JOURNAL OF MICROBIOLOGY & BIOLOGY EDUCATION, December 2016, p. 423-429
DOI: http://dx.doi.org/10.1128/jmbe.v17i3.1115

Curriculum

Metagenomic Approaches to Identify Novel Organisms from the Soil Environment in a Classroom Setting †

Sadia J. Rahman[1], Trevor C. Charles[2], and Parjit Kaur[1]*
[1]*Department of Biology, Georgia State University, Atlanta, GA, 30303, USA,*
[2]*Department of Biology, University of Waterloo, Waterloo, ON N2V 2P1, Canada*

Molecular Microbial Metagenomics is a research-based undergraduate course developed at Georgia State University. This semester-long course provides hands-on research experience in the area of microbial diversity and introduces molecular approaches to study diversity. Students are part of an ongoing research project that uses metagenomic approaches to isolate clones containing 16S ribosomal ribonucleic acid (rRNA) genes from a soil metagenomic library. These approaches not only provide a measure of microbial diversity in the sample but may also allow discovery of novel organisms. Metagenomic approaches differ from the traditional culturing methods in that they use molecular analysis of community deoxyribonucleic acid (DNA) instead of culturing individual organisms. Groups of students select a batch of 100 clones from a metagenomic library. Using universal primers to amplify 16S rRNA genes from the pool of DNA isolated from 100 clones, and a stepwise process of elimination, each group isolates individual clones containing 16S rRNA genes within their batch of 100 clones. The amplified 16S rRNA genes are sequenced and analyzed using bioinformatics tools to determine whether the rRNA gene belongs to a novel organism. This course provides avenues for active learning and enhances students' conceptual understanding of microbial diversity. Average scores on six assessment methods used during field testing indicated that success in achieving different learning objectives varied between 84% and 95%, with 65% of the students demonstrating complete grasp of the project based on the end-of-project lab report. The authentic research experience obtained in this course is also expected to result in more undergraduates choosing research-based graduate programs or careers.

INTRODUCTION

Most universities in the US like to provide at least one signature experience to their undergraduates during their four-year degree program. For biology majors, the most significant signature experience can come from hands-on involvement in a research project, either by becoming part of a research team or working with individual research faculty. Since it is often not practical for all biology majors to experience research in faculty members' laboratories, an alternate approach is to develop research-based laboratory courses. Most traditional undergraduate laboratory courses consist of a series of independent experiments with known outcomes designed to provide students with experience of the basic laboratory techniques. Such courses, however, do not foster active learning nor do they encourage new discoveries. On the other hand, the goal of a research-based course is to engage students in inquiry-based experiments. We have developed a new research-based undergraduate laboratory course (titled Molecular Microbial Metagenomics, or M^3) for biology students at Georgia State University (GSU). The central idea of the course is based on understanding and analyzing microbial diversity. This course encourages active learning through hands-on laboratory experiments, data collection, analysis, evaluation and synthesis, as well as class discussion and writing exercises, thus providing an authentic research experience which is beneficial to undergraduates for developing analytical and critical thinking skills (2, 7, 16). In addition, this semester-long course allows students to develop laboratory skills in the areas of microbiology, molecular biology, and bioinformatics, with the additional possibility of discovering novel species of bacteria and archaea.

Microorganisms occupy every niche on our planet; yet due to our limited ability to cultivate them, fewer than 1% of the organisms have been identified (19, 21). In recent years, however, metagenomic approaches have provided valuable tools to tap into the large diversity found in different ecological niches (6, 14). Metagenomics refers to culture-independent analysis of a community of organisms (9, 10) and involves characterization of community deoxyribonucleic acid (DNA) isolated from an environmental niche (22). The

