

Building Engineering Skills for the Genomics Revolution, a Genomics Technologies and Analysis Course for Biomedical Engineers

Dr. Karen R. Thickman, University of Washington

Karen R. Thickman is a lecturer in the Department of Bioengineering at the University of Washington. Previously, she was an assistant teaching professor at Carnegie Mellon University in the Computational Biology Department for five years. She received a Ph.D. in molecular biophysics from the Johns Hopkins University School of Medicine and an A.B. in biophysical chemistry from Dartmouth College. Thickman develops and teaches courses for an online professional masters program, courses in genomics and genomic technologies, and creates laboratory experiences. She also performs educational research and aimed at continually improving student learning and outcomes, and conducts research in online education to improve access to bioengineering education for students at various times in their careers.

Work in progress: Building Engineering Skills for the Genomics Revolution, a Genomics Technologies and Analysis Course for Biomedical Engineers

Over the past decade, development of next-generation sequencing technologies has led to dramatic changes in how genetic information is gathered and analyzed¹. These changes have led to substantial improvements in biomedical research and medical practices, including personalized cancer care and preventative procedures based on genetic backgrounds²⁻⁵. Biomedical and bioengineering students are entering a world where understanding how to obtain and analyze large-scale genetic data is a valuable skill. Bioengineers are uniquely positioned to address many of the large challenges in the genomics field, as these challenges require individuals with skills in biology, computation, and design. However, many degree programs do not include coursework in this area. Though there is literature describing genomics courses for first year undergraduates, nursing students and medical residents, none were found for engineering students⁶⁻⁹.

To address this need and provide our students with an opportunity to develop the skills necessary to advance genomics and its applications, I have developed a genomics technology and analysis course for advanced undergraduate students and graduate students: *Genomics Era Sequencing and Analysis*. This work-in-progress describes the design and implementation of the course, currently in its second offering. This course is innovative both in its content and structure. The course materials are available at <https://github.com/krthickman/Genomics-Technologies-and-Analysis-Course>, and may be used as a template for teaching bioengineers to enter and lead in genomic technology and analysis development. Pedagogical innovations of the course, such as low-stakes quizzes, are transferrable to any bioengineering course.

This course explores the technological advances that enabled the development of next-generation technologies, including the bioengineering constraints and criteria for developing these technologies. We focus on the currently available tools including NanoString, Illumina MiSeq and HiSeq, IonTorrent, Pacific Biosciences, and the Oxford NanoPore. With few computational and technical prerequisites, students acquire the skills

Table 1: Course Learning Objectives and Alignments with ABET Outcomes.

Learning Objectives		ABET outcomes
1	Identify and apply the genetic and biological terms that describe patterns of inheritance and genome contents.	A
2	Describe the types of libraries that are generated for DNA sequencing on genomic scales. Describe how to generate these libraries.	A
3	Distinguish between technologies being used to sequence DNA, and identify their strengths and weaknesses	A,B
4	Describe the history, development, and engineering challenges of these technologies. Explain the technological and engineering principles that underlie them	A
5	Determine what technology would be most useful for different biological and biomedical questions	A, B
6	Perform analysis of genetic and genomic data with and without software	A, K
7	Use common software to analyze different types of genomic data	A, K
8	Identify limitations on current technologies, and define the criteria of current engineering problems in the field of genomics	J
9	Discuss the opportunities and challenges posed by these technologies in the biomedical field	H
10	Understand the impact of high-throughput sequencing on the medical, computational, and bioengineering fields	H, J

to assess the pros and cons of each technology in isolation and in comparison to each other, to conduct quantitative analysis of large genomic data sets, and to consider technical advances that could facilitate the collection and analysis of new data to answer outstanding questions.

The course objectives for the course are listed in Table 1. These align with a number of the ABET objectives as indicated.

Student Profiles and Course Structure

This elective course is intended for advanced undergraduates and graduate students with backgrounds in bioengineering, though it is open to all advanced students in the university. In its offering in Spring 2016, 70% were bioengineering undergraduates and 30% bioengineering graduate students. In the current offering, Spring 2017, there are 50% each of undergraduate and masters students. Prerequisites for the course include an undergraduate biology course or an equivalent that provides the basics of genes and genomes. It is recommended that students have courses in biochemistry, genetics, or biological system analysis. It is not expected that students can write a program to analyze data, but they should be comfortable using a command line programs whether in Java, Python, MatLab or R.

The course provides an opportunity to gain analytical skills commonly used in this area of bioengineering by performing hands-on analysis of genomic data commonly used in this area of bioengineering. Statistical theories that underlie the analysis software are discussed and demonstrated with in-class activities. The course has two 80-minute lectures and a 3-hour computer lab section each week. The lab section is an addition from the first offering of the course, and is intended iteration to expand opportunities to analyze data. In these lab sessions, students use state-of-the-art analytical tools for genomic analysis that are used in current bioengineering research. These freely available tools run on Mac and Linux environments. These labs build familiarity with common genomics software and skills to analyze complex data. Students will be performing many basic genomic analyses, including de novo genome assembly and annotation, RNA-Seq alignment and analysis, and association studies with phenotypes. As we have limited access to high-performance computing facilities, we mostly work with smaller datasets. We discuss the computational needs for scaling data and students will complete analysis-based reports at the end of each analytical module. These reports are used to assess student mastery and achievement of the performative and analytical learning objectives.

