your sequences into first the MUSCLE part on the website to get an alignment of your species, and then to the PHYLIP part to generate a neighbor-joining tree.

Methods and Data Collection

Step 1. Permission

Obtain necessary permission for sample collection—your instructor may have already done this.

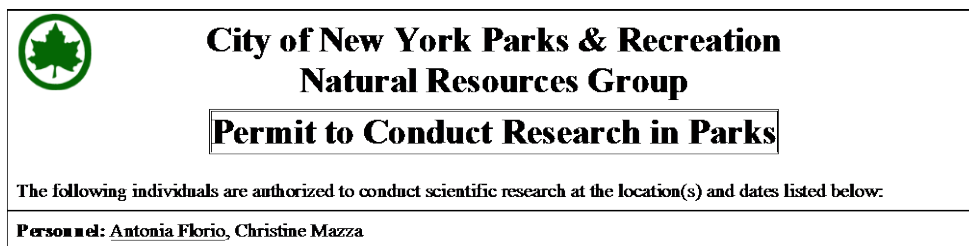


Figure 1. Sample collection permit.

Sampling permits for collection on city or state lands should be written for in advance because approval can take several months. For private lands, you can sometimes obtain permission directly from the owner (i.e., community garden board directors). See Figure 1 for sample permit.

Step 2. Sample Collection



Figure 2. At Brooklyn Bridge Park (left), students collected both animal and plant samples by seining. Only the plant samples were used for DNA barcoding (n=20). In Jamaica Bay, students collected leaves that had fallen in the parking lot (n=24).

Samples can be collected in test tubes, bags, jars, envelopes or other means. Two sample collection sites are shown in Figure 2, to give you an idea of what your collection sites might look like.

Step 3. Catalogue Samples



Figure 3. Sample catalogue.

All collected samples need to be catalogued before the wet lab portion of the procedure can begin. All samples should be given a unique identifier code and photos should be taken with a scale bar for later morphological identification. Figure 3 shows an example of how you could catalogue your specimen by taking a picture with your phone, and then saving it. Try to use a dichotomous key, such as this one available from the National Park Service, to identify your plant.

<https://www.nps.gov/goga/learn/education/upload/dichotomous.pdf>

Step 4. Wet Lab

DNA extraction, PCR, Gel Electrophoresis

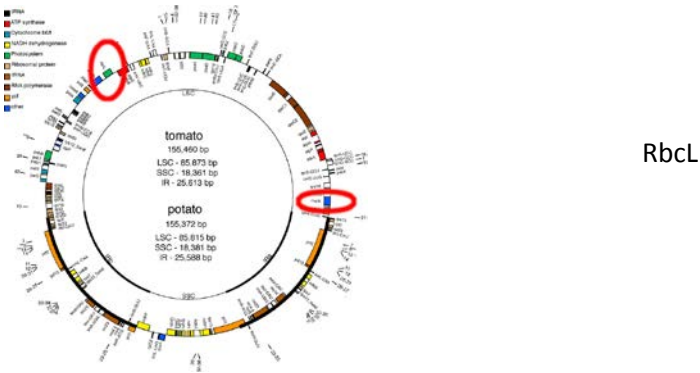


Figure 4. Chloroplast gene example.

Whole genomic DNA can be extracted using any protocol. We used a silica-based DNA extraction that can be ordered from Carolina Biological Co. Follow the protocol that comes with this kit.

In this case, we amplified the chloroplast *rbcL* gene using previously published primers and conditions (see Figure 4) (Kress, 2007) with a 54°C annealing temperature. Gel electrophoresis was used to confirm PCR success. The kit also comes with insect, fish, and mammal primers.

Step 5. Sanger Sequencing

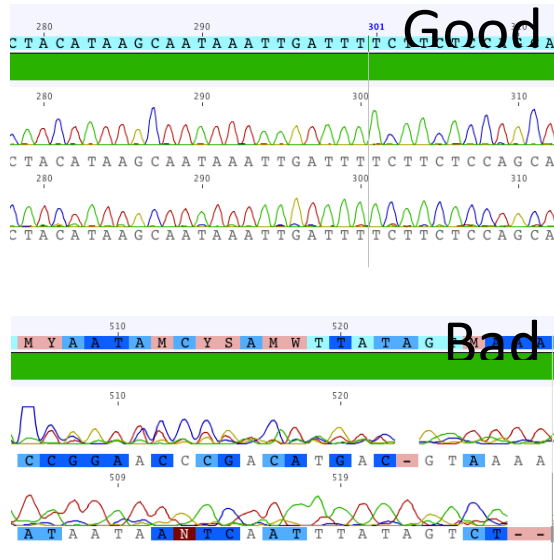


Figure 5. Examples of DNA sequence quality.