*Corresponding author. Mailing address: Department of Biology, Georgia State University, 58 Edgewood Ave., Atlanta, GA 30303. Phone: 404-413-5405. Fax: 404-413-5301. E-mail: pkaur@gsu.edu.
†Supplemental materials available at http://asmscience.org/jmbe

community genomic DNA is normally captured in cloning vectors, resulting in large metagenomic libraries, which can then be screened either for specific nucleotide sequences or by functional analysis (18). One sequence-based screening approach (used in this course) consists of using a small subunit ribosomal ribonucleic acid gene (16S rRNA gene) as a phylogenetic anchor (1). The 16S rRNA gene (or 18S in eukaryotes) is useful as an "anchor" because this molecule is found in all living organisms and has a highly conserved structure and function (1). Additionally, the 16S rRNA molecule contains variable regions that differ among species, genera, or domains of organisms (1, 9). Thus, diversity of 16S rRNA genes in community DNA can be studied by using universal, domain, or species-specific nucleotide probes (10), which serves as an indicator of the diversity of organisms in a particular niche. Identification of a novel (previously unidentified) 16S rRNA gene in the DNA pool can lead to identification of previously unreported lineages of organisms. This approach is very powerful, as evidenced by a similar study which used 16S rRNA gene-based approaches and showed the presence of the archaea-like light-absorbing rhodopsin gene in marine bacteria (3). Contrary to the previous belief that rhodopsin-dependent energy conservation is unique to Haloarchaea, the discovery of the bacteriorhodopsin gene in γ-Proteobacteria suggested that it likely occurs commonly in microorganisms found in ocean waters (3).

In the Molecular Metagenomics course, students work with a large soil metagenomic library containing about 80,000 DNA clones. The library was constructed by isolating community DNA from the top 10 cm of agricultural soil sample from a corn field (13). Recent phylogenetic analyses suggest that soil may contain bacteria spanning at least 13 different phyla as well as clades of archaea previously not known to exist in soil (4, 9). Most of this diversity in soil remains uncultured; therefore metagenomic analysis could reveal a variety of novel species belonging to many different phyla of bacteria, including Actinobacteria, Proteobacteria, Cyanobacteria, and Firmicutes, as well as archaea of methanogenic, halophilic, and thermophilic groups. Thus, the screening of clones in our soil library provides opportunities for open-ended investigations, as in a real research project.

Metagenomics-related laboratory modules are now being increasingly implemented in research-based and inquiry-based undergraduate instructional labs (5, 8, 12, 15, 17). An introductory biology course developed by Gibbens et al. (8) consists of four 100-minute modules that include both sequence-based and function-based analyses of the metagenome prepared from environmental samples. While this course provides an excellent overview of the potential applications of metagenomics, our semester-long M^3 course provides an in-depth experience focused on sequence-based analysis of the soil metagenome. Using a stepwise process to isolate 16S rRNA gene-containing clones from the library, our course imprints the analytical process, increases confidence and understanding of the research process, and promotes critical thinking. To our knowledge, this is the first exhaustive project-based undergraduate course using metagenomics to study microbial diversity. We expect that the skills gained in this course combined with the possibility of discovery will spark a lasting interest among undergraduates for scientific research, as discussed in a commentary by Weaver et al. (20).

Intended audience and prerequisite student knowledge

This course is currently being offered at GSU as a four-credit-hour, 4000-level, theme-based biology laboratory course. It was designed for upper-division undergraduate students majoring in biology, and prerequisites include 2000/3000-level introductory biology courses. Students should have a fundamental understanding of the scientific method, aseptic technique, and preparation of solutions as well as some experience in using basic laboratory equipment, such as micropipettes and microcentrifuges. These requirements could be waived if sufficient background information is provided by the instructor and more time is spent on explanation and demonstration of the basic laboratory techniques.

Learning time and learning objectives

The molecular metagenomics course offered at GSU is a 14-week course (one full semester) that meets for 2.5 hours twice a week. The major goal is to identify, isolate, and analyze clones containing a 16S rRNA gene from a soil metagenomic library. The learning objectives for this course are as follows:

1. Define and demonstrate understanding of the concept of microbial diversity and molecular phylogeny
2. Compare and contrast traditional culture-dependent methods and the culture-independent metagenomic approaches
3. Gain knowledge and demonstrate ability to perform basic molecular biology techniques
4. Record and interpret observations
5. Use bioinformatics tools to determine phylogenetic relationships
6. Develop analytical and critical thinking skills through synthesis of information and communication of findings
7. Design primers and follow-up experiments
8. Critically analyze scientific papers and engage in discussion

PROCEDURE

Materials and student instructions

A detailed list of all required materials for the course, sequences of primers, recipes for solutions, and protocols for each week are provided in the laboratory manual (Appendix

1) previously developed by us and printed through University Readers (11). The metagenomic library may be obtained from Dr. Trevor Charles or Dr. Parjit Kaur upon request. Further information about this library can be obtained from the Canadian MetaMicroBiome Library (13).

