Tested Studies for Laboratory Teaching
Proceedings of the Association for Biology Laboratory Education
Volume 38, Article 6, 2017

Examining Microbial Biodiversity in Soil: A Large-Enrollment Introductory Course-Based Undergraduate Research Experience

Stanley M. Lo[1,2] and Stephanie F. Mel[3]

[1]University of California San Diego, Section of Cell and Developmental Biology, 9500 Gilman Dr. #0355, La Jolla CA 92093-0355 USA
[2]University of California San Diego, Program in Mathematics and Science Education, 9500 Gilman Dr. #0355, La Jolla CA 92093-0355 USA
[3] University of California San Diego, Section of Molecular Biology, 9500 Gilman Dr. #0355, La Jolla CA 92093-0355 USA
(**smlo@ucsd.edu; smel@ucsd.edu**)

We describe an introductory course-based undergraduate research experience in which students collect and contribute original data to an on-going research project. The goal of this course is to help students develop an understanding for research in biology through hypothesis-driven laboratory experiments. The research project takes place at a biodiversity hotspot and examines how soil microbiomes associated with different plants may differ. Working in collaborative teams, students measure soil properties and correlate them to functional and genetic biodiversity of the microbiomes. The course addresses both macroscopic and molecular concepts, providing an opportunity for students to make connections across different biological disciplines.

Keywords: course-based undergraduate research experience, soil microbiome, biodiversity, ecology, molecular biology, sequence alignment, data analysis, Microsoft Excel®, CURE on soil microbiomes

Introduction

In the past few decades, national reports have called for research-based learning in undergraduate education (Boyer 1998, Rutherford 1991). The President's Council of Advisors on Science and Technology recommends the development of course-based undergraduate research experiences (CUREs) to engage all students in authentic research in the first two years of college (PCAST 2012). A variety of CUREs currently exist in the biology education research literature (Butler 2014, Hanauer 2006, Kloser 2011, Kloser 2013, Taylor 2010). However, many of these examples are designed for advanced undergraduates (Butler 2014, Caspers 2003, Murthy 2014, Taylor 2010), and some are limited to small numbers of students (Kloser 2011, Kloser 2013). Here, we describe a large-enrollment CURE for first-year undergraduates that assumes no prerequisite biology background or laboratory experience.

In introductory biology courses, discussions of biodiversity often focus on the taxonomic descriptions and characterizations of macroscopic organisms such as plants and animals (Freeman 2011, Reese 2014). Students may not realize that most species on Earth have yet to be identified, especially microbial species that are not visible or immediately accessible by conventional culture techniques (Fierer 2006, Riesenfeld 2004, Rout 2013). In this CURE, we introduce students to the microbial world and examine biodiversity in terms of functional and genetic parameters rather than taxonomy. While learning how to study biodiversity using a variety of approaches, students collect original data that contribute to a larger, on-going research project on soil microbiomes.

The research project has both discovery-based and hypothesis-driven objectives. In the long term, we aim to understand longitudinal changes in the characteristics of soil microbiomes at the Natural Reserve System at the University of California San Diego, which is located at one of the 25 biodiversity hotspots in the world (Myers 2000). Within each iteration of the course, we have short-term objectives driven by specific hypotheses, so that students can engage in hypothesis testing. Currently, we are examining how soil microbiomes associated with

different plant species may differ. Our working hypothesis is that soil microbiomes of native and invasive plants are functionally and genetically different (Wolfe 2005, Inderjit 2010), and students use basic inferential statistics to test the corresponding null hypotheses (Krzywinski 2013, Krzywinski 2014, Whitley 2002). As soil microbiome data from single plants do not allow us to draw conclusions about native vs. invasive plants in general, we plan to analyze soil samples from different native and invasive plants from the same natural reserve in coming years, such that more general conclusions can be drawn from these data.

Three key sets of research data are collected and analyzed to examine soil microbiomes: soil properties (moisture and pH), functional biodiversity (sole carbon source utilization by Ecoplate^{TM}), and genetic biodiversity (16S ribosomal RNA gene sequencing). With the anchoring research question on soil samples associated with native vs. invasive plants, this CURE connects concepts across biological scales from macroscopic to microscopic to molecular. While the individual components may be common to existing laboratory courses (Butler 2013, Mulcahy 2007, Skwor 2012), together these elements provide students the opportunity to synthesize their learning from potentially different biology courses. In addition, this CURE focuses on quantitative literacy and scientific reasoning, as students learn to analyze novel research data and draw conclusion from these data.

The CURE described in this paper constitutes an independent, on-going research program to which all introductory biology students contribute original data. We have established a set of scientific and learning objectives for this quarter-long course, and modular subsets of these objectives can easily be achieved in less time. With the general setup of this CURE, potentially any research questions related to microbial biodiversity can be examined by beginning undergraduate students.

Student Outline

Genomic DNA Extraction and Analysis

We will use the ZR Soil Microbe DNA MiniPrep kit from Zymo Research to extract genomic DNA from our soil samples. To maintain the composition of microbial communities as they are in the wild, we will freeze aliquots of soil samples until we are ready for DNA extraction. After DNA extraction, we will determine DNA concentration.

Extracting Metagenomic DNA from Soil

This protocol is written for two DNA samples (or any other even number of DNA samples) so that microtubes will always be balanced in centrifugation steps. It is important to make sure that samples are balanced in the centrifuge.

