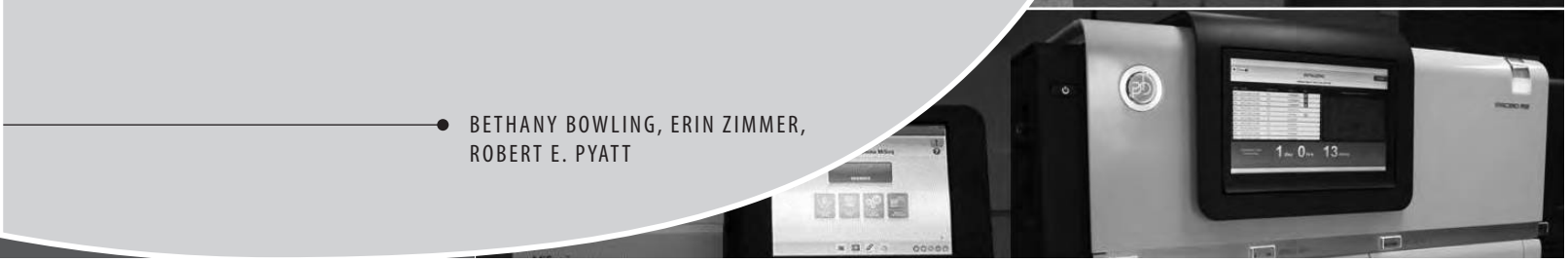
INQUIRY & INVESTIGATION

Bringing Next-Generation Sequencing into the Classroom through a Comparison of Molecular Biology Techniques



BETHANY BOWLING, ERIN ZIMMER, ROBERT E. PYATT

Abstract

Although the development of next-generation (NextGen) sequencing technologies has revolutionized genomic research and medicine, the incorporation of these topics into the classroom is challenging, given an implied high degree of technical complexity. We developed an easy-to-implement, interactive classroom activity investigating the similarities and differences between current sequencing methodology and three NextGen technologies. The activity uses existing materials created by each of the biotechnology companies that outline their instrumentation and chemistries. Following this activity, students will understand the molecular biology behind these NextGen applications and the similarities to existing Sanger sequencing methods.

Key Words: *Next-generation sequencing; molecular biology, Sanger sequencing.*

The current initiative in biology education is to transform the curriculum by developing student-centered teaching approaches that engage students in relevant interactive, inquiry-driven, cooperative, collaborative activities (AAAS, 2010). The challenge for science educators is to make the classroom current and dynamic without overwhelming students by going beyond their conceptual understanding. Research suggests that covering less material in greater detail is more effective than covering a host of topics within a course (Knight & Wood, 2005; Luckie et al., 2012). This is even more challenging in light of the overwhelming amount of new knowledge and advances in technology in the biological sciences and, in particular, the field of genetics. Therefore, undergraduate educators must find the appropriate balance between providing necessary depth and breadth of topics to cover.

Introducing students to the cutting-edge discoveries and technologies in a field can motivate and inspire undergraduates. In the field of genetics, the technology is rapidly evolving and changing. In the past decade, the demand for low-cost, high-throughput sequencing has driven the development of next-generation (NextGen) sequencing technologies. (The term "next-generation" is an industry standard that refers to technologies developed to replace Sanger sequencing and is not related, in this case, to science standards.) In these ultra-high-throughput sequencing methods, numerous operations can be run in parallel, which has led to the surge of whole-genome or whole-exome sequencing data we are now experiencing. NextGen sequencing is now a standard application for addressing questions in genetics, population dynamics, and anthropology. However, incorporation of NextGen technology in undergraduate curricula is likely far behind the pace of its use in research and clinical settings. Introducing students to NextGen technology is an important tool in that it helps them grasp the current advances in the field of genetics. This is particularly important as we train future medical professionals and researchers, as well as consumers of personalized genomics. The Genome Consortium for Active Teaching (GCAT; Buonaccorsi et al., 2011) also recognizes the importance of exposing undergraduate students to NextGen sequencing. One GCAT branch, entitled GCAT-SEEK(quence), gives undergraduate students access to NextGen data by establishing a regional research network through collaboration with a genome-sequencing core facility (Buonaccorsi et al., 2011), although this type of collaboration can only benefit institutions located near a genome-sequencing facility.