Faculty instructions

Project description. A detailed outline of the project and the steps of the procedure are provided in the lab manual (Appendix 1) and in Figure 1. Briefly, students are given a small aliquot of *Escherichia coli* cells containing the metagenomics library (13), which they dilute and plate on nutrient agar plates containing tetracycline (Fig. 1A). Each group of students then selects 100 colonies to work with, and by a process of elimination they will isolate one or more clones containing 16S rRNA gene sequences. The amplified 16S rRNA gene fragments are sequenced, followed by phylogenetic analysis.

Lab preparation. Detailed faculty instructions (timeline of each activity and a week-by-week guide for lab preparation) are provided in Appendix 2.1A. Topics covered in each class and time spent by students is provided in Appendix 2.1B. Initially, instructors should demonstrate dilution of the library and spreading of cells on agar plates, preparation of master plates, inoculation and growing cells in liquid media, isolation of cosmid DNA, and polymerase chain reaction (PCR) amplification. As the project progresses, students will have ample opportunities to master these methods.

Outcomes and issues for discussion with students. As in real research, the results for this project are not

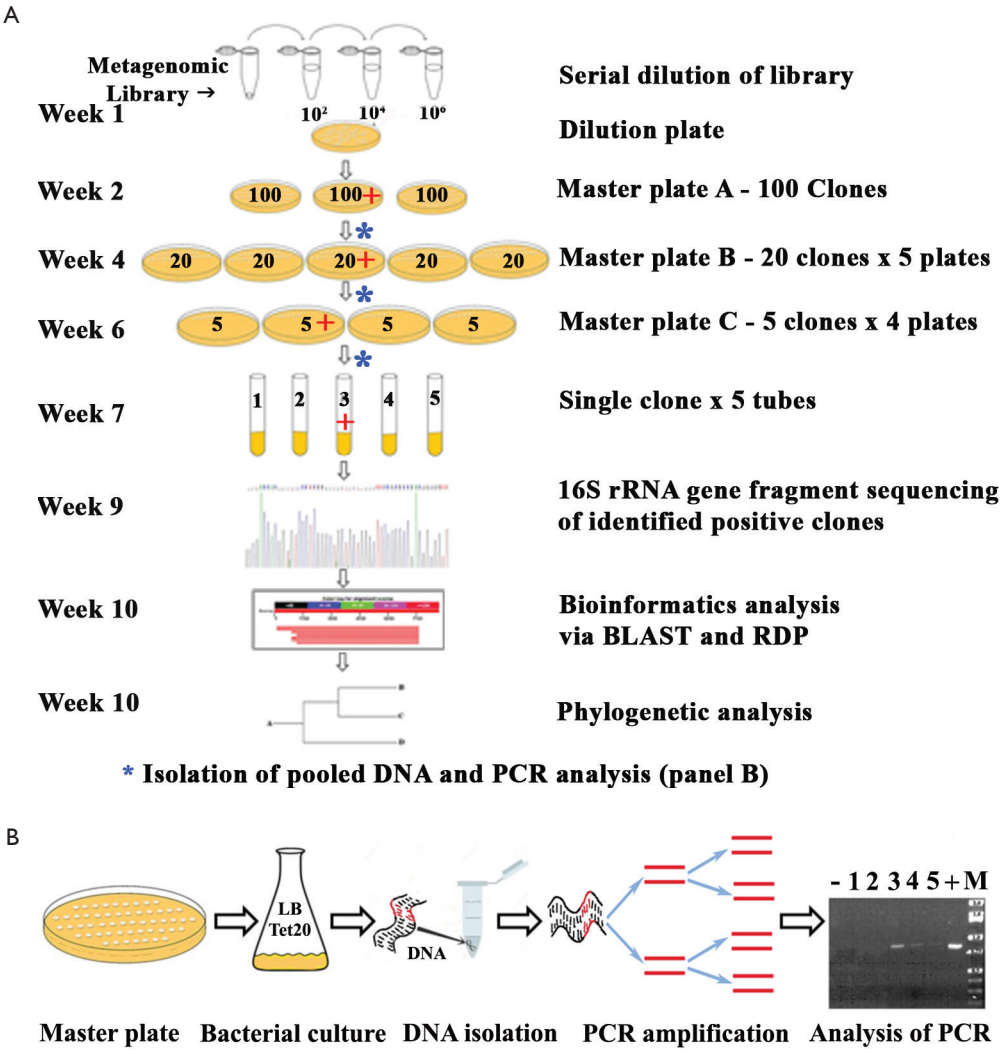

FIGURE 1. Flow of experiments in the project and analyses of PCR-amplified DNA. (A) Progression of experiments. A simplified flowchart of the experiments for identifying 16S rRNA gene-containing clones from the metagenomic library by a stepwise process of elimination. Red plus sign indicates the presence of a positive 16S-containing clone in the batch. (B) Isolation of pooled DNA and PCR analysis. This flowchart shows the experimental steps corresponding to each blue asterisk in panel A. DNA = deoxyribonucleic acid; rRNA = ribosomal ribonucleic acid; PCR = polymerase chain reaction; RDP = ribosomal database project.

guaranteed. Thus, it is possible that the batch of 100 clones selected by a group of students may not contain a 16S rRNA gene, which could prevent the group from moving forward with their planned experiments. Instructors should discuss this possibility with the students in advance and notify them that a positive clone will be provided to group(s) not able to find a 16S rRNA gene-containing clone in the selected batch. A positive clone is any previously identified metagenomic clone carrying the 16S rRNA gene. *E. coli* containing this plasmid should be provided to students during the division of 100 clones into batches of 20 (week 4) by replacing one of their clones but not telling them which clone was replaced, so that students may have the experience of identifying the single clone (or clones) carrying the 16S rRNA gene. A positive clone can be provided by us upon request. It is recommended, however, that if one or more clones containing the 16S rRNA gene are discovered by any groups, the instructors should make a stock of the *E. coli* cells carrying this clone for future use.

