Engaging First-Year Undergraduates in Hands-On Research Experiences: The Upper Green River Barcode of Life Project

By Jeffrey M. Marcus, Tia M. Hughes, Douglas M. McElroy, and Robert E. Wyatt

To improve retention and engagement, first-year college science majors enrolled in University Experience orientation courses participated in a hands-on laboratory research experience: a DNA barcoding project to facilitate species identification. Students collected arthropods and hypothesized morphology-based species identifications. Then they isolated DNA from specimens, performed polymerase chain reaction (PCR), sequenced the product, and performed a DNA-based identification, to test the morphological hypothesis. Participants in this exercise earned higher biology GPAs and were more likely to become a science major or remain enrolled at Western Kentucky University than students in a control section who did not participate in this research experience.

A great challenge for college science educators is to engage first-year undergraduates and encourage them to pursue science majors. This is a priority for many departments, yet attrition is high, with many students switching to non-science subjects or dropping out of college altogether (NRC 2006; Schroeder 1998). This has led many instructors to move away from unimaginative, cookbook-style laboratory exercises and large, uninspiring, and intentionally difficult lecture courses (Arwood 2004). Alternative strategies to interest and involve students include using group-based problem-solving exercises, problem-based learning, case studies, interactive computer learning, and virtual or inquiry-based laboratories (Handelsman et al. 2004; Hoskins, Stevens, and Nehm 2007; Lundeberg and Yadav 2006).

One promising way to communicate the excitement of scientific research and to engage students to become active learners, which should improve student motivation and retention, is to have students participate in novel research as part of their laboratory coursework (Buckner et al. 2007; DebBurman 2002). However, most implementations of “research in the teaching laboratory” occur in upper-level courses with smaller enrollments of students who have already committed to careers in science. Here we extend this approach to first-year students in order to improve retention and inspire students to become science majors. We are biologists by training and used a DNA barcoding project for this purpose but expect that similar projects could be designed for other science disciplines.
The issue of first-year attrition among prospective science majors is part of a broader problem. First-year attrition represents the majority of all undergraduate dropouts, regardless of course of study (Daugherty and Lane 1999). Extended semester-long freshman orientation courses that include information on academic policies, college regulations, career counseling, student services, and study skills have been shown to improve freshman retention and overall academic success of participants (Jones 1984; Noble et al. 2007). These courses typically involve small groups of students, teach important skills and convey useful information, assist students in creating academic relationships with peers, and help students to establish meaningful personal connections with faculty. In order to maximize the usefulness of these relationships to students, many institutions group students by discipline and assign them to orientation sections led by instructors from their intended major.

At our institution, first-year students are encouraged to take an optional semester-long University Experience orientation class of no more than 20 students that meets once a week for an hour and 50 minutes. Students can choose between generic and discipline-specific University Experience courses, and versions of both types are available for students enrolled in the university honors program (for which students are selected by high school GPA and American College Test [ACT] scores). Because these classes are small in size and have a flexible curriculum at our institution, we recognized that the discipline-specific versions of the University Experience courses are an unusual opportunity to incorporate a “research in the teaching laboratory” experience into the education of first-year science students.

In the fall 2006 semester, the Biology Department offered three discipline-specific sections of University Experience. All three sections included the basic University Experience curriculum, including explanations of university policies and services, study skills, and career counseling. In addition, for two sections containing 12 and 14 students, respectively, we implemented a novel research experience. A third section of the class (containing 20 students) included only the standard University Experience curriculum, did not participate in the research experience, and served as a comparison group. A different instructor was assigned to each of the three sections. Students were not told in advance about the project, nor did they know which sections were participating, so they chose sections irrespective of whether they were to participate in the research experience.

We designed a Barcode of Life project for the student research experience. The Barcode of Life initiative is a worldwide consortium of researchers who are attempting to collect unique DNA sequence identifiers for all species of eukaryotic organisms to facilitate taxonomy and specimen identification (Hebert et al. 2003). The ultimate goal is to have multiple DNA sequence isolates from geographically diverse populations of each species on Earth. Thus every additional sequence, particularly from localities not sampled by previous workers, represents a valuable addition to the Barcode of Life initiative. Thus far, work in animals has focused on a 648 base pair segment of mitochondrial cytochrome oxidase I gene (COI), for which nearly universal polymerase chain reaction (PCR) primers are available (Folmer et al. 1994). Projects within this initiative have generally focused either on covering a taxonomic group (e.g., birds, fish; Hebert et al. 2004; Kerr et al. 2007; Marshall 2005) or on creating an inventory of all of the species present in a particular locality (e.g., Great Smoky Mountains National Park, United States, and Área de Conservación Guanacaste, Costa Rica; Burns et al. 2007; Hajibabaei et al. 2006; White 2005).