NextGen sequencing is now a standard application for addressing questions in genetics, population dynamics, and anthropology.

We have developed an easy-to-implement, interactive classroom ("dry lab") activity that is appropriate for undergraduate genetics, molecular biology, cell biology, or biochemistry courses. The activity focuses on investigating the similarities and differences between the current Sanger sequencing methods (Applied Biosystems BigDye Chain Terminator) and three NextGen technologies (Illumina, Life Technologies Ion Torrent, and Pacific Biosciences; Figure 1). It utilizes existing commercial materials created by each of these biotechnology companies that outline

The American Biology Teacher, Vol. 76, No. 6, pages 396–401. ISSN 0002-7685, electronic ISSN 1938-4211.
DOI: 10.1525/abt.2014.76.6.7

Figure 1. An Applied Biosystems 3730, Life Technologies Ion PGM, Illumina MiSeq, and Pacific Biosciences PacBio RS instruments. Image of the PacBio RS courtesy of http://gqinnovationcenter.com/services/sequencing/technoPacBio.aspx?l=e.

their instrumentation and chemistries for DNA sequencing. These product brochures and videos are aimed at a broad audience, and consequently this activity is general enough to be appropriate for high school AP Biology or undergraduate (community or four-year college) courses. The activity aligns with recommended action items from *Vision and Change in Undergraduate Biology Education: A Call to Action* such as "engaging students as active participants" and "facilitate learning within a cooperative context" (AAAS, 2010). It allows students to work directly with professionally developed, real-world materials used in DNA sequencing methodologies. In addition, many goals from the National Research Council's (NRC) *Framework for K–12 Science Education* are addressed by the activity, which could be modified for use at the high school level (Table 1; NRC, 2012).

We first presented this activity at the Undergraduate Faculty Education Workshop at the 62nd Annual Meeting of the American Society of Human Genetics (ASHG; http://www.ashg.org/education/). Each year, ~50 educators participate in this workshop, which is focused on undergraduate genetics education. The workshop is highlighted by a scientific presentation by a well-renowned geneticist on a cutting-edge topic. While the educators enjoy learning about recent research, one of the challenges experienced in the classroom is finding the time to incorporate current and authentic activities when there are so many basic concepts included in any one course. Also, while many would consider whole-exome sequencing and next-generation sequencing technologies to be advanced topics, they could be utilized as connectors that link several core topics included in a course and as reinforcement of basic genetics concepts. The ASHG workshop attendees were challenged to identify the basic concepts addressed or reinforced through inclusion of whole-exome sequencing and next-generation sequencing. After brainstorming within groups, a number of embedded basic genetics concepts were identified (Table 2). Feedback from the faculty attendees concerning using this activity in the classrooms was very positive.

○ Procedure

Part 1: Comparison of Commercially Available Machines

1. Students are divided into four groups, and each group is assigned one of the DNA sequencing instruments to be covered in the activity (Applied Biosystems 3730, Illumina MiSeq, Life Technologies Ion Torrent PGM/Proton, or Pacific Biosciences PacBio RS; Figure 1).

Table 1. Dimensions and practices from *Framework for K–12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (National Research Council, 2012) addressed by the NextGen sequencing activity.

