

ARTICLE

A guided-inquiry investigation of genetic variants using Oxford nanopore sequencing for an undergraduate molecular biology laboratory course

Princess D. Rodriguez^{1,2} | Michael Mariani^{1,2} | Jamie Gay¹ | Tyler C. Hogan¹ | Eyal Amiel^{1,2} | Paula B. Deming^{1,2,3} | Seth Fretze^{1,2,3} 

¹Department of Biomedical and Health Sciences, University of Vermont, Burlington, Vermont

²Cellular Molecular Biomedical Sciences Program, University of Vermont, Burlington, Vermont

³The University of Vermont Cancer Center, Burlington, Vermont

Correspondence

Seth Fretze, University of Vermont, Department of Biomedical and Health Sciences, 106 Carrigan Drive, Burlington, VT 05405.

Email: seth.fretze@med.uvm.edu

Funding information

National Institutes of Health, Grant/Award Numbers: T32AI055402, U54GM115516; Office of Integrative Activities, Grant/Award Number: OIA-1736097

Abstract

Next Generation Sequencing (NGS) has become an important tool in the biological sciences and has a growing number of applications across medical fields. Currently, few undergraduate programs provide training in the design and implementation of NGS applications. Here, we describe an inquiry-based laboratory exercise for a college-level molecular biology laboratory course that uses real-time MinION deep sequencing and bioinformatics to investigate characteristic genetic variants found in cancer cell-lines. The overall goal for students was to identify non-small cell lung cancer (NSCLC) cell-lines based on their unique genomic profiles. The units described in this laboratory highlight core principles in multiplex PCR primer design, real-time deep sequencing, and bioinformatics analysis for genetic variants. We found that the MinION device is an appropriate, feasible tool that provides a comprehensive, hands-on NGS experience for undergraduates. Student evaluations demonstrated increased confidence in using molecular techniques and enhanced understanding of NGS concepts. Overall, this exercise provides a pedagogical tool for incorporating NGS approaches in the teaching laboratory as way of enhancing students' comprehension of genomic sequence analysis. Further, this NGS lab module can easily be added to a variety of lab-based courses to help undergraduate students learn current DNA sequencing methods with limited effort and cost.

KEYWORDS

cancer, next-generation sequencing, teaching laboratory, variants

1 | INTRODUCTION

The first complete draft of the human genome was finished in 2003, initiating a period of rapid advancement in our knowledge of the role of genomic variation in human

disease. Developments in next generation sequencing (NGS) technologies has laid the framework for expanding genetic research and has opened up an entire industry of new diagnostic applications. Notable new diagnostic tests are now being widely used in the clinical setting and are

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2021 The Authors. *Biochemistry and Molecular Biology Education* published by Wiley Periodicals LLC on behalf of International Union of Biochemistry and Molecular Biology

contributing to our understanding of diverse diseases. In particular, various NGS-based applications are aiding in the understanding of the pathogenesis and progression of cancer. Genomic sequence analysis can inform distinctive responses to drugs (pharmacogenomics) and the various mechanisms underlying different types of cancer. Genomic knowledge therefore permits more personalized health care as NGS-based analytical approaches are becoming increasingly relevant to all medical and biomedical professionals.^{1–3} Educating the next generation of physicians, pathologists, biomedical, and laboratory scientists will be instrumental in the workforce development necessary for the continued delivery and advancement of genomic medicine. With reduced costs in real-time sequencing, new opportunities exist for the application of NGS in the teaching laboratory. However, there remains a challenge in implementing cost-effective contemporary diagnostic tools in the undergraduate teaching laboratories.

The introduction of NGS to undergraduates interested in pursuing careers in the biomedical sciences presents an opportunity to integrate active 'wet-lab' experiences with bioinformatics analysis to promote the analytical skills required for the inquiry of data output from the sequencer. Strong evidence exists to suggest that immersing students in a hands-on and investigative learning experience leads to better student outcomes.⁴ While there are several examples of integrating genomic data, gene expression and Sanger sequencing data in the classroom, to date there are few guidelines available for laboratory instructors interested in bringing NGS technology into the classroom. There are also several considerations that must be accounted for prior to establishing an NGS-friendly module in the teaching laboratory including, (1) the time required to generate and sequence libraries, (2) analysis and interpretation of bioinformatically-heavy sequencing data, and (3) high-cost associated with running next generation sequencing machines. Finally, the 2019 coronavirus pandemic has introduced a new hurdle for laboratory instructors which includes executing a college-level molecular biology module off-site, to encourage distance-learning. For the purpose of introducing NGS concepts in molecular pathology, we created a 4-week laboratory module that highlights variant detection in cancer cells using an affordable, applicable NGS technique, which can be fulfilled with a mix of in person and remote/online experiences.