Suggestions for determining student learning

Several methods of evaluation can be used to assess student learning during this course, including pre- and post-class surveys (Appendix 3.1), quizzes, regular lab notebook checks, class exercises/assignments, participation, discussion of scientific literature, and a final lab report (Table 1). Assessment rubrics for the lab report, lab notebook checks, exercises/assignments, paper discussion, and participation are available in Appendix 2.3; a guide for preparation of the lab report and lab notebook are in Appendix 3.2 and 3.3; and samples of students' work, including a partial lab notebook, design exercise, and lab reports, are available in Appendix 4.

Sample data

The data obtained from the Undergraduate Summer Research Pilot (see Field testing) are shown under supplementary data in Appendix 5. Individual clones containing 16S rRNA gene were isolated from analysis of 500 clones in five batches in a stepwise manner, as described under faculty instructions. Of the five batches, three (batches 3, 4, and 5) resulted in the expected 825-bp 16S rRNA gene PCR product (Appendix 5, Fig. S1, panel A). Batch 5 was used for further analysis and was reduced from 100 colonies to 10 groups of 10 colonies each. PCR analysis resulted in the expected fragment in group 61 to 70 (Appendix 5, Fig. S1, panel B). Individual clones from group 61 to 70 provided three candidates which resulted in the expected 825-bp fragment: clone 61, 69, and 70 (Appendix 5, Fig. S1, panel C).

Using BLAST, all three metagenomic 16S rRNA sequences were assigned to domain Archaea with 93 to 99% identity to uncultured archaea of the phylum Crenarchaeota. Ribosomal database project (RDP) Classifier confirmed this prediction with high confidence of prediction (55–100%) (Appendix 5, Table S1). This analysis also predicted, albeit with low confidence, that clones 61, 69, and 70 belonged to the genuses *Thermocladium, Fervidicoccus,* and *Caldisphaera*, respectively. The low confidence of prediction suggests that these sequences may represent either novel species or even novel genera of archaea. Moreover, similar sequences found in the databases were all from uncultured archaea, which demonstrated to the students the importance of culture-independent methods of studying the microbial world. To further study the molecular relationship between these sequences and the known 16S sequences of the thermophilic archaea in the National Center for Biotechnology Information (NCBI) database, a preliminary phylogenetic tree was developed which suggests that the three sequences may form a distinct branch (Appendix 5, Fig. S2).