1. Thaw the soil aliquots at room temperature if frozen.
2. Transfer each soil aliquot to a ZR Bashing Bead™ lysis tube.
3. Add 750 µL lysis solution to each tube. Screw caps onto tubes tightly.
4. Vortex at maximum speed for 5 minutes.
5. Centrifuge lysis tubes at 10,000 g for 1 minute to separate the supernatant from the pellet. Be careful not to disrupt the pelleted Bashing Beads™, soil, and cell debris.
6. Prepare 2 debris filters: Snap off the base of 2 Zymo-Spin™ IV spin filters and insert each filter into a clean collection tube.
7. Transfer 400 µL of each supernatant to the Zymo-Spin™ IV spin filters that are sitting in collection tubes. Afterward, set aside the Bashing Bead™ tubes, as the tubes are no longer needed.
8. Centrifuge the Zymo-Spin™ IV spin filters at 7,000 g for 1 minute to remove residual Bashing Beads™, soil, and cell debris. Remove Zymo-Spin™ IV spin filters from collection tubes, as the filters are no longer needed. Keep the collection tubes with filtrate.
9. Add 1,200 µL of soil DNA binding buffer to the filtrate in each collection tube.
10. Twist off lids from stock microtubes and cap collection tubes with these lids. Vortex briefly for 5-10 seconds to mix.
11. Insert 2 Zymo-Spin™ IIC DNA binding columns into clean Collection Tubes and then transfer 800 µL of each filtrate into these columns.
12. Centrifuge at 10,000 g for 1 minute. Discard flow-through.
13. Return the DNA Binding columns to the collection tubes and transfer the remaining 800 µL of each filtrate into the DNA binding columns.
14. Centrifuge again at 10,000 g for 1 minute and discard the flow-through.
15. Return the DNA binding columns to the collection tubes and add 200 µL DNA pre-wash buffer to each Zymo-Spin™ IIC columns.
16. Centrifuge at 10,000 g for 1 minute. Discard flow-through.
17. Return the DNA binding columns to the collection tubes and add 500 µL DNA wash buffer to each Zymo-Spin™ IIC column.
18. Centrifuge at 10,000 g for 1 minute. Discard flow-through.
19. Transfer Zymo-Spin™ IIC DNA binding columns to new 1.5-mL collection tubes.
20. Add 40 µL DNA elution buffer directly to each DNA binding column.
21. Centrifuge at 10,000 g for 30 seconds to elute DNA.
22. Remove the DNA binding columns from the collection tubes and set them aside. Do not discard the eluted DNA that is in the collection tubes.
23. Prepare 2 clean-up columns: Snap off the base of 2 Zymo-Spin™ IV-HRC spin filters. Insert each into a collection tube. Centrifuge Zymo-Spin™ IV-HRC spin filters at 8,000 g for 3 minutes to remove the buffer that they are stored in.
24. Transfer each spin filter (clean-up column) to new 1.5-mL collection tube.
25. Transfer eluted DNA from Step 22 to the Zymo-Spin™ IV-HRC spin filters.
26. Centrifuge at 8,000 g for 1 minute.
27. Combine the 2 purified DNA samples that came through the clean-up columns into a single 1.5-mL microtube. The filtered DNA is suitable for subsequent applications.
28. Label microtubes appropriately on the lids.

Determining DNA Concentration

We will determine DNA concentration using a compact spectrophotometer called NanoDrop Lite. DNA concentration and purity are determined by absorbance (or optical density, OD) at 260 and 280 nm.

1. Blank the NanoDrop Lite with 1.5 µL DNA elution buffer.
2. Add 1.5 µL of your DNA sample to the NanoDrop Lite pedestal. Read OD at 260 and 280 nm.
3. Record DNA concentration and OD values in the research notebook and shared Google spreadsheet.
4. Store the remaining genomic DNA at -20°C until the experiment.

16S Ribosomal RNA Gene Sequence Analysis

We will submit our transformed *E. coli* colonies for DNA sequencing at a local company, Eton Biosciences, which will perform the sequencing reactions using the universal T7 promoter primer. This primer sequence is located on the plasmid about 50 base pairs from where the PCR product is inserted.

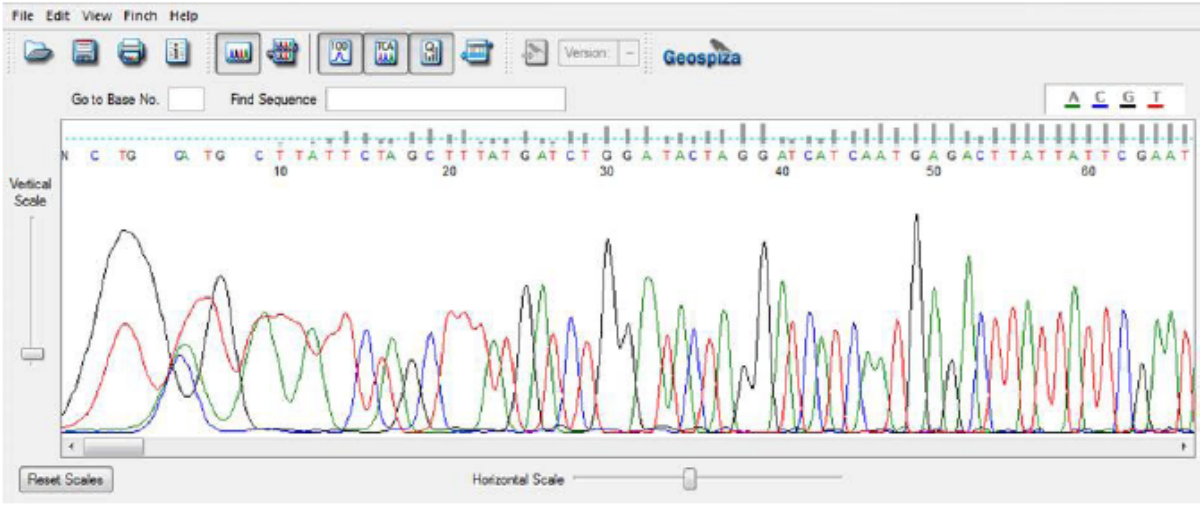
We will analyze the DNA sequence results to characterize genetic biodiversity in the soil samples. First, we will check that the DNA sequencing reactions have generated high-quality data using the chromatogram viewing program Finch TV (http://www.geospiza.com/ftvdlinfo.html). We will then clean up all the usable sequences and convert them to a format that can be read by an online sequence alignment software called SINA (http://www.arb-silva.de/aligner/). Next, we will use SINA to compare the 16S ribosomal RNA gene sequences from our samples to an existing database of known 16S sequences. The output from SINA will be the taxonomic classification (kingdom, phylum, class, order, genus, and species) of the microbial organisms found in our soil samples.