Samples can be sent away to various companies for sequencing. We used GENEWIZ, LLC. It usually takes a few days to get results. Your instructor will do this for you. Above (Figure 5) are examples of “good” and “bad” sequence quality

Step 6. BLAST Searches on NCBI GenBank

Description	Max score	Total score	Query cover	E value	Ident
<input type="checkbox"/> Ozoroa mucronata voucher A. Koopa 20 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds; chloroplast	928	928	98%	0.0	99%
<input type="checkbox"/> Cotinus coggygria ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, complete cds; chloroplast	928	928	98%	0.0	99%
<input type="checkbox"/> Brucea javanica ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, complete cds; chloroplast	928	928	98%	0.0	99%
<input type="checkbox"/> Toxicodendron sylvestri ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, complete cds; chloroplast	922	922	98%	0.0	99%
<input type="checkbox"/> Toxicodendron succedaneum ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, complete cds; chloroplast	922	922	98%	0.0	99%
<input type="checkbox"/> Toxicodendron trichocarpum ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, complete cds; chloroplast	922	922	98%	0.0	99%
<input type="checkbox"/> Toxicodendron vernicifluum ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, complete cds; chloroplast	917	917	98%	0.0	99%
<input type="checkbox"/> Mangifera indica cultivar Okrong ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds; chloroplast	911	911	97%	0.0	99%
<input type="checkbox"/> Mangifera indica cultivar Thunthawai Tha Wai ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds; chloroplast	911	911	97%	0.0	99%
<input type="checkbox"/> Mangifera indica cultivar Namdokmai ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds; chloroplast	911	911	97%	0.0	99%
<input type="checkbox"/> Mangifera indica cultivar Kao ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds; chloroplast	911	911	97%	0.0	99%
<input type="checkbox"/> Mangifera indica cultivar Khiaosawoey ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds; chloroplast	911	911	97%	0.0	99%
<input type="checkbox"/> Mangifera indica cultivar Haeo ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds; chloroplast	911	911	97%	0.0	99%

Figure 6. Sample BLAST results.

After the sequence data is retrieved from GENEWIZ, students then directly BLAST their results to search for similarities in GenBank on the National Center for Biotechnology Information (NCBI) website <https://www.ncbi.nlm.nih.gov/>.

To do this, put >sample DNA #___ at the top of your sequence, copy it, and paste it into the white box on the BLAST website. You then should obtain results similar to those in Figure 6.

Students sometimes become frustrated when they cannot get a conclusive result; for example, you can see in Figure 6 that the top BLAST matches are to three different plant genera, all with identical scores. There were several instances where we were unable to identify samples to the species-level using rbcL. If this happens, you can use another gene (such as matk or ITS) to help identify the plant, or you can rely on morphology to pinpoint the species identity. This is why taking sample photos is so important. For example, sample #7 (below) was identified (with identical NCBI scores) as belonging to potentially three genera, as seen below.