Safety issues

Students will work with a nonpathogenic *E. coli* strain classified as a biosafety level 1 organism. Please refer to the American Society for Microbiology (ASM) Guidelines for Biosafety in Teaching Labs for safe handling of microorganisms in the teaching lab. Students are required to maintain aseptic technique and wear lab coats, goggles, and gloves at all times during experimentations. Instructors should follow university guidelines for fire safety in labs when students are working with Bunsen burners. All chemicals in this course, including tetracycline and boric acid, are low risk as biohazardous agents, but should be discarded according to the university biohazard waste disposal guidelines when necessary.

DISCUSSION

Field testing

After the development of protocols, proof-of-concept was first tested in an eight-week Summer Pilot Program in 2012 with five students chosen from diverse backgrounds. Based on the success of the pilot, the semester-long course was then offered during the spring semesters in 2014 with nine students and in 2015 with seven students.

Evidence of student learning

Assessment measures for the Summer Pilot Program included pre/post-class surveys completed by the five students chosen as participants. The survey results (Fig. 2) revealed that even though students had some familiarity with the concept of microbial diversity and molecular techniques before the pilot, their knowledge was limited. Most students also had a limited knowledge of metagenomics or its application to the study of microbial diversity. At the end of the program, however, every student showed full understanding of metagenomics and how these approaches differ from traditional culturing methods. The median score

TABLE 1.
Learning objectives and the corresponding methods of assessment for the semester-long course.

Learning Objectives	Assessment Methods	Average Scores	Materials Provided
1. Define and demonstrate understanding of microbial diversity and molecular phylogeny	**Quiz** (Students were graded individually)	**84.5**	Appendix 4.1 – Examples of quiz questions and answers
2. Compare and contrast traditional and metagenomics approaches			
3. Gain knowledge and demonstrate ability to perform molecular biology techniques	**Participation / Attendance** (Students were graded individually)	**95.5**	Appendix 2.3 – Grading rubric for attendance and participation
4. Record and interpret observations	**Lab Notebook Checks** (Students were graded individually)	**89**	Appendix 2.3 – Grading rubric for lab notebook Appendix 3.3 – Guide for lab notebook Appendix 4.5 – Partial sample of student's lab notebook
5. Use bioinformatics tools to determine phylogenetic relationships	**Lab Report** (Students were graded individually)	**89.5**	Appendix 2.3 – Grading rubric for lab report Appendix 3.2 – Guide for preparing lab report Appendix 4.3 and 4.4 – Student lab report I and II
6. Develop analytical and critical thinking skills through synthesis and communication			
7. Design primers and follow-up experiments	**Exercises / Class Assignments** (Students were graded individually or as part of the group)	**90**	Appendix 2.3 – Grading rubric for exercises/assignments Appendix 4.2 – Experimental design exercise and sample of student work
8. Critically analyze and engage in discussion of scientific literature	**Literature Paper Discussion** (Students were graded as part of the group)	**92**	Appendix 2.3 – Grading rubric for class discussions Appendix 2.2 – Suggested reading material

on the pre-class survey was 25%, which increased to 100% on the post-class survey. A sample of student responses on the surveys is provided in Appendix 3.1.