We took the latter approach, initiating a DNA barcode project for the arthropods found at the Upper Green River Biological Preserve in Hart County, Kentucky. The preserve encompasses approximately 741 acres (300 ha), lies 1.9 miles (3 km) upstream of Mammoth Cave National Park, and is home to six federally listed endangered species. A DNA barcoding project is ideally suited to first-year undergraduates; it allows students to choose an organism that interests them, it is designed to include both fieldwork and laboratory work, the experimental techniques involved are not technically difficult, and the experimental protocols are sufficiently “high tech” to impress students.

Materials and methods
Experimental procedures
An in-depth description of the class exercise is available elsewhere (Marcus et al. 2009), so it will not be described in detail here. Briefly, a special four-hour-long Saturday class meeting was arranged to do the field collections on September 16, 2006. For those students unable to meet on Saturday, the Friday before was made available as an alternative. Arthropods were collected with handheld butterfly nets and a battery powered ultraviolet light trap, which operated overnight (Winter 2000). Arthropods were selected because they are diverse, abundant, easy to capture, and require no veterinary oversight, and a large library of taxonomic keys and field guides was available for these organisms. Global Positioning System (GPS) coordinates were recorded for each specimen collected. Each student chose one specimen for the barcoding project. Specimens were taken back to the laboratory, mounted on insect pins, digitally photographed, and identified based on morphology using keys and field guides; a single leg from each specimen was removed and frozen for DNA extraction. The morphology-based species determinations made by each of the students served as the identification hypothesis that students then tested via analysis of DNA sequence data.

All molecular biology manipula-
were used (Marcus et al. 2009). PCR failed to produce a product for a given fragment of the COI gene by agarose gel electrophoresis. If PCR failed to produce a product for a given sample, alternative PCR primer sets were used (Marcus et al. 2009). PCR products were sequenced directly using an automated fluorescent capillary DNA sequencer. Sequences were edited and compared with the National Center for Biotechnology Information (NCBI) database by BLASTN (Altschul et al. 1997) and with sequences in the BOLD database (Ratnasingham and Hebert 2007) in order to identify each specimen on the basis of DNA sequence. DNA sequence-based identifications were then used to test the hypothesized species identification based on morphology.

### Example of a species web page produced by a student participant in the Upper Green River Barcode of Life project.

**Bio 175**

**Upper Green River Barcode of Life Project**

**Collector:** Nicole Long

**Collection Date:** September 16, 2005

**Scientific Name:** Apantesis phlaraisa

**Common Name:** Harassed Tiger Moth

**TAXONOMY**

**Phylum:** Arthropoda

**Class:** Insecta

**Order:** Lepidoptera

**Family:** Arctiidae

**Genus:** Apantesis

**Species:** phlaraisa

**Collection Locality:** Upper Green River Biological Preserve, Gallatin Parish, South Side, Hermit County, Kentucky, USA

**Organizer Image:** A picture of a species of moth.

**Organizer Biology:**

- Body length: 25-28 mm
- Color: orange and black dots along the inner margin
- Head: pinkish color
- Wings: black and orange

**Dna Sequences:**

- ATCATAGAC CAAATACCT TATGTTG AGCGGTTGGA ATTACAGCT TrTTATTACT CTTTCATTrAG
- GAACAGGATG AACCGTGTAC CCCOCACTTT
- TrTATAACCTA AACATTATAT TATTIATrTG GAATTTGAGC AGGTATAGTA

**Location Map:**

- **Coordinates:** 37°16.6668$, 86°30.8935
- **Google Maps Location:**

**Assessment of outcomes**

In December 2007, one year after completing the University Experience course, each student’s composite ACT score upon entrance to college and the GPA of each student in his or her first-year biology course was determined. Two-tailed t-tests were used to compare students between participating sections and between participating and nonparticipating students. Linear regression and analysis of covariance (ANCOVA) were used to examine the relationship between composite ACT score and first-year biology GPA. In addition, the major and enrollment status of each student in participating and nonparticipating sections were assessed. Students were classified as being science majors, non-science majors, or no longer enrolled at the university, and the number of students in each category was determined for participating and nonparticipating students. A chi-squared contingency test was used to compare the participating and nonparticipating student groups, using the relative frequencies of the nonparticipating students in each category (science, non-science, no longer enrolled) to calculate expected values.