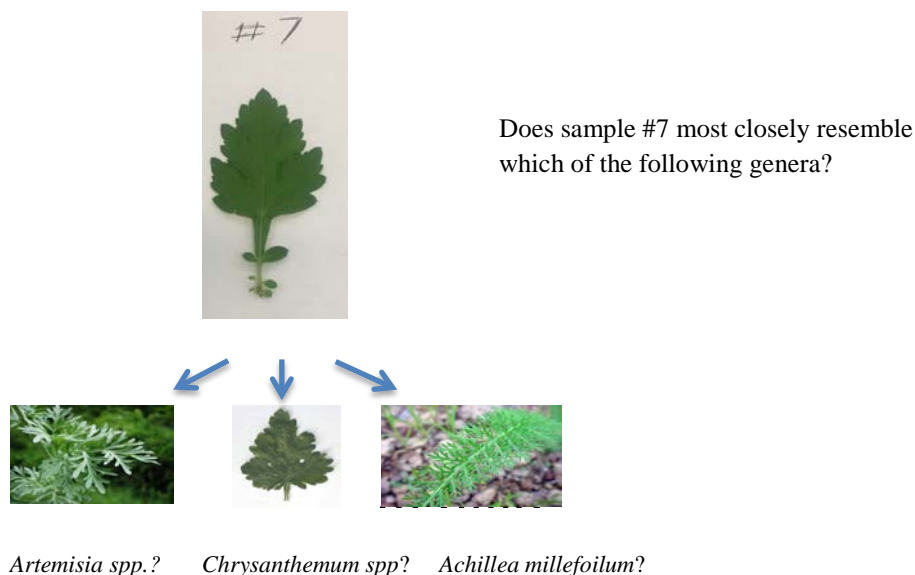


Figure 7. Choosing the correct plant.

You should be able to see that this sample most closely resembles a *Chrysanthemum* sp.! The next step is to identify the sample down to the species level using a taxonomic guide. As a word of caution, multiple species should be examined for each genus, because plant morphology can differ much even within a genus.

Step 7. Sequence Editing

There are many software options for sequence editing including

- Geneious
- MacClade
- DNA Subway is useful for an introduction to sequence editing



Figure 8. DNA subway (<https://dnasubway.cyverse.org/>).

In this case, we used DNA subway (see Figure 8). Directions for editing are in the Carolina Biology DNA Barcoding kit protocol. If those are not sufficient for you, try:

https://learning.cyverse.org/projects/dnasubway_guide/en/latest/step5.html

Once you have tentatively identified your DNA sequences to genus or species, replace the >sample DNA #_____ at the top of your sequence with its name; for example, >chrysanthemum. You should transfer all your sequences to a Word document with their names (either Latin or common). When prompted to copy all sequences into DNA subway, you can copy and paste from this master document.

Step 8. Data Analysis

DNA subway will allow you to first paste in your renamed sequences, then go to MUSCLE that will depict an alignment with a consensus sequence. If you click on box in upper right-hand corner of your picture of your alignment, a table will display that shows sequence similarity among each sequence. Then go to PHYLIP and click on “neighbor joining” and a tree will be displayed that will group your organisms through sequence similarity. In your analysis, you can state which species are the most similar and which ones are the most distant. You can also note which ones are introduced and which are native.

Discussion

This lab projects will take a few weeks---your instructor will give you the schedule. By the end of the project, which will result in either a group lab report or a poster (again, your instructor will decide), you will have collected and isolated DNA from organisms of your choosing, and, in doing so have learned more about local biodiversity and DNA analysis. Make sure that you familiarize yourself with terms that you do not know or remember, such as rbcL, CO1, DNA, PCR, gel electrophoresis, markers, Sanger sequencing, BLAST, consensus sequence, alignment, and a general understanding of neighbor joining. In your lab report, you should have pictures of your organisms, pictures of DNA sequences you obtained, a picture of your BLAST results, a picture of your MUSCLE alignment from DNA subway, and a neighbor-joining tree from DNA subway. Do some background checking on organisms, and tell us why they are here (introduced or native) and how well they survive with other organisms. For example, the white mulberry tree was introduced from Asia at the same time that silk worms were imported

here for economic purposes. The climate was not suitable for the silkworms, and they did not thrive, whereas the mulberry has done well.

Sometimes you will not get good DNA amplification, or the quality of the sequence will be poor. This seems to be unavoidable in the DNA world. This is why it is good that each person checks their lab partner's work, and that you work in teams to optimize data collection. Also, databases to compare your sequences against are always growing; Zúñiga (2017) attests to this point.

Cited References

<https://dnasubway.cyverse.org/>

https://learning.cyverse.org/projects/dnasubway_guide/en/latest/step5.html

<https://www.ncbi.nlm.nih.gov/>

<https://www.nps.gov/goga/learn/education/upload/dichotomous.pdf>

Kress WJ, Erickson DL. 2007. A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnh-psba* spacer region. *PLoS One*. 6, e508:1-10.