Checking the Quality of DNA Sequencing Reactions

1. Open each .ab1 sequence file in Finch TV (already installed on the laboratory computers).
2. Check the .ab1 sequences for quality of the DNA sequencing reactions.

The color peaks in the chromatogram represent actual sequence of the PCR product. There are four colors, each representing a different DNA base. Typically in a usable sequence, the first 20-50 peaks have overlapping colors, and the subsequent peaks are well resolved. The gray bars above each base tell us the quality of the sequence at that particular base: the higher the bar, the more certain the computer program is about identifying that base. There is also a horizontal dotted line that indicates the threshold for identifying a base. In an unusable sequence, many peaks overlap, and very few gray bars are above the threshold.

Example of a usable sequence:



Removing Ambiguous Peaks from Usable Sequences

1. For each usable sequence, determine the low-quality bases within the first 20-50 peaks. Highlight these low-quality peaks and delete them.
2. Determine the low-quality bases toward the end of the sequencing reaction, typically after 500-800 bases. Highlight these low-quality peaks and delete them.
3. Go to File, Export, DNA Sequence: FASTA.
4. Set up a single .txt file in Notepad with all of the trimmed usable sequences together in the FASTA format, which is recognized by SINA and other sequence alignment programs. Save the .txt file.

   In the FASTA format, each sequence begins with ">", followed by the sequence in the next line. Combine the usable sequences from the corresponding .seq files into a single .txt file. Label each sequence with section number, group number, "seq", and sequence number (example below).

   >A01Aseq001
   TAGTCGTAGTCGTAGTCGCATGCTCCGGCCGCCATGGCGGCCGCGGGAATTCGATTAGAGTTTGAT
   …

Identifying Microbial Organisms Using SINA

1. Go to the SINA website (http://www.arb-silva.de/aligner/).
2. Upload the .txt file containing the FASTA sequences.
3. Set the basic alignment parameters to "SSU" and "attached to the last aligned base".
4. Check the search and classify box.
5. Set the minimum identity with query sequence to 0.85
6. Set the number of neighbors per query sequence to 10.
7. Select output settings as FASTA format and no compression.
8. Set "reject sequence below identity" to 70%.
9. Type in the .txt file name (section number, group number, and "16S") as the job name.
10. Run the SINA aligner.
11. Toggle between "display scores" and "display classification" to examine our results. The "LCA tax. SILVA" column identifies the best taxonomic classification of a particular input sequence.
12. Export the data to CSV. Open the file in Excel.
13. Highlight the first column in Excel. Reorganize the data into columns by doing the following steps. This will change each line of text into separate columns for each data field.
    a. Select "Text to Columns" under the data menu.
    b. Select the option "Delimited".
    c. Select the option "Semicolon".
14. Record these data in the research notebook.
15. Further expand the last column (column U: "lca_tax_slv") into separate columns using the same method as above.
16. Record column C and columns U to AA in the shared Google spreadsheet. Enter "Unclassified" for cells in columns U to AA that are missing information.

Column	Excel column name	Taxonomic rank	Example
C	sequence_identifier	---	01Aseq2
U	lca_tax_slv	Kingdom	Bacteria
V	---	Phylum	Actinobacteria
W	---	Class	Rubrobacteria
X	---	Order	Rubrobacterales
Y	---	Family	Rubrobacteriaceae
Z	---	Genus	Rubrobacter
AA	---	Species	Unclassified

Genetic Biodiversity

Richness and evenness are two measurements that we can use to quantify biodiversity within a community. In the context of our soil microbiome research project, richness is measured by the number of different phyla identified for each soil sample. Evenness is a measured by the relative proportions of the different phyla represented. The Shannon diversity index H is a composite measurement that takes both richness and evenness into account. A higher H value indicates a more diverse community.

Shannon Diversity Index (H) = $-\Sigma\ p_i \times \ln(p_i)$
p_i = the proportion of a particular phylum being represented out of the total number of observed sequences

Calculating Shannon H

1. From the class aggregate data, identify each phylum present in all soil samples.
2. Count the number of sequences represented in each phylum for each soil type. Use the COUNTIF function to do this.
3. Calculate richness (S) by counting the total number of phyla in the soil samples taken from a given plant. Again, use the COUNTIF function to do this.
4. Add up the total number of sequences in the soil sample. Use the SUM function to do this.
5. Divide the number of sequences in each phylum by the total number of sequences to determine the proportion (p_i) of each phylum in the community.
6. Take the natural log of each p_i value. Calculate $p_i \times \ln(p_i)$ for each phylum.
7. Calculate Shannon H by multiplying the sum of all $p_i \times \ln(p_i)$ values by -1.
8. Repeat calculations for the soil samples taken from other plant types.
9. Compare the Shannon H values between the different soil types.

Materials

Materials for all three portions of the soil microbiome projects are listed below (Table 1). In our setup, each laboratory section has up to 32 students working in eight teams, and the required materials are counted based on the number of teams.

Table 1. Materials for 32 students in 8 teams of 4 students each.