Dimension	Practice	Goals
Scientific and Engineering Practices	8. Obtaining, evaluating, and communicating information	Read scientific and engineering text, including tables, diagrams, and graphs, commensurate with their scientific knowledge and explain the key ideas being communicated.
Crosscutting Concepts	1. Patterns 4. Systems and System Models	1. Observed patterns of forms and events guide organization and classification, and they prompt questions about relationships and the factors that influence them. 4. Defining the system under study – specifying its boundaries and making explicit a model of that system – provides tools for understanding and testing ideas that are applicable throughout science and engineering.
Disciplinary Core Ideas	Life Sciences Engineering, Technology, and Applications of Science	LS3.A: Inheritance of traits ETS2.A: Interdependence of science, engineering, and technology

Table 2. Embedded genetics topics in NextGen whole-genome and whole-exome sequencing.

Terms	Definition
Next-generation (NextGen) sequencing	The process of determining the exact order of nucleotides within a DNA molecular using technology developed after Sanger sequencing.
Exome	Sequencing only the parts of a genome that code for and are translated into protein.
Read length	The size of a single DNA strand of sequence produced through NextGen sequencing that is measured in the total number of bases.
Enzyme	A complex protein produced by plants and animals that is actively involved in the chemical reactions of life.
Dideoxynucleotide triphosphates (ddNTPs)	Compound consisting of a nucleoside combined with a phosphate group, but lacking a 3′ hydroxyl group, which, when incorporated into a newly synthesized strand of DNA, blocks the further addition of nucleotides to the DNA strand.

Table 3. Sequencing instrument name and web address of datasheet.

Instrument	Web Address
Applied Biosystems 3730	http://www3.appliedbiosystems.com/cms/groups/mcb_marketing/documents/generaldocuments/cms_041894.pdf
Illumina MiSeq	http://www.illumina.com/documents/products/datasheets/datasheet_miseq.pdf
Life Technologies Ion Torrent/Proton PGM	http://www3.appliedbiosystems.com/cms/groups/applied_markets_marketing/documents/generaldocuments/cms_096460.pdf
Pacific Biosciences PacBio RS	http://www.mscience.com.au/upload/pages/pacbio/pacbio-rs-brochure-2012.pdf

2. Each group is then given the product datasheet on their specific instrument. These documents are brief descriptions of the machines and their chemistries, created by each company to commercially promote them to potential buyers. Because they are written for a broad audience, they are appropriate for students who have at least an introductory background in molecular biology. The web addresses for each pdf document are presented in Table 3.
3. Students are given a blank copy of the worksheet "Comparison of sequencing machines and chemistries" (Table 4). Using the instrument datasheets, students work in groups to complete items 1 (Name the sequencing chemistry/technology), 2 (What is the read length?), and 3 (What is detected in the sequencing process?) of the worksheet for their assigned instrument.
4. Students can then regroup as a class or be shuffled into new groups of four with representatives of each technology and share their findings for each machine so that each student can complete the first half of the table.

Part 2: Comparison of Sequencing Chemistries

1. Having identified the specific sequencing chemistry for each platform in Part 1, students will view short movies created by each company to explain the molecular biology of DNA sequencing on each platform. The links to the animations for each chemistry are listed in Table 5.

Table 4. Comparison of sequencing machines and chemistries. Student worksheet for parts I and II of the activity.

	Applied Biosystems 3730	Illumina MiSeq	Life Technologies Ion PGM/Proton	Pacific Biosciences PacBio RS
1. Name the sequencing chemistry/technology.				
2. What is the read length?				
3. What is detected in the sequencing process?				
4. How is that detected?				
5. What enzyme is used in the sequencing process?				

Table 5. Sequencing chemistries and web addresses for illustrative animations.

Instrument	Web Address
BigDye Chain Terminator	http://media.invitrogen.com.edgesuite.net/ab/applications-technologies/pharma-biotherapeutics/DNA_sequencing.swf
Sequencing by Synthesis	http://www.youtube.com/watch?v=I99aKKHcxC4
Semiconductor Sequencing	http://www.youtube.com/watch?v=yVf2295JqUg
Single Molecule Real Time Sequencing (SMRT)	http://www.youtube.com/watch?v=NHCJ8PtYCFc

2. After viewing each video, students will answer worksheet items 3 (What is detected in the sequencing process?), 4 (How is that detected?), and 5 (What enzyme is used in the sequencing process?) in Table 4.