Various methods used to assess learning gains in the two semester-long courses offered in 2014 and 2015 are shown in Table 1. The data in Figure 3A and Table 1 show that average scores for different learning objectives using six assessment methods ranged from 84% to 95%, with an overall average grade of B+ in the two courses. Further breakdown of the assessment scores for the quizzes (Fig. 3B) and lab reports (Fig. 3C) was also carried out. Analysis of selected quiz questions showed that the average score varied between 67% and 97%, with the lowest average score seen on Q5 (purpose of restriction digestion in protocol). Although the students generally knew what restriction enzymes do (also seen in survey results, Fig. 2), they did not fully grasp the purpose of restriction digestion of their cosmid DNA, which needs better explanation. The end-of-project lab report was graded on the content of each section of the report and the overall style. The average score using the provided assessment rubrics varied between 74% and 97%, with the lowest average score seen in the Results section

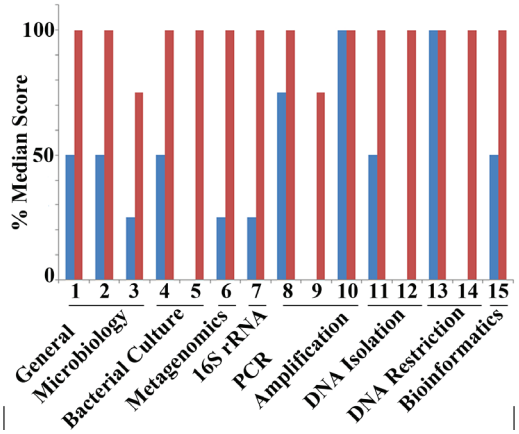

FIGURE 2. Learning gains evident from pre/post-class surveys given during the Summer Pilot Program. A pre-class survey (Appendix 3) was given on the first day of the program to test knowledge of the topics related to the course. The same survey was provided to the students at the end of the program. The surveys were completed by five students and were evaluated on accuracy as well as the extent of knowledge displayed. Blue = pre-class survey; red = post-class survey; RNA = ribonucleic acid; PCR = polymerase chain reaction; DNA = deoxyribonucleic acid.

(Fig. 3C), indicating that students had difficulty communicating rationale and summarizing results; these will need more emphasis in the future. Two additional criteria were established to assess student learning as reflected in the lab reports: understanding the flow of the project, and the ability to describe the concepts of metagenomics, microbial diversity, and phylogeny in detail. The data in Figure 3D show that 90% of students earned at least 1 out of 2 on the rubric, with 65% meeting both criteria. Samples of feedback received from students on a post-class questionnaire as well as additional outcomes and impacts on student success are provided in Appendix 3.1B.

Possible modifications

At GSU, Molecular Microbial Metagenomics is a 14-week, semester-long course. However, this course can be adapted as a module for microbiology labs or as part of a mini-semester. To complete the course over fewer weeks, the project can be concluded after the bioinformatics