Experiment	Equipment and supplies	Count
General	Balance	4
	Bunsen burner	8
	Drying oven: 105°C	1
	Incubator: 4-37°C	1
	Microcentrifuge	8
	Micropipette	8 sets
	Microplate reader: OD_{590}	1
	Microtube	Stock
	NanoDrop Lite	1
	Vortex mixer	8
	Water bath	1
Soil collection	Soil aliquots	6 g per group
	Soil sampler: 1” diameter	1 per soil type
Soil moisture	Glass vial	8
Soil pH	Glass vial	8
	pH meter	4
DNA extraction	ZR Soil Microbe DNA MiniPrep kit	8 preps
Gel electrophoresis	1-kb DNA ladder	Stock
	1× TAE buffer	Stock
	Agarose	Stock
	Gel apparatus and power supply	8
	SYBR safe DNA stain	Stock
	UV light box and camera system	1
PCR	16S primers: 8F and 1492R	10 µM stock
	GoTaq Green® master mix	32 reactions
PCR cleanup	Column-Pure PCR Clean-Up kit	16 reactions
Ligation	pGEM®-T Easy Vector system	32 reactions
Transformation	Mix-and-Go *E. coli* Transformation kit	400 µL of cells
	LB/Amp/X-gal plates	64
DNA sequencing	DNA sequencing company	1
	Inoculation loop	16
	LB/Amp/X-gal plates	16
Sequence analysis	Computers with Finch TV	16
	Internet connection	1

Notes for the Instructor

Course-Based Undergraduate Research Experience

CUREs are defined by five critical elements: the use of scientific practices in inquiry, collaboration among students, iteration of scientific processes, discovery of novel findings, and broader relevance of such discoveries to the larger research community (Brownell 2015). In this CURE, we deliberately build in all five components throughout the course (Table 2).

Table 2. Alignment of CURE elements and course activities.

CURE elements	Course activities
Scientific practices	Students collect and analyze original data to draw conclusions
Collaboration	Small teams of students collaborate and share research data across teams
Iteration	The first paper is used as a draft and is incorporated into the second paper
Discovery	Original research data on soil microbiomes are collected by students
Relevance	Research question is of interests to professional researchers (Rout 2013)

Learning Objectives

This introductory CURE aims to help students develop an understanding of research in biological sciences through hypothesis-driven laboratory experiments. Students work in collaborative teams to collect, analyze, and present original research data while learning laboratory methods common to a variety of disciplines in biological sciences. Specifically, the learning objectives are listed below and are aligned with corresponding course activities and student assessments (Table 3).

- Understand the multifaceted nature of biodiversity
- Develop testable hypotheses and design controlled experiments
- Analyze research data using basic descriptive and inferential statistics
- Draw conclusions based on data and supported by reasoning
- Communicate scientific results in laboratory reports written as journal-style papers

Table 3. Alignment of learning objectives, activities, and assessments.

Learning objectives	Course activities	Student assessments
Understand functional and genetic biodiversity	Discuss material in lecture and laboratory	Quizzes in lectures, laboratory reports
Develop testable hypotheses and design controlled experiments	Carry out scientific processes in the context of the course	Data collection, laboratory reports, quizzes in lectures
Analyze research data using basic statistics	Analyze and discuss data within and across collaborative teams	Figures and results in laboratory reports, quizzes
Drawing conclusions based on data and reasoning	Compare soil microbiomes using statistics	Results and discussion in laboratory reports, quizzes
Communicate scientific results in publication-style papers	Write, review, and revise laboratory reports	Laboratory reports, peer review of reports

Structure of the Research Project

The research project on soil microbiomes is divided into three major portions: soil properties, functional biodiversity by Ecoplate™ (Choi 1999, Garland 1997, Preston-Mafham 2002, Zak 1994), and genetic biodiversity by 16S ribosomal RNA gene sequencing (Tringe 2005, Weisburg 1991). In the previous section, we include the student outline for parts of the genetic biodiversity portion, as presented at the 38th Annual ABLE Meeting at the University of Houston. In this section, we briefly describe all three portions of the research project.

Soil Properties

Plant soil samples are collected by teaching assistants and then divided into aliquots for student teams. This ensures the quality of sample collection and prevents the ecosystems from being disrupted by a large number of students. Nonetheless, students take a guided tour to the Natural Reserve System, where we discuss native plant adaptation and identify plants at specific sites from which soil samples are collected. In principle, soil samples can be from any source, and the proximity to a Natural Reserve System is not required to implement this CURE or elements of this CURE.

The soil moisture and pH measurements involve basic laboratory equipment such as balances and vortex mixers. These experiments are scheduled at the beginning of the course so that students can be oriented to the laboratory early on. Students also begin using statistics in comparing soil moisture and pH, and similar statistical analyses are then practiced throughout the course. In addition, the percent moisture calculation and the concept of pH relate to topics commonly learned in General Chemistry courses, providing opportunities for students to make cross-disciplinary connections in their learning.

Functional Biodiversity

The Ecoplate™ system can be used to measure functional biodiversity of a microbiome (Zak 1994). It contains a tetrazolium salt as an indicator of metabolism. As metabolism is activated, a non-specific reductase reduces the colorless tetrazolium salt to a purple compound, which absorbs light optimally at 590 nm. Absorbance at 590 nm is an indirect measurement of metabolism activated by a specific carbon source. Some caveats of such a coupled reporter system may include variable expression of the reductase across microbial species, which cannot be controlled for in this type of experiment.

Genetic Biodiversity

In the current setup, each team of up to four students analyzes 6-8 16S ribosomal RNA gene sequences. This distribution of labor was determined based on the number of DNA sequences required to make comparisons among microbiomes from different soil samples. Typically, about 100 sequences are required per sample. Instructors adopting this CURE, especially those who are interested in the potential of collecting meaningful research data, should consider the number of DNA sequences and soil samples in relation to the number of students in the course.

For PCR amplification of the 16S ribosomal RNA gene, the 8F and 1492R primer set (pre-made stocks from IDT DNA) seems to be quite robust. We tested the PCR protocol with annealing temperatures ranging from 45°C to 65°C. The quantity and specificity of PCR products generated at 45°C to 55°C annealing temperatures are not noticeably different on an agarose gel. At 65°C annealing temperature, the expected product is still generated but in much lower quantity.