**Electronic resources**

Students generated a web page about their organism, including classification, digital photograph, GPS coordinates (with links to Google Earth), maps, DNA sequence, BLASTN or BOLD identification results, a description of the biology of the organism, and references. To standardize the web pages produced by students, a set of ASP (Active Server Pages) applications was developed to allow students to easily enter, edit, and view their web pages without any prior knowledge of web programming. This helped students build basic computer literacy skills and develop confidence working with unfamiliar software. The gateway to this online resource is available at [http://bioweb.wku.edu/faculty/Marcus/Barcode.html](http://bioweb.wku.edu/faculty/Marcus/Barcode.html), and an example of a species page produced by a student participant is shown in Figure 1.
Results

Scientific outcomes
A total of 26 students participated in the first implementation of the Upper Green River Barcode of Life project, and sequences for 28 specimens were generated (see Figure 2). Most specimens (25 of 28) could be amplified and sequenced successfully using the COI primers. The remaining 3 specimens required alternative primers for successful PCR amplification. Sequences generated by this project have been deposited in Genbank, accession numbers EU271647-EU271674.

Of the 25 sequences that amplified successfully with the standard COI primers, 17 of them definitively matched sequences in the Genbank database, the BOLD database, or both, allowing identification at the species level by DNA sequence. Two additional sequences matched at the level of genus but are the first sequenced exemplars for their species. A further four sequences did not match with anything in the DNA databases even at the level of genus and are probably the first sequences for this part of the COI gene for their respective genera. Molecular identifications generally confirmed students' identifications based on morphology, though in the case of several spiders and lepidopteran larvae, the molecular identifications suggested that the specimens belonged to different families than was originally hypothesized.

Educational outcomes
The mean first-year biology GPAs for the two sections that participated in the Barcode of Life exercise were 3.64 (N = 12, standard deviation = 0.36) and 3.57 (N = 14, standard deviation = 0.23), whereas the mean for the nonparticipating section was 2.36 (N = 20, standard deviation = 1.20). A t-test between the two participating sections showed no significant difference between them (p = .608), so they were combined for further analyses. A t-test between the combined participating sections and the nonparticipating section showed a statistically significant difference between them (p = .000186).

However, the mean composite ACT score for participating sections (mean = 28.08, standard deviation = 2.19) was higher than that of the nonparticipating section (mean = 23.15, standard deviation = 4.41), and the difference is significant (two-sample, two-tailed t-test with unequal variances, p = .000007). To compare the participating and nonparticipating treatments, controlling for this correlation, an ANCOVA was performed with composite ACT score as the concomitant variable and first-year biology GPA as the dependent variable. The ANCOVA showed that the participating and non-participating treatments are still significantly different from one another after taking composite ACT into account (F = 7.27, p = .00997). To be valid, ANCOVA assumes homogeneity among within-treatment regressions. For this analysis, within-treatment regressions were not significantly different from each other (F = 1.34, p = .25358) and can therefore be considered homogeneous, meeting the assumptions of ANCOVA.

The numbers of science majors, nonscience majors, and students no longer enrolled at our institution for participating and nonparticipating sections is summarized in Table 1. A chi-squared contingency test comparing these student success data suggests that participants in the project were significantly more likely to remain a science major or to remain at
Engaging First-Year Undergraduates

Students inspecting vegetation for insect herbivores.

Our institution than nonparticipants ($\chi^2 = 7.46, df = 2, p = .024$).

Discussion

University Experience and similar extended first-year orientation courses are a good opportunity to introduce classroom-based research experiences into undergraduate curricula either as pilot projects for later introduction into courses with larger enrollments or as a permanent component of these classes. The small class size and flexible curriculum make doing experimental manipulations tractable and allow the instructor to easily telegraph enthusiasm about science to students. The attributes of a successful structured inquiry exercise of this type include the availability of many samples from which students can choose, each of which can be expected to produce unique and individually useful results; the manipulations that each student performs should be substantially similar and amenable to performance by novices; and the availability of a medium (such as a website, database, or repository) through which the student data can be shared beyond the confines of the classroom. We chose to use DNA barcoding of species in our course, but students could just as easily document and interpret the emission spectra of stars in an astronomy course, or map the contamination plume of mine tailings in an environmental chemistry course.

For biology courses, DNA barcoding experiments are extremely well suited for use as exploratory research projects in a teaching laboratory. By including both fieldwork and molecular biology techniques, they engage students with interests that vary from environmental science to biochemistry, and they connect the age-old tradition of natural history study with modern research methods. This exercise shows students interdisciplinary connections between biology and fields as diverse as geography, photography, and computer science. The experiment allows students to determine what organism they wish to study, providing an investigative component to the exercise (McKenzie and Glasson 1997). Students seem even more involved and enthusiastic if they collect and select which samples to analyze.