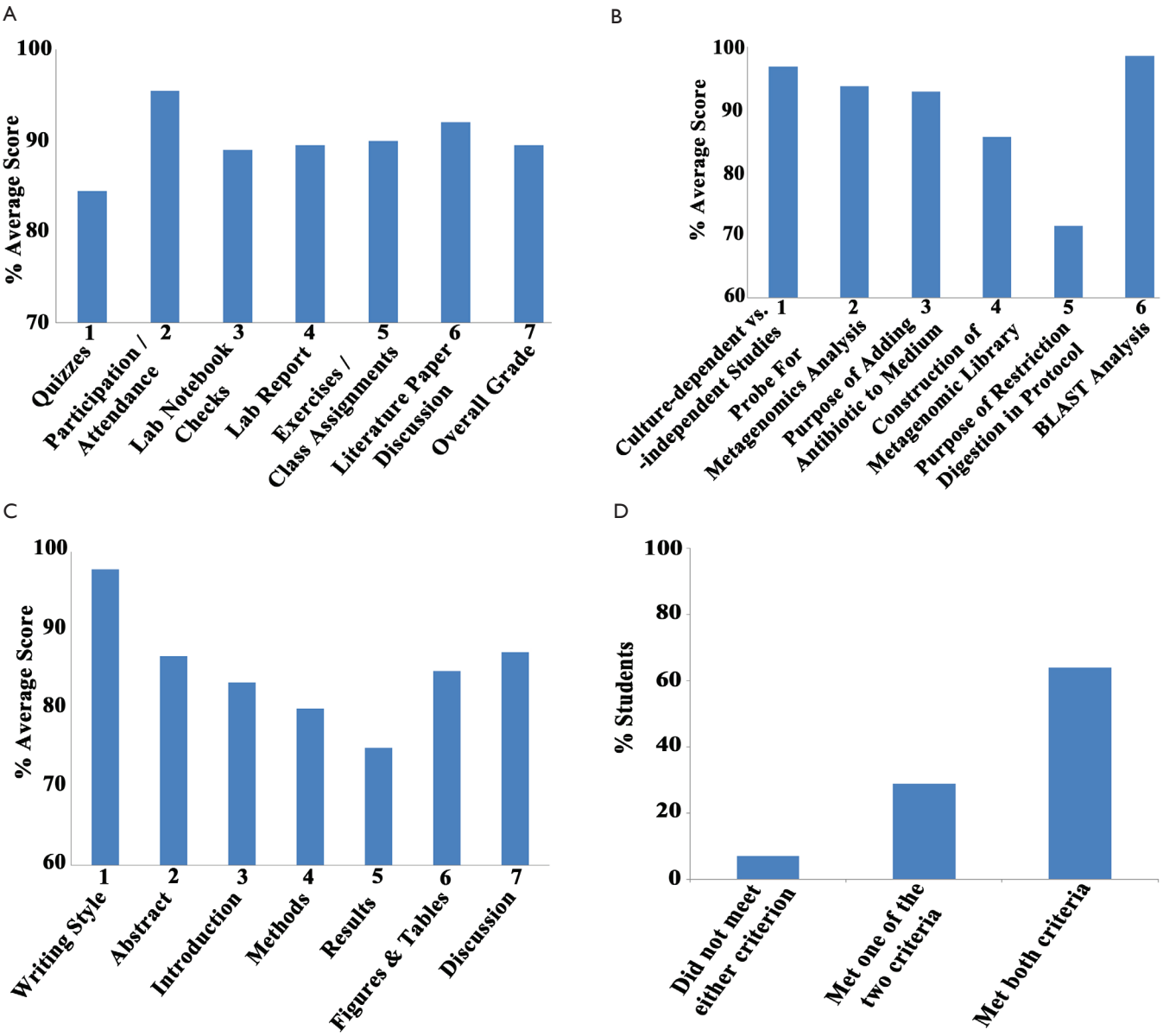

FIGURE 3. Assessment of learning gains from two semester-long courses with a total of 16 students. (A) Average scores on all assessment methods (1–6) and the average overall course grade over two semesters (7). (B) Average scores on selected quiz questions (provided in Appendix 4.1). (C) Detailed assessment of student lab reports: writing style follows scientific paper (1); abstract summarizes purpose of the project (2); introduction includes appropriate content (3); method section provides concise narrative (4); results include rationale and summary of experiments (5); figures and tables are included in results (6); discussion includes analysis of results (7). (D) Overall assessment of the lab reports using a two-point rubric: i) understanding flow of the project, and ii) understanding central concepts. BLAST = basic local alignment search tool.

analysis of the 16S rRNA gene sequence in week 10 and the collection of lab reports. Additionally, depending on the time and pace of the classroom, some experiments may be combined and media/reagents may be provided by instructors.

Though this course was initially designed for upper-division biology majors, instructors can accommodate lower-division students as well as other science majors by supplying sufficient background information on microbial diversity and demonstrating basic techniques. This course may also be advantageous for postgraduate students of biology who are seeking to gain hands-on research experience in the area of molecular microbiology and are interested in exploring microbial diversity.

SUPPLEMENTAL MATERIALS

Appendix 1: Laboratory manual
Appendix 2: Faculty instructions, reading materials, and assessment rubrics
Appendix 3: Pre/post-class surveys and student responses, a guide for lab report, and a guide for lab notebook
Appendix 4: Samples of students' work
Appendix 5: Supplementary data

ACKNOWLEDGMENTS

Thanks are due to the Georgia Perimeter College and the Bio-Bus program (http://biobus.gsu.edu/) of Georgia State University for financial support of the Summer Pilot and to the Molecular Basis of Disease and the McNair programs of Georgia State University for stipends to the participants. Thanks are also due to Georgia State University for award of the Instructional Innovation grant to develop this course. The authors declare that there are no conflicts of interest.