Zúñiga JD, Gostel MR, Mulcahy DG, Barker K, Hill A, Sedaghatpour M, Vo SQ, Funk VA, Coddington JA. 2017. Data Release: DNA barcodes of plant species collected for the Global Genome Initiative for Gardens Program, National Museum of Natural History, Smithsonian Institution. *Phytokeys* 88:119-122.

Materials

Collection tubes, jars, bags
 Permanent markers
 Nets or other collecting devices
 Forceps/scissors/scalpels
 Ice bucket
 Computer with Internet
 Vortexor
 Camera or photo-documentary system
 Electrophoresis reagents need to be purchased separately.
 (50X TAE buffer and agarose)
 Electrophoresis equipment (gel boxes, trays, dams,
 combs, power supplies)
 Microcentrifuges
 Micropipettors and tips
 Heating block or incubator (37 C and 65 C)
 Microfuge tubes (1.5 ml)
 PCR tubes
 PCR machine
 UV transilluminator
[SYBR™ Safe DNA Gel Stain](#) by Invitrogen
 Safety goggles

DNA Barcode Amplification Kits(with prepaid coupon) from Carolina Biological Company Item # 211385P. These kits have been developed in conjunction with the Cold Spring Harbor DNA Learning Center. This kit contains primers to isolate DNA from plants, mammals, fish, or insects (*rbcl* or *COI* genes).

Fungal Primer Set (item #211388, is sold separately).

Carolina Biological has partnered with, GENEWIZ, LLC., which offers an educational discount for sequencing services. it contains materials for 25 DNA extractions and PCR reactions.

Kit contents:

Lysis buffer
rbcl primer plus loading dye
Fish COI primer plus loading dye
Insect-mammal primer plus loading dye
 pBR322 /BstNI markers
 Silica resin
 Wash buffer concentrate
 Sterile distilled water
 Ethanol, 95%
 Mineral oil (not needed if PCR tubes have caps)
 Ready-to-Go™ PCR Beads in 0.2 ml PCR tubes
 Plastic pestles for grinding tissue
 Teacher's Manual with Reproducible Student Guide

Notes for the Instructor

DNA barcoding can easily be implemented in a lab course by following the steps outlined here. For the fall semester, we suggest the following timeline:

Obtain permit by end of summer: Step 1
 Week 1: Steps 2+3—Planning project, collecting and cataloguing samples
 Week 2: Step 4 (DNA extraction + PCR)
 Week 3: Step 4 (Gels) + Step 5 (Send samples)
 Week 4: Steps 6-8 Data analysis

Sample Results

Overall Amplification Success

Gel electrophoresis was used to confirm PCR success: in the Brooklyn Bridge Park, 8/20 (40%) samples were amplified and in Jamaica Bay, 18/24 (75%) samples were amplified.

Brooklyn Bridge Park Samples

6/8 samples (75%) could be identified to the genus level. 2/8 samples (25%) could not even be identified to the family level. Identified genera included:

- *Lepidium* spp.: pepperwort
- *Oxalis* spp.: “false shamrocks”
- *Parietaria* spp.
- *Ulva* spp.: “sea lettuce”

See Figure 9.



Ulva spp. *Lepidium* spp.

Figure 9. Specimens from Brooklyn Bridge Park

Jamaica Bay Wildlife Refuge Samples

5/18 samples (28%) could be identified to the genus level, while most plants 13/18 (72%) could not be identified to the genus-level, and many not even to the family-level. Those that could be identified to the genus-level included:

- *Ilex* spp.: Holy
- *Lonicera* spp.: Honeysuckle

- *Prunus* spp.: includes plums, peaches, nectarines

See Figure 10.



Ilex spp. *Lonicera* spp. *Prunus* spp.