○ Results & Discussion

Part 1

The pdf documents for each instrument can easily be printed for students or accessed directly online. With some reading, students should be able to identify the name of the sequencing chemistry used on their instrument and the sequencing read length (the total length of a single sequence in nucleotides produced by the machine) (Table 6). Unlike the other instruments, the Applied Biosystems 3730 machine is not a sequencer itself but is used in the electrophoretic separation of fragments of DNA based on size. BigDye Chain Terminator sequencing chemistry (also known classically as Sanger sequencing) is done in a thermocycler before the products are loaded onto the 3730 instrument for base detection. By contrast, the other machines are themselves sequencers that house both the sequencing and base-detection steps. One of the main advantages of NextGen sequencing methods over traditional Sanger sequencing is the throughput. NextGen methods allow exponentially more sequencing output than Sanger. However, Sanger sequencing generates reads that are typically much larger than most, but not necessarily all, NextGen sequencing methods.

Additionally, students can explore online additional applications that are available for each of these instruments (SNP genotyping, linkage mapping, microsatellite analysis, methylation, loss of heterozygosity, and restriction fragment length polymorphism [RFLP] analysis).

Part 2

These animations distill each complex chemistry down to its basics, which students can then easily use to draw comparisons. Both BigDye Chain Terminator and Sequencing by Synthesis methods use dideoxynucleotide triphosphates (ddNTPs). However, the ddNTPs used in BigDye Chain Terminator are nonreversible and terminate the growing DNA strand by preventing the addition of other nucleotides, whereas the ones used in Sequencing by Synthesis are reversible. By contrast, both Semiconductor Sequencing and Single Molecule Real Time (SMRT) Sequencing use deoxynucleotides incorporated into the growing DNA strand. BigDye Chain Terminator, Sequencing

Table 6. Student worksheet for parts I and II of the activity with corresponding answers.

	Applied Biosystems 3730	Illumina MiSeq	Life Technologies Ion PGM/Proton	Pacific Biosciences PacBio RS
1. Name the sequencing chemistry/technology.	*BigDye Chain Terminator*	*Sequencing by Synthesis*	*Semiconductor Sequencing*	*Single Molecule Real Time Sequencing (SMRT)*
2. What is the read length?	*400–900 bp*	*36–250 bp*	*100 bp (35–400 bp)*	*Average 2200 to 3100 bp*
3. What is detected in the sequencing process?	*Irreversible ddNTP incorporated into DNA strand*	*Reversible ddNTP incorporated into DNA strand*	*dNTP incorporated into DNA strand*	*dNTP incorporated into DNA strand*
4. How is that detected?	*Fluorescent label of ddNTP*	*Fluorescent label of ddNTP*	*Change in pH (H+)*	*Fluorescent label of dNTP*
5. What enzyme is used in the sequencing process?	*DNA polymerase*	*DNA polymerase*	*DNA polymerase*	*DNA polymerase*

Table 7. Questions for short-term follow-up for the activity and additional activities that could be used immediately after this one.

Immediate Follow-up Questions	
What is consistent among the machines?	Why do you think there are some consistencies?
Why would machines be produced that use different technologies?	Why are NextGen sequencing applications such as whole-exome sequencing and whole-genome sequencing useful?
What ideas can you see in research and medicine for NextGen sequencing applications?	For a specific topic (forensic use of DNA, diagnosis of a genetic disease, identification of a gene responsible for a syndrome, etc.), have your students write a grant showing how NextGen sequencing could be beneficially used.
Extensions to Broader Concepts	
Alaie, A., Teller, V. & Qui, W.-G. (2012). A bioinformatics module for use in an introductory biology laboratory. *American Biology Teacher, 74,* 318–322.	
Taylor, D.L., Campbell, A.M. & Heyer, L.J. (2013). Illuminating the black box of genome sequence assembly: a free online tool to introduce students to bioinformatics. *American Biology Teacher, 75,* 572–577.	
Falteisek, L., Černý, J. & Janštová, V. (2013). A simplified technique for evaluating human *CCR5* genetic polymorphism. *American Biology Teacher, 75,* 704–707.	