REFERENCES

1. **Amann, R. I., W. Ludwig, and K. H. Schleifer.** 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol. Rev. **59**(1):143–169.
2. **American Association for the Advancement of Science.** 2011. Vision and change in undergraduate biology education: a call to action: a summary of recommendations made at a national conference organized by the American Association for the Advancement of Science, July 15–17, 2009. Washington, DC.
3. **Beja, O., et al.** 2000. Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. Science **289**(5486):1902–1906.
4. **Bintrim, S. B., T. J. Donohue, J. Handelsman, G. P. Roberts, and R. M. Goodman.** 1997. Molecular phylogeny of Archaea from soil. Proc. Natl. Acad. Sci. USA **94**(1):277–282.
5. **Donato, J. J., M. A. Kllimstra, J. R. Byrnes, R. J. White, and T. C. Marsh.** 2012. The introduction of metagenomics into an undergraduate biochemistry laboratory course yielded a predicted reductase that decreases triclosan susceptibility in *Escherichia coli*. DNA Cell Bio. **31**(6):968–973.
6. **Fuhrman, J. A.** 2012. Metagenomics and its connection to microbial community organization. F1000 Biol. Rep. **4**:15.
7. **Gasper, B. J., and S. M. Gardner.** 2013. Engaging students in authentic microbiology research in an introductory biology laboratory course is correlated with gains in student understanding of the nature of authentic research and critical thinking. J. Microbiol. Biol. Educ. **14**(1):25–34.
8. **Gibbens, B. B., C. L. Scott, C. D. Hoff, and J. L. Schottel.** 2015. Exploring metagenomics in the laboratory of an introductory biology course. J. Microbiol. Biol. Educ. **16**(1):34–40.
9. **Handelsman, J.** 2004. Metagenomics: application of genomics to uncultured microorganisms. Microbiol. Mol. Biol. Rev. **68**(4):669–685.
10. **Hugenholtz, P., and N. R. Pace.** 1996. Identifying microbial diversity in the natural environment: a molecular phylogenetic approach. Trends Biotechnol. **14**(6):190–197.
11. **Kaur, P., and S. J. Rahman.** 2015. Molecular microbial metagenomics. BIOL 4905 laboratory manual. University Readers, Georgia State University, Atlanta, GA.
12. **Muth, T. R., and C. M. McEntee.** 2014. Undergraduate urban metagenomics research module. J. Microbiol. Biol. Educ. **15**(1):38–40.
13. **Neufeld, J. D., et al.** 2011. Open resource metagenomics: a model for sharing metagenomic libraries. Stand. Genomic Sci. **5**(2):203–210.
14. **Pace, N. R.** 1997. A molecular view of microbial diversity and the biosphere. Science **276**(5313):734–740.
15. **Rios-Velazquez, C., et al.** 2011. Summer workshop in metagenomics: one week plus eight students equals gigabases of cloned DNA. J. Microbiol. Biol. Educ. **12**(2):120–126.
16. **Russell, S. H., M. P. Hancock, and J. McCullough.** 2007. Benefits of undergraduate research experience. Sci. Educ. Forum **316**:548–549.
17. **Sanders, E. R., and A. M. Hirsch.** 2014. Immersing undergraduate students into research on the metagenomics of the plant rhizosphere: a pedagogical strategy to engage civic-mindedness and retain undergraduates in STEM. Plant Sci. **5**(157):1–4.
18. **Sharpton, T. J.** 2014. An introduction to the analysis of shotgun metagenomic data. Front. Plant Sci. **5**:209.
19. **Ward, B. B.** 2002. How many species of prokaryotes are there? Proc. Natl. Acad. Sci. USA **99**(16):10234–10236.
20. **Weaver, G. C., C. B. Russell, and D. J. Wink.** 2008. Inquiry-based and research-based laboratory pedagogies in undergraduate science. Nat. Chem. Biol. **4**(10):577–580.
21. **Whitman, W. B., D. C. Coleman, and W. J. Wiebe.** 1998. Prokaryotes: the unseen majority. Proc. Natl. Acad. Sci. USA **95**(12):6578–6583.
22. **Zarraonaindia, I., D. P. Smith, and J. A. Gilbert.** 2013. Beyond the genome: community-level analysis of the microbial world. Biol. Philos. **28**(2):261–282.