Though this course is not flipped, I am using some of these techniques to provide low-stakes assessments to help students acquire metacognition of their own learning^{10,11}. In particular, students take weekly online reading quizzes before class through the course Canvas site. Students are able to retake these quizzes up to five times to enable them to identify material they need to review. The quizzes are graded to incentive students to complete them. These quizzes provide low-stakes practice opportunities demonstrated to improve student learning¹¹⁻¹³. In addition to providing feedback to the students about their learning, these quizzes identify areas that need extra emphasis in lecture^{14,15}. They are not intended to determine if students have met the learning objectives of the course, so are not used in evaluating the course.

Students perform a final project that focuses on synthesizing the material they learned in the class and exploring their interests in this area more fully. These projects include defining

engineering design criteria and constraints of current or proposed genetic sequencing technologies, or analysis of a chosen data set. Undergraduate students work in teams of 2-3 and graduate students work alone. Graduate students present their project to the class. These projects demonstrate the skills gained by the students during the course, and are used in evaluation of the conceptual learning objectives.

During the first offering, students completed a variety of final projects. Some students focused on technical descriptions of current DNA sequencing tools, including the Agena Mass Array system and GnuBio's single-cell sequencing platform. These evaluations demonstrated strong understandings of the technology underlying these tools and their strengths and limitations. A number of students performed analyses on publically available data sets, including a search for SNPs to help explain the high incidence of cancer in dogs, differences in metabolic processes in human microbiomes, and visualizations of genetic changes associated with glioblastomas. Many of these analyses were fairly basic, as expected given a single course on genomic analysis. These projects clearly demonstrated creativity, analytical thinking, and broad interests in biological and biomedical questions. Finally, one graduate student proposed a new means of sequencing mRNAs through monitoring ribosomal synthesis. Though this proposal had clear challenges, it was creative and the student struggled with some of the complexities of pushing forward this field.

Assessment

In addition to analytical lab reports and final projects, student exams will be used to evaluate student mastery of the learning goals. Preliminary analysis from the first course offering indicates most students met many of the learning goals. In particular, using the final exam from 2016, it appeared most undergraduate students gained mastery in four of the seven learning objectives tested on the exam (Table 2). Many students struggled with mastery of strategies for data analysis and the technological needs of sequencers. To better address these, I have added a computational lab, to enable more extensive data analysis, and have allocated more class time to discussing the technological needs and problems of the current sequencing tools. Several innovations are planned for the second offering in Spring quarter 2017, and analysis of these innovations will be presented. First, student progress will be evaluated systematically through the use of the midterm and final exams. These exams have been redesigned to better balance the learning objectives tested and to align each question more clearly with an objective. Second, student perceptions of interest in and importance of genomics technologies will be assessed through surveys administered at the beginning and end of the course. Third, as questions of ethics, consent, and privacy are central to appropriate use of these genomic technologies, in-class discussions and student reflections on these issues will be used to assess student understanding of the non-technical challenges associated with these tools.

Table 2: Mastery of Learning Objectives of 2016 Undergraduate Students Based on Performance on Final Exam

Learning Objective	Number of Questions	Average Fraction Correct
1	4	0.7
2	2	0.8
3	4	0.8
4	7	0.6
5	2	0.9
6	9	0.6
8	1	0.9

1. Mardis, E. R.; *Trends Genet.* **24**, 133–141 (2008).
2. Hood, L. & Flores, M.; *N. Biotechnol.* **29**, 613–624 (2012).

3. Tursz, T., *et al*; *Nat. Rev. Clin. Oncol.* **8**, 177–183 (2011).
4. Disease, Committee on A Framework for Developing a New Taxonomy of Disease, Board of Life Sciences, Division of Earth and Life Sciences, N. R. C. of E. N. A. *Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease. Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease* (2011). doi:10.17226/13284
5. Chivers, T. Genomics: the revolution that's transforming medicine. *The Telegraph*
6. Jordan, T. C. *et al.*; *mBio* **5**, (2014).
7. Daack-Hirsch, S. *et al.* Integrating Genetics and Genomics into Nursing Curricula. *Nurs. Clin. North Am.* **48**, 661–669 (2013).
8. Sharoff, L.; *J. Nurs. Educ. Pract.* **5**, 13–18 (2015).
9. Haspel, R. L., Ali, A. M. & Huang, G. C.; *J. Grad. Med. Educ.* 80–84 (2016).
10. Dunlosky, J., *et al*; *Psychol. Sci. Public Interes. Suppl.* **14**, 4–58 (2013).
11. McLaughlin, J.E., *et al*; *Acad. Med.* **89**, 236–243 (2014).
12. Zalewski, D. & Schneider, K. Assessing Learning Outcomes and Evaluating Graduate Student Perceptions of a Flipped Classroom. *ASEE Natl. Conf. Proceedings. New Orleans, LA* (2016).
13. Beidler, K. & Panton, L.; *J. Interact. Technol. Pedagog.* (2013).
14. Muzyka, J. L.; *Journal of Chemical Education* (2015). doi:10.1021/ed500904y
15. Geetracianne, K. L. & Neilsen, B.; *Proc. Mtgs. Acoust.* **18**, 025002 (2012); doi:10.1121/1.4770093