For DNA sequencing, we use a local company, Eton Biosciences, with a negotiated educational rate. Other companies provide similar DNA sequencing services, and most of them have standard sequencing primers available that are suitable for sequencing from the pGEM®-T Easy vector and other standard vectors. In our experiments, we use a standard primer for the T7 promoter, which primes near the multiple cloning site on the pGEM®-T Easy vector and sequences the insert region.

For 16S ribosomal gene sequence analysis, a specialized database and algorithm, SINA, is used instead of a more conventional tool such as BLAST. SINA directly provides identification of operational taxonomic units based on existing sequences that most closely match the query sequence (Pruesse 2012). On the other hand, BLAST simply returns all available matches without curation, which could be overwhelming for introductory students. SINA also outputs the data in a format that is easily transformed and organized into a Google spreadsheet for students to share their data across teams.

Laboratory Protocols

To mimic the type of protocols that biologists would encounter in research laboratories, protocols are provided in a laboratory manual organized by experiments rather than individual weeks. The

protocols are intentionally concise while still providing enough information for introductory students. The purpose is to help students learn to read protocols similar to those provided by vendors. To facilitate this learning, the instructor discusses in lectures the theory underlying specific laboratory methods, and teaching assistants explain the experiments and demonstrate different equipment at the beginning of each laboratory session. Students are also asked to read the protocols before laboratory sessions and submit graphic summaries of the protocol as pre-laboratory assignments. These graphic summaries are then used as guides in the laboratory. As students work in teams, they are asked to compare their graphic summaries with teammates and discuss the protocols before beginning the experiments, with the intention of promoting active learning of laboratory methods and developing collaborative process skills.

The protocols are designed to visit repeatedly the idea of controls in experimental design. In many different experiments (such as Ecoplate™, PCR, ligation, and transformation), students are asked to determine if a specific reaction is an experiment or control. The goal is to provide deliberate practice (Ericsson 1993) to help students understand the importance of positive and negative controls in biological experiments and to engage them in thinking about well-controlled experiments.

Collaboration is a key aspect of CUREs and authentic research (Brownell 2015) and is thus a focal point of the laboratory protocols. Students work in teams and are asked to keep a single, shared electronic research notebook within each team. Each team records all their data and thought processes in a Google document, and these research notebooks are communal property accessible by all students and teaching assistants, as well as the instructor, in the course. Teams are responsible for carrying out their own protocols but are asked to analyze data collected by and shared across all the teams. This provides a unique opportunity for students to analyze larger data sets unlike conventional laboratory courses, in which students typically analyze their own data. Some activities also promote collaboration between teams, such as cross-checking Ecoplate™ calculations. Together, these activities aim to foster active learning of data analysis and presentation.

Prerequisite Student Knowledge and Skills

As an introductory CURE, the context of the research project is designed to be approachable by students with minimal background in biological sciences. The research question focuses on soil microbiomes associated with different plants, which connect the microscopic and macroscopic scales of biology. The central focus of biodiversity is discussed twice in the course, first as functional biodiversity measured by Ecoplate™ and then as genetic biodiversity measured by 16S ribosomal RNA gene sequencing. This approach also provides students with an opportunity to learn the material using different examples, thus constructing their own understanding of core ideas (Gick 1983). Students can also revise their knowledge when they encounter these concepts for a second time (Chi 2014), therefore minimizing the requirement of prerequisite knowledge.

The experiments are designed to help students develop a set of core laboratory skills to promote learning situated in an authentic research context (Lave 1991, Wenger 1998). We make no assumptions that students have previous experiences in the laboratory setting. Students learn to use basic equipment such as micropipettes, balances, centrifuges, vortex mixers, and spectrophotometers early in the course and continue to practice these skills throughout the research project. These skills are intentionally interleaved throughout the course to promote contextual and inductive learning (Goode 1986, Kornell 2008). Students learn fundamental laboratory methods in modern biology, such as colorimetric assays, DNA extraction, PCR, gel electrophoresis, molecular cloning, and DNA sequencing.

Students learn to analyze and present original research data using Microsoft Excel®, a skill that is likely useful in many careers beyond the laboratory setting. Again, we make no assumptions that students have previous experience with the software. The course is designed such that students learn to use Microsoft Excel® for a wide range of purposes with increasing sophistication through the research project, such as performing basic arithmetic and statistics, generating graphs and figures with appropriate labels and legends, and using counting and logic functions.

Student Assessments

Summative assessments include two papers written in teams in the style of research journal publications. The papers are built collaboratively throughout the course, and substantial laboratory time is dedicated to discussions on how to write different sections of these papers. The first paper focuses on the functional biodiversity experiments. Student teams receive feedback from teaching assistants and can incorporate changes into their final papers. The second paper consists of both the functional and genetic biodiversity experiments, as well as the soil properties data, thus challenging students to synthesize

information and draw conclusions about soil microbiomes using multiple sources of data.

The papers are written in teams of up to four students and are limited to one or four pages, for the first and second papers respectively. Both of these mechanisms allow for scaling up of such a course in terms of grading, especially since the power of the research project lies in the parallel data collection and analysis by a large number of students.

Formative assessments are built into many laboratory activities, such as comparing graphic summaries of protocols within teams and comparing Ecoplate™ data analysis across teams. In addition, some laboratory sessions in the course are not fully scheduled with experiments and data analysis associated with the soil microbiome research projects. During some weeks, the majority of laboratory time is dedicated to discussions and workshop activities. These workshops provide a scaffold for students to develop their understanding of experimental design, data analysis, data presentation, statistics, and scientific writing. Laboratory time is also scheduled for peer review of student-generated graphs and figures.