by Synthesis, and SMRT sequencing all detect the addition of a nucleotide into a DNA strand using a fluorescent label on their respective ddNTPs or dNTPs. Semiconductor sequencing detects the change in pH caused by the release of hydrogen when a nucleotide is incorporated into the newly synthesized DNA strand. Finally, while all these methods differ in their operation or chemistry, they all utilize the enzyme DNA polymerase in the sequencing reaction (Table 6). Additional questions to stimulate further discussion are provided in Table 7, along with suggestions for classroom activities that would build on the principles in this exercise.

○ Conclusion

In summary, incorporating relevant and current trends in the classroom is a goal of all educators. The challenge for biology education is to keep up with the fast-pace, exciting, and novel breakthroughs, particularly in genomic technology and research. Too often, undergraduate biology educators take a traditional approach in teaching genetics, even though they recognize the importance of incorporating new scientific advances and producing a genetically literate public (McElhinny et al., 2014). Therefore, this engaging, inquiry-based activity will introduce NextGen technologies while at the same time reinforcing the core concepts in the biology curriculum. In addition, the use of industry-produced materials exposes students to real-world documentation and decision making. This is an important aspect of this classroom activity because it allows instructors to include basic genetics course material while introducing students to current techniques. This exposure will provide students the necessary background and preparation as they move forward in their biology career.

Web Resources

Applied Biosystems DNA Sequencing
https://products.appliedbiosystems.com/ab/en/US/adirect/ab?cmd=catNavigate2&catID=600522

Illumina MiSeq
http://www.illumina.com/systems/miseq.ilmn

Ion Torrent
http://products.invitrogen.com/ivgn/product/4462917

Pacific Biosciences
http://www.pacificbiosciences.com/products/

References

AAAS. (2010). *Vision and Change in Undergraduate Biology Education: A Call to Action.* Washington, DC: AAAS. http://visionandchange.org/.

Buonaccorsi, V.P., Boyle, M.D., Grove, D., Praul, C., Sakk, E., Stuart, A. et al. (2011). GCAT-SEEKquence: Genome Consortium for Active Teaching of undergraduates through increased faculty access to next-generation sequencing data. *CBE Life Sciences Education, 10,* 342–345.

Knight, J.K. & Wood, W.B. (2005). Teaching more by lecturing less. *CBE Life Sciences Education, 4,* 298–310.

Luckie, D.B., Aubry, J.R., Marengo, B.J., Rivkin, A.M., Foos, L.A. & Maleszewski, J.J. (2012). Less teaching, more learning: 10-yr study supports increasing student learning through less coverage and more inquiry. *Advances in Physiology Education, 36,* 325–335.

McElhinny, T.L., Dougherty, M.J., Bowling, B.V. & Libarkin, J.C. (2014). The status of genetics curriculum in higher education in the United States: goals and assessment. *Science & Education, 23,* 445–464.

National Research Council. (2012). *Framework for K–12 Science Education: Practices, Crosscutting Concepts, and Core Ideas.* Washington, DC: National Academies Press.

BETHANY BOWLING is an Associate Professor in the Department of Biological Sciences, Northern Kentucky University, Highland Heights, KY 41099. ERIN ZIMMER is an Associate Professor in the Department of Biology, Lewis University, Romeoville, IL 60446. ROBERT E. PYATT is Assistant Director, Cytogenetics/Molecular Genetics, at Nationwide Children's Hospital, 700 Children's Drive, Columbus, OH 43205; e-mail: robert.pyatt@nationwidechildrens.org.