The MinION sequencing platform introduced by Oxford Nanopore Technologies in 2014, offers many advantages over second-generation sequencers making it an optimal teaching tool.⁵ The MinION operates by using a nanometer wide protein pore that allows single DNA or RNA molecules to pass through. As these molecules cross, the unique current measurement for each nucleotide is recorded and

translated into sequence information in real-time.^{6,7} The MinION device is portable, easily interfaces with a laptop, and has a startup investment estimated at ~\$1000.⁸ In addition, the MinION has been used for clinical applications such as summarizing the gut metagenome of head and neck cancer patients, detecting structural variants in pancreatic cancer cell-lines, and assessing the TP53 mutational status in chronic lymphocyte leukemia patients.^{9–11} The reasonable cost and ease of use, together with a wide range of applicability, make the MinION sequencing device an ideal platform for the teaching laboratory.

Here, we present a molecular laboratory course exercise that combines both wet-lab experimentation and variant interpretation after sequencing on the MinION platform. The overall goal for student groups was to correctly identify an unknown non-small cell lung cancer (NSCLC) cell-line based on its unique genomic profile. Students isolated and then PCR amplified genomic DNA from NSCLC cells using primers they created to amplify known variants in *KRAS*, *EGFR*, and *TP53*.^{12–16} Each forward and reverse primer was tagged at the 5' end with a unique 12-nucleotide molecular index and combined into a single reaction to develop a multiplex PCR genotyping assay that can be interpreted through bioinformatics analysis. These unique index tags or barcodes allowed for each group to combine their reaction products into a single sample that was then sequenced in real-time on the MinION device during the laboratory session. Following data processing by demultiplexing, the step where the barcode information is used to generate unique sequence files from each group after being sequenced together, and alignment of each file to a reference genome sequence, students were able to interpret variants via annotation using the freely available Integrative Genomics Browser software.¹⁷ We show that our students were able to accurately detect characteristic variants in three NSCLC cell-lines. Further, we provide pre-laboratory assignments (see Appendix S1) for each unit that can be easily modified by instructors as needed. Overall, this 4-week module can be implemented at a low-cost, with minimal bioinformatic background, using available equipment from a molecular biology laboratory.

2 | MATERIALS AND METHODS

2.1 | Class structure

The course, "Applied Molecular Biology Laboratory" was open to senior undergraduates or graduate students enrolled in the Medical Laboratory Science program at the University of Vermont. Twenty undergraduates and four graduate students were registered for the course. The course met once weekly for 3 h per session. For the

duration of the 4-week module, students worked in groups of three and were asked to take turns during hands-on experimentation.

2.2 | Cell culture

HCC827 (CRL-2868), H1975 (CRL-5908), and H441 (HTB-174) lung cancer cell-lines were purchased from the American Type Culture Collection (ATCC, Manassas, VA). Cell-lines were cultured in RPMI-1640 medium with 10% fetal bovine serum and 1% penicillin–streptomycin. All cells were cultured at 37°C and 5% CO₂. Cultured cells were counted and resuspended to a final concentration of one million cells per milliliter. One million cells were aliquoted into 1.5 ml tubes, washed once with PBS, and the cell pellets were stored in the –80°C. These steps were not performed by the students but were a part of the pre-laboratory preparation.

2.3 | Genomic DNA extraction and quantification for PCR

Genomic DNA was isolated by students from thawed cell pellets using the GeneJet Genomic DNA Purification Kit (Thermo Fisher Scientific, USA) as instructed by the manufacturer for adherent cells. All DNA pellets were eluted in 200 µl of distilled water. The concentration of the DNA was determined by students using a NanoDrop-2000 (Thermo Fisher Scientific, USA).