Possible Modifications

Sources of Microbial Samples

The current CURE is set up to examine soil microbiomes associated with different plant species and to collect longitudinal data on soil microbiomes at the University of California Natural Reserve System. However, natural reserves and soil samples are not necessarily required for these experimental approaches. Aquatic and marine sources are also rich in microbial organisms (Gibbons 2014, Moran 2015), and in principle, any source with microbial organisms can be chosen based on the interests of the instructor and students, as well as local availability of such sources. For example, fermented foods such as cheese have been used as model systems to study microbiomes, as well as microbial functions and interactions (Wolfe 2014, Wolfe 2015). Winogradsky columns, commonly used in microbiology education, represent another accessible source of microbiomes for this type of experiment (Esteban 2015, Rundell 2014).

Additional Experiments and Data Analysis

Because of time constraints in the 10-week quarter system, we are unable to perform more in-depth analysis on the 16S ribosomal DNA gene sequences. For example, as an extension of the data analysis, students can construct phylogenetic trees to examine the evolutionary relationships among the prokaryotic organisms in different soil samples (Kaplan 2013, Turner 1999, Woose 1991). This added experiment would engage students in more sophisticated bioinformatics and connect the soil microbiome research project to the concept of evolution, a core concept in biology.

Data analysis from the Ecoplate™ experiments can also be extended to enrich the learning experience for students. At present, we perform simple calculations on biodiversity (Marshall 2008, Pruvis 2000). However, data from the Ecoplate™ system represents a wealth and depth of quantitative analyses that go beyond the basic descriptive and inferential statistics used in this CURE, for example, principle component analysis to compare Ecoplate™ data from different samples (Weber 2007, Weber 2009, Weber 2010).

The functional and genetic biodiversity experiments can be combined to examine potential enrichment of microbiomes based on exposure to different environments. For example, after incubation in Ecoplates™, the microbial samples are essentially enriched based on the sole carbon source available. The genetic approach of determining 16S ribosomal RNA gene sequences can then be used to identify changes that may have taken places in such enrichments (Ros 2008).

Additional Scientific Practices

One major drawback of this CURE is that students carry out a relatively straightforward series of experiments. Because of the large enrollment and time constraints, we are unable to provide students the opportunity to develop their own research projects. A potential solution, currently in a pilot phase, is to challenge students to develop research proposals. In our case, these research proposals are written collaboratively in teams and are presented as posters in the style of a research conference. Each team identifies a topic to study hypothetically and proposes experiments to investigate that topic using the methods learned in the course. While not ideal, these research proposals still allow students to develop their own ideas and engage with primary literature without the need for additional laboratory resources. We also invite faculty, postdoctoral researchers, and graduate students from across campus to serve as judges for the poster session, providing students with more authentic interactions with scientists.

Cost of Equipment and Reagents

The cost of some equipment, such as the NanoDrop Lite spectrophotometer, can be a potential barrier to implementing this CURE. The procedures can be modified depending on available resources. It is not necessary to determine DNA concentration or visualize the samples on agarose gels at every step. The pGEM®-T Easy vector systems, which is substantial in cost, can be replaced by more economic systems such as the CloneJET PCR cloning kit or traditional cloning with restriction enzymes. PCR primers can be modified to include restriction sites to allow for traditional cloning. The Mix-and-Go Transformation kit can be replaced by traditional transformation if time is less of a constraint. To reduce costs, plasmids can be isolated in the laboratory and sent for sequencing instead of using direct colony sequencing services provided by the vendor.

Class and Preparation Time

Many of the experiments are optimized for time (Table 4), so that the entire research project can be completed by introductory students within one three-hour laboratory session and one 80-minute lecture per week. After testing a number of soil genomic DNA extraction kits, we found that the ZR Soil Microbe DNA MiniPrep™ protocol can be completed by students in 60-90 minutes with consistent results, as assayed by DNA concentration and intactness of genomic DNA. The PCR protocol is shortened to 25 cycles of amplification with minimal denaturing, annealing, and elongation times in each cycle. With a run-time of under 60 minutes, the protocol allows students to run the reactions and observe the products on an agarose gel in one the laboratory session. The cloning protocol with pGEM®-T Easy vector allows for ligation of PCR products directly from the reaction without additional treatment with restriction enzymes. For transformation, we use the Mix-and-Go *E. coli* Transformation kit, which requires no incubation or heat-shocking and recovery steps. Transformed colonies are sent for direct colony sequencing to eliminate the extra time required for plasmid purification.

Table 4. Estimated time for students to perform each experiment.

Experiment	Estimated student time
Soil collection	1 hour to visit the Natural Reserve System
Soil moisture	15 minutes on the first day, 15 minutes on the second day
Soil pH	1 hour including a 30-minutes incubation period
Ecoplate™ setup	1 hour to suspend and dilute soil microbes and set up Ecoplate™
Ecoplate™ analysis	2 hours to analyze data on Microsoft Excel®, 1 hour to generate figures
DNA extraction	1 hour to extract DNA and determine DNA concentration
Gel electrophoresis	30 minutes to pour gel and load samples, 30 minutes to run gel
Polymerase chain reaction	45 minutes to set up reactions, 60 minutes to run reactions
PCR clean-up	30 minutes to purify products and determine DNA concentration
Ligation	30 minutes to set up reactions, 30 minutes for incubation
Transformation	30 minutes to set up reactions and transfer directly onto plates
DNA sequencing	30 minutes to restreak plates, DNA sequencing by vendor
Sequence analysis	2 hours to analyze data, 1 hour to generate figures

The experiments are designed to minimize preparation time for laboratory staff. Most of the functional and genetic biodiversity experiments use commercially available kits, and the soil properties experiments only require simple equipment. Two full laboratory sessions are dedicated to data analysis on computers which requires virtually no laboratory preparations.