Figure 10. Species found at Jamaica Bay Wildlife Refuge

Overall. We found that:

- PCR success rate varied: 40% (BBP) vs. 75% (Jamaica Bay)
- rbcL is not great for identifying species (<50% were identified to even the genus-level).
- The identified genera in the two localities varied
- The projects were successfully completed over 10 days with high school students, and in 4 lab sessions with college students
- All participating students were introduced to taxonomic methods by completing real research

Cited References

- Hebert PDN, Ratnasingham S, de Waard JR. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. R. Soc. Lond. B* 270 S96-S99.
- <https://dnasubway.cyverse.org/>
https://learning.cyverse.org/projects/dnasubway_guide/en/latest/step5.html
<https://www.ncbi.nlm.nih.gov/>
<https://www.nps.gov/goga/learn/education/upload/dichotomous.pdf>
- Kress WJ, Erickson DL. 2007. A two-locus global DNA barcode for land plants: the coding rbcL gene complements the non-coding trnH-psbA spacer region. *PLoS One* 6 e508:1-10.
- Lindemann-Matthies P. 2002. The Influence of an Educational Program on Children's Perception of Biodiversity. *J Environ Educ* 33:22-31.

Novacek MJ. 2008. Colloquium paper: engaging the public in biodiversity issues. *Proc Natl Acad Sci U S A* 105 Suppl:11571-11578.

Tosh J, James K, Rumsey F, Crookshank A, Dyer R, Hopkins D. 2016. Is DNA barcoding child's play? Science education and the utility of DNA barcoding for the discrimination of UK tree species. *Botanical Journal of The Linnean Society* 181(4):711-722.

Zúñiga JD, Gostel MR, Mulcahy DG, Barker K, Hill A, Sedaghatpour M, Vo SQ, Funk VA, Coddington JA. 2017. Data Release: DNA barcodes of plant species collected for the Global Genome Initiative for Gardens Program, National Museum of Natural History, Smithsonian Institution. *Phytokeys* 88:119-122.

Acknowledgments

We are thankful for the help and guidance provided by the following people: Dr. Alison Dell at St. Francis College, Dr. Christine Marizzi, Melissa Lee and Alison Cucco from Cold Spring Harbor Laboratory's DNA Learning Center, Isa Del Bello and Christina Tobitsch from Brooklyn Bridge Park, and Geri Korbyn-Blatterl from the Jamaica Bay Wildlife Center.

About the Authors

Antonia Florio, Ph.D. developed this exercise while she was an assistant professor at St. Francis College. She is currently a science educator at the National Park Service in Florida and is an associate scientist at the American Museum of Natural History

Noemi Rivera graduated from St. Francis College in 2012. She is a professional advisor and departmental assistant. She teaches general biology labs and is an instructor in the St. Francis College Summer Science Academy. This is her first ABLE, and she is very excited to learn about new materials to use in her teaching.

Kathleen A. Nolan, Ph.D. is a professor of biology and Chair of the Biology and Health Sciences Department at St. Francis College. She has been a long-time ABLE member and has presented numerous major and mini-workshops at ABLE conferences. She is interested in a wide variety of topics, including fish population genetics, animal vocalizations, and biology laboratory education.

Mission, Review Process & Disclaimer

The Association for Biology Laboratory Education (ABLE) was founded in 1979 to promote information exchange among university and college educators actively concerned with teaching biology in a laboratory setting. The focus of ABLE is to improve the undergraduate biology laboratory experience by promoting the development and dissemination of interesting, innovative, and reliable laboratory exercises. For more information about ABLE, please visit <http://www.ableweb.org/>.

Papers published in *Tested Studies for Laboratory Teaching: Peer-Reviewed Proceedings of the Conference of the Association for Biology Laboratory Education* are evaluated and selected by a committee prior to presentation at the conference, peer-reviewed by participants at the conference, and edited by members of the ABLE Editorial Board.

Citing This Article

Florio A, Rivera N, Nolan KA. 2018. The use of DNA barcoding to teach students the importance of classifying biodiversity. Article 67 In: McMahon K, editor. *Tested Studies for Laboratory Teaching*. Volume 39. Proceedings of the 39th Conference of the Association for Biology Laboratory Education (ABLE). <http://www.ableweb.org/volumes/vol-39/?art=1>

Compilation © 2018 by the Association for Biology Laboratory Education, ISBN 1-890444-17-0. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the copyright owner. ABLE strongly encourages individuals to use the exercises in this proceedings volume in their teaching program. If this exercise is used solely at one's own institution with no intent for profit, it is excluded from the preceding copyright restriction, unless otherwise noted on the copyright notice of the individual chapter in this volume. Proper credit to this publication must be included in your laboratory outline for each use; a sample citation is given above.