

COMMENTARY – Professional Development

Personal microbiomes and next-generation sequencing for laboratory-based education

Mark R. Hartman^{1,†}, Kristin T. Harrington¹, Candice M. Etson¹,
Matthew B. Fierman¹, Donna K. Slonim² and David R. Walt^{1,*}

¹Department of Chemistry, Tufts University, Medford, MA, 02155, USA and ²Department of Computer Science, Tufts University, Medford, MA, 02155, USA

*Corresponding author: Department of Chemistry, 62 Talbot Avenue, Medford, MA 02155, USA. Tel: 1 (617) 627-2013; E-mail: david.walt@tufts.edu

One sentence summary: This commentary discusses approaches and challenges for bringing personal microbiome and next-generation sequencing experiments into the classroom laboratory.

Editor: Beatrix Fahnert

[†]Mark R. Hartman, <http://orcid.org/0000-0003-1564-5285>

ABSTRACT

Sequencing and bioinformatics technologies have advanced rapidly in recent years, driven largely by developments in next-generation sequencing (NGS) technology. Given the increasing importance of these advances, there is a growing need to incorporate concepts and practices relating to NGS into undergraduate and high school science curricula. We believe that direct access to sequencing and bioinformatics will improve the ability of students to understand the information obtained through these increasingly ubiquitous research tools. In this commentary, we discuss approaches and challenges for bringing NGS into the classroom based on our experiences in developing and running a microbiome project in high school and undergraduate courses. We describe strategies for maximizing student engagement through establishing personal relevance and utilizing an inquiry-based structure. Additionally, we address the practical issues of incorporating cutting edge technologies into an established curriculum. Looking forward, we anticipate that NGS educational experiments will become more commonplace as sequencing costs continue to decrease and the workflow becomes more user friendly.

Keywords: microbiome; next-generation sequencing; education

INTRODUCTION

In the past decade, next-generation sequencing (NGS) technologies have transformed sequencing from a method practiced only by experts in dedicated facilities to a general purpose research tool available to researchers in a wide range of fields. The maturation of NGS has driven an information boom in biology, concomitant with the development of new bioinformatics tools. As NGS and bioinformatics become increasingly widespread, these technologies are playing a growing role in fields such as human health and microbiology (Shendure and Aiden 2012). Thus, there is a need for students to gain familiarity with NGS and bioinformatics as part of modern biological science educa-

tion. To address this need, novel NGS and bioinformatics educational experiments have recently been developed for undergraduate classrooms and laboratories (Drew and Triplett 2008; Makarevitch and Martinez-Vaz 2016).

Microbiology has greatly benefited from advances in NGS and bioinformatics that allow for the rapid study of entire microbial communities. The study of microbiomes is naturally engaging to students because the topic is novel, has far-ranging implications and affects them on a personal level. In addition to these merits, the topic of microbiomes represents an excellent vehicle for bringing NGS into the classroom (Wang et al. 2015). In this commentary, we discuss our experience in developing and

implementing a 'Microbiome Portrait' laboratory module, which we have piloted through our NIH SEPA-funded program 'Bioinformatics Inquiry through Sequencing' (BioSeq 2016). Our microbiome laboratory module has allowed us to introduce NGS and bioinformatics to many high school and undergraduate classrooms. We believe that our experience may be informative to others considering the implementation of similar experiments.

An overview of our Microbiome Portrait Experiment is provided in Table 1. In this laboratory-based experiment, students investigate their own and their classmates' microbiomes to address the research question: 'Is the makeup of the microbiome influenced more strongly by individual or by body site?' To find the answer, students collect samples of microbes from two sites on their bodies and prepare amplicons from the microbial 16S ribosomal RNA genes for sequencing (Klindworth et al. 2013). Using results (relative abundance of identified microbes) from the Illumina MiSeq metagenomic analysis workflow, students employ simple bioinformatics concepts (such as determining the Jaccard index) to quantify the similarity between samples and propose an answer to the research question. We run this experiment in a 'structured inquiry' format, where the students are given a research question but draw their own conclusions from their data (while microbes exhibit body site preferences, each person's microbiome is unique, so the outcome of the experiment is unknown in advance). Our major learning objectives are for students to understand that microbiomes and their host environments can influence one another and to understand how these communities can be sampled and compared. Student knowledge and attitudes are assessed pre- and post-implementation to evaluate the effectiveness of our approach and to identify areas that need improvement. Below, we discuss the benefits of this topic and the experimental structure while also addressing the difficulties of incorporating NGS and bioinformatics into the classroom.

ENGAGING STUDENTS THROUGH PERSONAL MICROBIOMES

Topic selection and personal samples

Microbiomes provide an ideal topic for introducing NGS and bioinformatics techniques and technologies. This topic is

easily accessible to students, and the ubiquity of microbiomes allows the project's focus to be suited to a range of classes or for students to select the samples they find the most interesting. We chose to focus our project on *personal* microbiomes to maximize students' connection to their sample and discuss recent research tying personal microbiomes to individual health.