Laboratory Safety

Even though this course is designed with minimal use of hazardous materials, a substantial portion of the first week is dedicated to laboratory safety, including the use of personal protective equipment and an orientation to general waste disposal. Instructions for specific hazardous waste disposal are provided to students each week by teaching assistants and laboratory staff at the beginning of the laboratory session. For example, the soil samples contain microbial organisms, and especially after incubation in

Ecoplates™, certain organisms could be enriched. Therefore, Ecoplates™ are discarded as biohazardous waste. For the genetic biodiversity experiments, the PCR purification involves small quantities of 70% ethanol, and the wash solutions with ethanol are discarded in a dedicated container. Agarose gel electrophoresis is performed with SYBR Safe DNA gel stain, which is less hazardous than the traditional ethidium bromide. Finally, as general good practice, students are instructed to wash their hands with soap and clean the bench with 70% ethanol at the beginning and the end of each laboratory session.

Acknowledgements

We thank Deborah Wong, Lorlina Almazan, and other laboratory staff who have made this course-based undergraduate experience possible. We are grateful to all the students and instructional assistants who have participated in this course. Since its conception, they have provided important feedback to the operations of the course and contributed data to the on-going soil microbiome research project. Finally, we thank Dr. Lisa McDonnell and Dr. Melissa Micou for their continuous discussions that help improve the course for thousands of students each year.

About the Authors

Stanley M. Lo is an Assistant Teaching Professor in the Section of Cell and Developmental Biology and the Program in Mathematics and Science Education at University of California San Diego. His research aims to understand how faculty teaching promotes student learning and what factors support sustainable classroom innovations. Dr. Lo received his B.S. and M.S. in Molecular Biophysics and Biochemistry from Yale University and Ph.D. in Biochemistry from Harvard University. Dr. Lo was a Senior Research Associate in Learning Sciences at the Searle Center for Advancing Learning and Teaching at Northwestern University, prior to moving to San Diego. Between undergraduate and graduate studies, Dr. Lo taught biology and chemistry at Phillips Exeter Academy, where he first developed his interests in science education. Since 2011, Dr. Lo has been a National Academies Education Fellow and Mentor in the Life Sciences.

Stephanie F. Mel is an Associate Teaching Professor in the Section of Molecular Biology at the University of California San Diego. She received her B.S. in Genetics from University of California Davis and her Ph.D. from the University of California San Francisco, where she worked in the Macromolecular Structure Group. After graduate school, Dr. Mel did a post-doctoral fellowship at Harvard Medical School studying cholera. She joined the teaching faculty at University of California San Diego in 1998 and has taught a wide variety of both laboratory and lecture courses in fields including structural biochemistry, molecular biology, physiology, and genomics. One of Dr. Mel's main interests has been to introduce research into the undergraduate curriculum.

Cited References

Boyer Commission on Educating Undergraduates in the Research University. 1998. Reinventing Undergraduate Education: A Blueprint for America's Research Universities.

Brownell SE, Kloser MJ. 2015. Toward a conceptual framework for measuring the effectiveness of course-based undergraduate research experiences in undergraduate biology. Studies in Higher Education 40 :525-544.

Butler M, Henter H, Mel S. 2014. From bugs to barcodes: Using molecular tools to study biodiversity. Tested Studies for Laboratory Teaching 35: 41-55.

Caspers ML, Roberts-Kirchhoff ES. 2003. An undergraduate biochemistry laboratory course with an emphasis on a research experience. Biochemistry and Molecular Biology Education 31:303-307.

Chi MTH, Wylie R. 2014. The ICAP framework: Linking cognitive engagement to active learning outcomes. Educational Psychologist 49: 219-243.

Choi KH, Dobbs FC. 1999. Comparison of two kinds of Biolog microplates (GN and ECO) in their ability to distinguish among aquatic microbial communities. Journal of Microbiological Methods 36: 203-213.

Ericsson KA. 1993. The role of deliberate practice in the acquisition of expert performance. Psychological Review 100: 363-406.

Esteban DJ, Hysa B, Bartow-McKenney C. 2015. Temporal and spatial distribution of the microbial community of Winogradsky columns. PLoS ONE 10: e0134588.

Fierer N, Jackson RB. 2006. The diversity and biogeography of soil bacterial communities. Proceedings of the National Academy of Sciences 103: 626-631.

Freeman S. 2011. Biological Science. Benjamin Cummings.

Garland JL. 1997. Analysis and interpretation of community-level physiological profiles in microbial ecology. FEMS Microbiology Ecology 24: 289-300.

Gibbons SM, Jones E, Bearquiver A, Blackwolf F, Roundstone W, Scott N, Hooker J, Madsen R, Coleman ML, Gilbert JA. 2014. Human and environmental impacts on river sediment microbial communities. PLoS ONE 9: e97435.

Gick ML, Holyoak KJ. 1983. Schema induction and analogical transfer. Cognitive Psychology 15: 1-38.

Goode S, Magill RA. 1986. Contextual interference effects in learning three badminton serves. Research Quarterly for Exercise and Sport 57: 308-314.

Hanauer DI, Jacobs-Sera D, Pedulla ML, Cresawn SG, Hendrix RW, Hatfull GF. 2006. Teaching scientific inquiry. Science 314: 1880-1881.

Inderjit, van der Putten WH. 2010. Impacts of soil microbial communities on exotic plant invasions. Trends in Ecology and Evolution 25: 512-519.

Kaplan D, Maymon M, Agapakis CM, Lee A, Wang A, Prigge BA, Volkogon M, Hirsch AM. 2013. A survey of the microbial community in the rhizosphere of two dominant shrubs of the Negev Desert highlands, *Zygophyllum dumosum* (*Zygophyllaceae*) and *Atriplex halimus* (*Amaranthaceae*), using cultivation-dependent and cultivation-independent methods. American Journal of Botany 100: 1713-1725.

Kloser MJ, Brownell SE, Chiariello NR, Fukami T. 2011. Integrating teaching and research in undergraduate biology laboratory education. PLoS Biology 9: e1001174.

Kloser MJ, Brownell SE, Shavelson RJ, Fukami T. 2013. Effects of a research-based ecology lab course: A study of nonvolunteer achievement, self-confidence, and perception of lab course purpose. Journal of College Science Teaching 42:72-81.