2.4 | PCR and amplicon sequencing

PCR was performed by students against two regions of the human *EGFR* gene and one region of the human *KRAS*, *TP53*, and *GAPDH* gene. Primers were designed to amplify regions containing known variants in NSCLC cell-lines.^{12–16} Forward and reverse primers were tagged at the 5' end with a unique 12-nucleotide sequence designed using Barcode Generator.¹⁸ The minimum number of mismatches between tags was set to 6 bp.¹⁹ Each group was given their own unique combination of tagged forward and

reverse primers for each target. The primers and the expected amplicon sizes are shown in Table 1. The 50 µl PCR reactions were performed using AccuPrime Pfx DNA polymerase (Thermo Fisher Scientific, USA) with two-step cycling conditions. Initial denaturation was performed for one cycle at 95°C for 2 min followed by 25 cycles at 95°C for 15 s and 68°C for 1 min. Students then ran 12 µl of PCR-amplified DNA in a 1.5% agarose gel. As two or more groups received the same starting DNA material, representative samples from each cell-line were chosen for sequence analysis to ensure read depth required for variant detection. Thus, the remaining 38 µl of a set of 12 total amplicons (four gene regions for each cell-line) were selected for pooling. Fragments below 150 bp were removed using a 0.9X bead cleanup with Ampure XP beads (Beckman Coulter, USA) and eluted in 20 µl of water. The final sequencing pool submitted for Oxford Nanopore 1D ligation for long-read sequencing was quantified at 222.8 ng/µl.

2.5 | Data analysis

Raw data sequencing data in the form of a FASTQ file was retrieved from the MinION sequencing run and transferred to the UVM Vermont Advanced Computing Core (VACC) for bioinformatic analysis. The NanoPack software tools were used to calculate the sequencing run metrics (Table 2).²⁰ FASTQ files were used to align raw sequence reads against the GRCh38 human reference genome using BWA-MEM.²¹ Aligned reads were then sorted and indexed using SAMtools²² and visualized with IGV.¹⁷

3 | RESULTS

3.1 | Module overview: Cancer variant detection through MinION amplicon sequencing

A new inquiry-based curriculum was designed and implemented for a molecular biology laboratory course at the University of Vermont with the overall goal to help students learn and apply practical NGS sequence

TABLE 1 Target specific primer sequences used for PCR amplification

Gene	Forward sequence (5'–3')	Reverse sequence (5'–3')	Amplicon size (bp)
<i>KRAS</i>	GGACCCTGACATACTCCCAAGG	AAGCGTCGATGGAGGAGTTTGT	544
<i>GAPDH</i>	TTCCGGAAACCAGATCTCCCAC	AAACCTGGGGGAATACGTGAGG	296
<i>EGFR15del</i>	TGTGGCACCATCTACAATTGC	AACAATACCCATGCTCCAGGCT	462
<i>EGFR</i>	TAATGGTCAGCAGCGGGTTACA	GGGCCTTTCCTGTTCCTTAGGT	721
<i>TP53</i>	TTCTTCTTGGCTGGGGAGAGG	CCCCTCCTCAGCATCTTATCCG	1400

analysis. In this project, senior level students generated and analyzed DNA sequences of known gene polymorphisms from unknown NSCLC cell-line samples. To support in the learning process outside of the laboratory, we developed a set of core teaching units to provide students with an introduction to nucleic acid analysis and variant detection using NGS technologies and bioinformatics sequence analysis (Figure 1). The four units included: (1) primer design for multiplex targets, (2) PCR amplification and purification of amplicon targets, (3) MinION sequencing, and (4) data analysis. Each unit was

accompanied by a corresponding pre-laboratory exercise and the project culminated in an oral and written report where students were asked to identify unknown NSCLC samples based on genotype information.

3.2 | Unit 1: Primer design for dual-index PCR amplification

The primary goal of the introductory unit was to offer students the opportunity to learn and apply PCR primer design that would be used to amplify target genomic loci of interest. Accordingly, students were provided with a table of genes and variants of interest in a prelab tutorial (see Appendix S1) on how to design PCR primers using the freely available Benchling software suite.²³ The Benchling platform enables database searching to retrieve gene sequence and annotation, as well as, use of the Primer3 software to perform primer design to capture each target of interest.²⁴ Specifically, our students were asked to design target primers at specific genetic loci according to NSCLC genotypes (i.e., *EGFR*, *TP53*, *KRAS*) (see Table S1). The primers designed by students amplified approximately

TABLE 2 Data summary of MinION sequencing run

Parameter measured	Mapped
Mean read length	747.9
Mean read quality	8.6
Median read length	506
Median read quality	8.5
Total number of reads	310,500
Read length N50	760
Total bases	232,217,962