Our DNA barcoding project allowed students to go through an entire cycle of the scientific method (Gower 1996). First, they used preliminary observations of the morphology of an organism to generate a hypothetical species assignment. Then, they conducted an experiment (PCR and sequencing) from which they collected data to test their hypothesis. Following that, students interpreted the data by comparing the DNA sequences that they generated with those that other researchers collected to come up with a DNA-based species assignment. Comparing the two species determinations allowed them to test their morphology-based hypothesis. If the two species

<table>
<thead>
<tr>
<th>TABLE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of participants and nonparticipants in the Bar Code of Life project, December 2007.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Students</th>
<th>Science majors</th>
<th>Nonscience majors</th>
<th>No longer enrolled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participating</td>
<td>19</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Nonparticipating</td>
<td>10</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

assignments disagreed, then students needed to account for the disagreement by reevaluating all of the available data and generating a further hypothesis. By generating web pages for each of their species, students take part in a forum for making these determinations as well as receive credit and "authorship" for their contributions as individuals. More formally, students are also coauthors on a manuscript submission describing this project to a regional peer-reviewed journal (Marcus et al. 2009). Finally, the data collected by students was deposited in the international web-based sequence repository Genbank and can now be used by the worldwide community of researchers as part of the genetic barcode library. The best verifications of morphological identifications were in those taxa for which much sequence data were already available, but the greatest contributions students made toward the international barcoding initiative (Hebert et al. 2003) were in those taxa for which few sequences are available.

In terms of educational outcomes, participating students had statistically significant higher first-year biology GPAs than did nonparticipants. Although students were not told which sections were participating in the project and could not directly self-select, the different sections of the class met on different days and at different times, so it is possible that scheduling conflicts with other classes may have inadvertently stratified enrollments according to ability, as evidenced by the significant difference between the composite ACT scores in participating and nonparticipating sections. Because there is a significant positive correlation between composite ACT score and first-year biology GPA, ANCOVA was used to show that participating students performed significantly better than did nonparticipants, after controlling for differences in composite ACT scores.

Although the data are clear that students in participating sections were more academically successful, we do not wish to overinterpret these results. The three sections (two participating sections and one nonparticipating section) of our University Experience course were taught by three different faculty members, and we did not control for differences in instructor effectiveness. We feel that differences in instructor effectiveness are likely to be a minor factor, because the two participating sections did not differ significantly from each other, and all three of the instructors have consistently received favorable teaching evaluations and assessments. Further, the faculty member assigned to the nonparticipating (and low-performing) section is widely recognized as an extremely effective biology educator and has received multiple awards for her teaching. All three instructors have since been reassigned to other teaching responsibilities and no longer teach sections of University Experience, so we cannot repeat this comparison to better control for these factors. In spite of this limitation, our overall interpretation of the data is that participation in the Upper Green River Barcode of Life project has a measurable positive effect on student success.

The Barcode of Life project has now been integrated as part of the laboratory component of one of our mainstream (nonhonors) first-year biology courses for majors, and students continue to add species to the database. The Upper Green River Biological Preserve has been very appreciative of students' work and is eager for the project to continue to include additional species. The accumulation of data and web pages from each class will gradually create a rich resource that workers at the Upper Green River Biological Preserve and elsewhere can consult for specimen identification and species distribution information.

**Acknowledgments**

We thank Scott Grubbs, Albert Meier, Ouida Meier, and Mike Stokes for facilitating our work at the Upper Green River Biological Preserve. We also thank Ann Cutler, Bruce Grant, Joanne Steff, and three anonymous reviewers for helpful suggestions.

**References**

Engaging First-Year Undergraduates

Jeffrey M. Marcus (marcus@cc.umanitoba.ca) is an adjunct professor in the Department of Biology at Western Kentucky University, Bowling Green, Kentucky, and an associate professor in the Department of Biological Sciences, University of Manitoba, Winnipeg, Canada. Tia M. Hughes was a research associate in the Department of Biology at Western Kentucky University and is now a research assistant II at the Center for Human Genetics Research, Vanderbilt University Medical Center, Nashville, Tennessee. Douglas M. McElroy is a professor, and Robert E. Wyatt is an associate professor in the Department of Biology at Western Kentucky University.
COPYRIGHT INFORMATION

TITLE: Engaging First-Year Undergraduates in Hands-On Research Experiences: The Upper Green River Barcode of Life Project
SOURCE: J Coll Sci Teach 39 no3 Ja/F 2010

The magazine publisher is the copyright holder of this article and it is reproduced with permission. Further reproduction of this article in violation of the copyright is prohibited. To contact the publisher: http://www.nsta.org/college