In our experience, the efficacy of this module is increased by allowing students to work with their own personal microbiome samples. However, using personal samples for educational purposes also raises a host of practical and ethical issues (Walt et al. 2011; Callier 2012; Sanderson et al. 2013; LaBonte and Beers 2015; Garber, Hyland and Dasgupta 2016). Before undertaking any lessons or experiments involving human-derived samples, it is extremely important to communicate with the appropriate institutional review board (IRB), which has jurisdiction over human subjects research and can provide guidance in prioritizing student privacy and addressing health implications. We obtained IRB approval for both the use of student microbiome samples and the students' participation in educational research. We note that while human microbiome data are often not considered to be specific to a person, recent publications have demonstrated that microbiomes may be used to uniquely identify individuals (Franzosa et al. 2015; Meadow et al. 2015).

In our implementation, we obtain student consent prior to their participation, and students who wish to opt out of the study can do so at any time. Students are assured that their decision regarding consent will not affect their grade, and their teacher will not have access to their educational assessment test or survey results. Students who decline to collect their own samples are provided third party samples as alternatives. We support student privacy throughout the experiment by anonymizing all sample identities and data with randomized codes. After sequencing results are obtained, students receive their own data back for analysis of their microbiome and, optionally, comparison with the microbiomes of their classmates. Throughout the experiment, we emphasize several health disclaimers, including (1) the mere detection of potentially pathogenic bacteria does not signify a health risk, and (2) data obtained in an educational setting are not equivalent to a diagnostic result and cannot be considered health information. We also explore the limitations of interpreting data by discussing how our computational sequence classification methods lack the

Table 1. Summary of the Microbiome Portrait Experiment as implemented by BioSeq.

Overview of the microbiome experimental module	
Outline of the experiment	Specific activities performed by students
<p>Background. Introductory lessons give context and motivate the experiment. Students use a research question that is provided to them.</p>	<ul style="list-style-type: none"> • Survey current literature on microbiomes, including popular science articles • Discuss research questions and hypothesis
<p>Sample preparation. The laboratory workflow engages students through hands-on experiences. Students can perform all aspects of the procedure until the final library preparation and loading the sequencer. Students see these final steps via a laboratory visit or video.</p>	<ul style="list-style-type: none"> • Collect samples and extract DNA • Perform PCR for targeted amplification of 16S region • Clean-up and check PCR product • Use indexing PCR to attach adapters and indexes • Perform second PCR clean-up and quantification
<p>Data analysis. Data are preprocessed for student analysis. Complexity of the analysis depends on the depth of bioinformatics desired for the course.</p>	<ul style="list-style-type: none"> • Use automated primary/secondary analysis tools • Analyze similarities among samples • Propose answers to the original research questions • Present results to classmates

resolution to distinguish pathogenic from harmless bacterial strains, and we stress that this fact helps to explain disclaimer (1) above.

Benefits of inquiry-based design

The study of microbiomes lends itself well to inquiry-based projects. Inquiry-based learning uses open-ended questions or scenarios to provide students with a glimpse of genuine research, while also providing enough direction and structure to focus their learning (Hofstein and Lunetta 2004; Blanchard et al. 2010; Puttick, Drayton and Cohen 2015). Inquiry-based microbiology experiments have been reported for a wide range of topics selected to enhance student engagement, including beer brewing (Sato et al. 2015), oral hygiene (Geiger-Thornsberry 2016) and endophytes from rainforest plants (Strobel and Strobel 2007). Inquiry projects provide ample opportunities for hands-on experiences and self-direction, while allowing the level of support to be tailored to the needs of a particular class. This project could also be run as an ‘open inquiry’ experiment by having students pose their own questions and sample the microbiomes necessary to answer that question. For example, a student may ask, ‘Can the microbiome on my phone be identified as mine?’ (Lax et al. 2015). Because microbiomes are a new area of research, it is possible for students to ask many questions that lack well-defined answers. Nonetheless, the processing workflow for microbiome samples is well established (Illumina 2014), allowing students to focus on designing questions and analyzing data without pioneering new techniques.

COST CONSIDERATIONS FOR NGS TECHNOLOGY

The greatest impediment to educational use of NGS remains cost. A typical sequencing run for a microbiome experiment with an Illumina MiSeq System will handle 100–200 samples. In our experience, 96 samples represent a reasonable upper limit based on typical classroom size and convenience of processing in 96-well plates. On this basis, the total cost of our typical experiment was approximately \$4000 (roughly \$3000 for consumables, \$500 for equipment operational costs and \$500 for technician time). This total excludes the capital expense of equipping an NGS lab, which despite the continued decrease in equipment costs may be prohibitive for teaching labs (the Illumina MiSeq currently costs over \$100 000). However, as NGS has grown in popularity, so have opportunities for accessing this technology. We encourage educators interested in incorporating NGS into their teaching to collaborate with colleagues who have access to the technology. It should be noted that the per sample cost is expected to come down as sequencing costs decrease and as more samples can be run per lane.