Kornell N, Bjork RA. 2008. Learning concepts and categories: Is spacing the "enemy of induction"? Psychological Science 19: 585-592.

Krzywinski M, Altman N. 2013. Points of significance: Significance, P values and t-tests. Nature Methods 10: 1041-1042.

Krzywinski M, Altman N. 2014. Points of significance: Nonparametric tests. Nature Methods 11: 467-468.

Lave J, Wenger E. 1991. Situated Learning: Legitimate Peripheral Participation. Cambridge University Press.

Marshall PA, Sweat KG. 2008. Integrating mathematics into microbiology: Statistical analysis of Biolog community level phenotype profiling. Journal of the Arizona-Nevada Academy of Science 40: 157-159.

Moran MA. 2015. The global ocean microbiome. Science 350: 1330, aac85445.

Mulcahy MP, Edenborn HM. 2007. Environmental microbial ecology. Tested Studies for Laboratory Teaching 28: 28-51.

Murthy PPN, Thompson M, Hungwe K. 2014. Development of a semester-long, inquiry-based laboratory course in upper-level biochemistry and molecular biology." Journal of Chemical Education 91: 1909-1917.

Myers N, Mittermeier RA, Mittermeier CG, Fonseca GAB, Kent J. 2000. Biodiversity hotspots for conservation priorities. Nature 403: 853-858.

President's Council of Advisors on Science and Technology (PCAST). 2012. Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics.

Preston-Mafham J, Boddy L, Randerson PF. Analysis of microbial community functional diversity using sole-carbon-source utilization profiles: A critique. FEMS Microbiology Ecology (2002) 42: 1-14.

Pruesse E, Peplies J, Glockner FO. 2012. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. Bioinformatics 28: 18-23-1829.

Pruvis A, Hector A. Getting the measure of biodiversity. Nature (2000) 405: 212-219.

Reece JB, Urry LA, Cain ML, Wasserman SA, Minorsky PV, Jackson RB. 2014. Campbell Biology. Pearson.

Riesenfeld CS, Schloss PD, Handelsman J. 2004. Metagenomics: Genomic analysis of microbial communities. Annual Review of Genetics 38: 525-552.

Ros M, Goberna M, Pascual JA, Klammer S, Insam H. 2008. 16S rDNA analysis reveals low microbial diversity in community level physiological profile assays. Journal of Microbiological Methods 72: 221-226.

Rout ME, Southworth D. 2013. The root microbiome influences scales from molecules to ecosystems: The unseen majority. American Journal of Botany 100: 1689-1691.

Rundell EA, Banta LM, Ward DV, Watts CD, Birren B, Esteban DJ. 2014. 16S rRNA gene survey of microbial communities in Winogradsky columns. PLoS ONE 9: e104134.

Rutherford FJ, Ahlgren A. 1991. Science for All Americans. Oxford University Press.

Skwor T. 2012. The use of DNASTAR Lasergene educational software with molecular techniques to support bacterial identification. Tested Studies for Laboratory Teaching 33: 327-334.

Taylor EV, Fortune JA, Drennan CL. 2010. A research-inspired laboratory sequence investigating acquired drug resistance. Biochemistry and Molecular Biology Education 38: 247-252.

Tringe SG, Rubin EM. 2005. Metagenomics: DNA sequencing of environmental samples. Nature Review Genetics 6: 805-814.

Turner S., Pryer KM, Miao VPW, Palmer JD. 1999. Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. Journal of Eukaryotic Microbiology 46: 327-338.

Weber KP, Grove JA, Gehder M, Anderson WA, Legge RL. 2007. Data transformations in the analysis of community-level substrate utilization data from microplates. Journal of Microbiological Methods 69: 461-469.

Weber KP, Legge RL. 2009. One-dimensional metric for tracking bacterial community divergence using sole carbon source utilization patterns. Journal of Microbiological Methods 79: 55-61.

Weber KP, Legge RL. 2010. Community-level physiological profiling. Methods in Molecular Biology: Bioremediation 559: 263-281.

Weisburg WG, Barns SM, Pelletier DA, Lane DJ. 1991. 16S ribosomal DNA amplification for phylogenetic study. Journal of Bacteriology 173: 697-703.

Wenger Etienne. 1998. Communities of Practice: Learning, Meaning, and Identity. Cambridge University Press.

Whitley E, Ball J. 2002. Statistics review 3: Hypothesis testing and P values. Critical Care 6: 222-225.

Whitley E, Ball J. 2002. Statistics review 6: Nonparametric methods. Critical Care 6: 509-513.

Wolfe BE, Button JE, Santarelli M, Dutton R. 2014. Cheese rind communities provide tractable systems for in situ and in vitro studies of microbial diversity. Cell 158: 422-433.

Wolfe BE, Dutton RJ. 2015. Fermented foods as experimentally tractable microbial ecosystems. Cell 161: 49-55.

Wolfe BE, Klironomos. 2005. Breaking new ground: Soil communities and exotic plant invasion. BioScience 55: 477-487.

Woose CR, Fox GE. 1991. Phylogenetic structure of the prokaryotic domain: The primary kingdoms. Proceedings of the National Academics of Sciences 74: 5088-5090.

Zak JC, Willig MR, Moorhead DL, Wildman HG. 1994. Functional diversity of microbial communities: A quantitative approach. Soil Biology and Biochemistry 26: 1101-1108.

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

Lo, S. M. and Mel, S. F. 2017. Examining Microbial Biodiversity in Soil: A Large-Enrollment Introductory Course-Based Undergraduate Research Experience. Article 6 In: McMahon K, editor. Tested studies for laboratory teaching. Volume 38. Proceedings of the 38th Conference of the Association for Biology Laboratory Education (ABLE). **http://www.ableweb.org/volumes/vol-38/?art=6**