CONCLUSION

Laboratory experiences involving personal microbiomes represent an excellent opportunity to make NGS and bioinformatics part of modern science education. The relatively high cost of sequencing has traditionally discouraged its use in the classroom, but it is now increasingly feasible to obtain access to NGS technology and design experiments with a relatively low per sample cost. In this era of high-throughput technologies, there are many similar examples of high capital cost, centralized instrumentation that is typically inaccessible to students. We believe our

model can be extended to many of these high-cost instruments (e.g. flow cytometers, mass spectrometers, computing clusters) for use in inquiry-based educational experiences.

FUNDING

The authors acknowledge support from the National Institutes of Health [award number R25OD010547-01].

Conflict of interest. David R. Walt is the scientific founder and holds equity in Illumina Inc.

REFERENCES

- BioSeq. *Bioinformatics Inquiry through Sequencing*. 2016. <http://ase.tufts.edu/chemistry/walt/sepa/index.html> (23 November 2016, date last accessed).
- Blanchard MR, Southerland SA, Osborne JW et al. Is inquiry possible in light of accountability?: a quantitative comparison of the relative effectiveness of guided inquiry and verification laboratory instruction. *Sci Educ* 2010;**94**:577–616.
- Callier SL. Swabbing students: should universities be allowed to facilitate educational DNA testing? *Am J Bioethics* 2012;**12**:32–40.
- Drew JC, Triplett EW. Whole genome sequencing in the undergraduate classroom: outcomes and lessons from a pilot course. *J Microbiol Biol Educ* 2008;**9**:3–11.
- Franzosa EA, Huang K, Meadow JF et al. Identifying personal microbiomes using metagenomic codes. *P Natl Acad Sci USA* 2015;**112**:E2930–8.
- Garber KB, Hyland KM, Dasgupta S. Participatory genomic testing as an educational experience. *Trends Genet* 2016;**32**:317–20.
- Geiger-Thornsberry GL. How does oral hygiene affect bacterial counts in saliva? *Am Biol Teach* 2016;**78**:258–60.
- Hofstein A, Lunetta VN. The laboratory in science education: foundations for the twenty-first century. *Sci Educ* 2004;**88**:28–54.
- Illumina. *16S Metagenomic Sequencing Library Preparation*. 2014. http://support.illumina.com/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf (23 November 2016, date last accessed).
- Klindworth A, Pruesse E, Schweer T et al. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* 2013;**41**:e1.
- LaBonte ML, Beers MA. An alternative laboratory designed to address ethical concerns associated with traditional TAS2R38 student genotyping. *Biochem Mol Biol Educ* 2015;**43**:100–9.
- Lax S, Hampton-Marcell JT, Gibbons SM et al. Forensic analysis of the microbiome of phones and shoes. *Microbiome* 2015;**3**:21.
- Makarevitch I, Martinez-Vaz B. Killing two birds with one stone: model plant systems as a tool to teach the fundamental concepts of gene expression while analyzing biological data. *Biochim Biophys Acta* 2016, DOI: 10.1016/j.bbaggm.2016.04.012.
- Meadow JF, Altrichter AE, Bateman AC et al. Humans differ in their personal microbial cloud. *PeerJ* 2015;**3**:e1258.
- Puttick G, Drayton B, Cohen E. A study of the literature on lab-based instruction in biology. *Am Biol Teach* 2015;**77**:12–8.
- Sanderson SC, Linderman MD, Kasarskis A et al. Informed decision-making among students analyzing their personal genomes on a whole genome sequencing course: a longitudinal cohort study. *Genome Med* 2013;**5**:113.

- Sato BK, Alam U, Dacanay SJ et al. Brewing for students: an inquiry-based microbiology lab. *J Microbiol Biol Educ* 2015;**16**:223–9.
- Shendure J, Aiden EL. The expanding scope of DNA sequencing. *Nat Biotechnol* 2012;**30**:1084–94.
- Strobel SA, Strobel GA. Plant endophytes as a platform for discovery-based undergraduate science education. *Nat Chem Biol* 2007;**3**:356–9.
- Walt DR, Kuhlik A, Epstein SK et al. Lessons learned from the introduction of personalized genotyping into a medical school curriculum. *Genet Med* 2011;**13**: 63–6.
- Wang JTH, Daly JN, Willner DL et al. Do you kiss your mother with that mouth? An authentic large-scale undergraduate research experience in mapping the human oral microbiome. *J Microbiol Biol Educ* 2015;**16**:50–60.