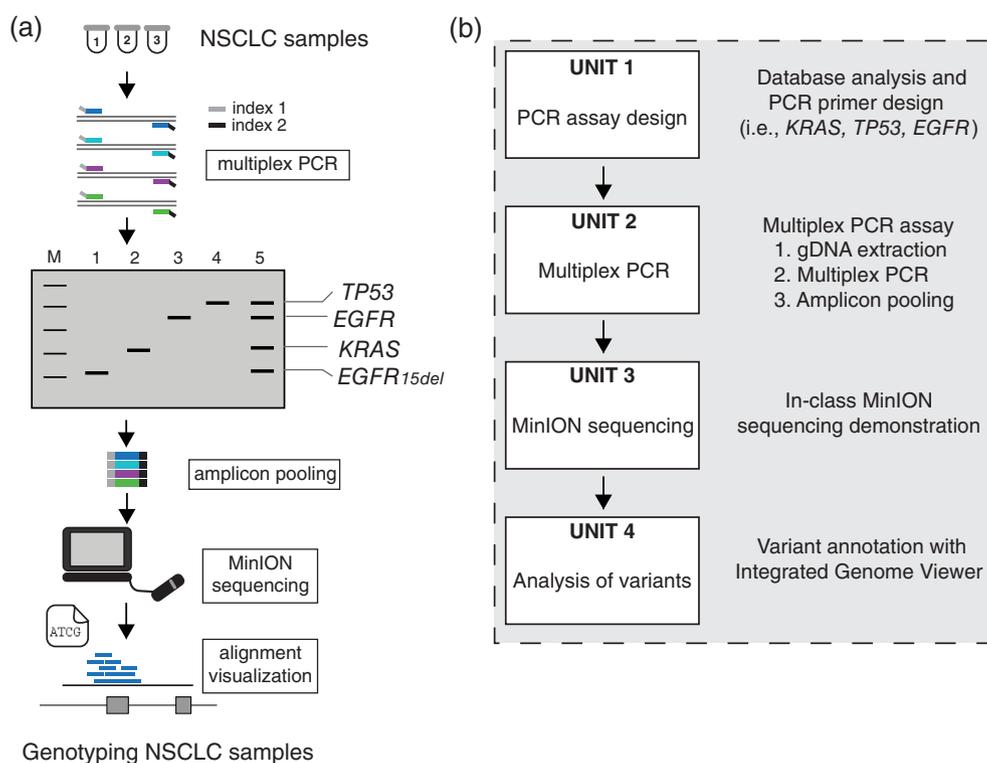


FIGURE 1 An overview of the 4-week module implemented at the University of Vermont to teach variant analysis with the MinION sequencing platform. (a) Students used primers they created to amplify DNA regions known to flank genetic variants in the *TP53*, *EGFR*, and *KRAS* gene in non-small cell lung cancer (NSCLC) cells using both single target and multiplex PCR. A portion of the PCR-product was then run on an agarose gel for size validation while the rest of the sample was pooled and sequenced on the MinION platform. After sequencing, students were asked to identify the starting NSCLC cell-line by genotype analysis using integrative genome viewer (IGV). (b) Flowchart summarizing the learning objectives and outcomes for the four units taught [Color figure can be viewed at wileyonlinelibrary.com]

1000 bp of genomic DNA containing the genomics variants of interest (Table 1). Each group was then assigned a 12-nucleotide group-specific DNA sequence as a barcode to incorporate into dual-index primer pairs. A list of the 12-nucleotide molecular barcodes to be added to the 5' end of each primer is provided (see Appendix S1). Each index primer averaged 37-nucleotides in length and the average cost per primer was \$7.50.

3.3 | Unit 2: Genomic DNA extraction and multiplex PCR

The second unit of this module focuses on the PCR method, including the extraction and quantification of

genomic DNA as a PCR template, preparation of working primer dilutions, and setting up individual or multiplex PCR reactions. The prelab material provided for this unit included a brief tutorial on the methods and principles of DNA extraction and quantification, as well as, the protocol used to perform primer dilutions for PCR (see Appendix S1). In the laboratory, students extracted genomic DNA from frozen cell pellets from NSCLC cell-lines labeled 'A', 'B' or 'C' (sample identities were blind to students). Students then quantified the total genomic DNA yield of each sample via the Nanodrop-2000 spectrophotometer. The typical yields of genomic DNA isolated by students was 20 ng/ μ l with a 260/280 absorbance range of 1.8–2.5 (see Figure S1).

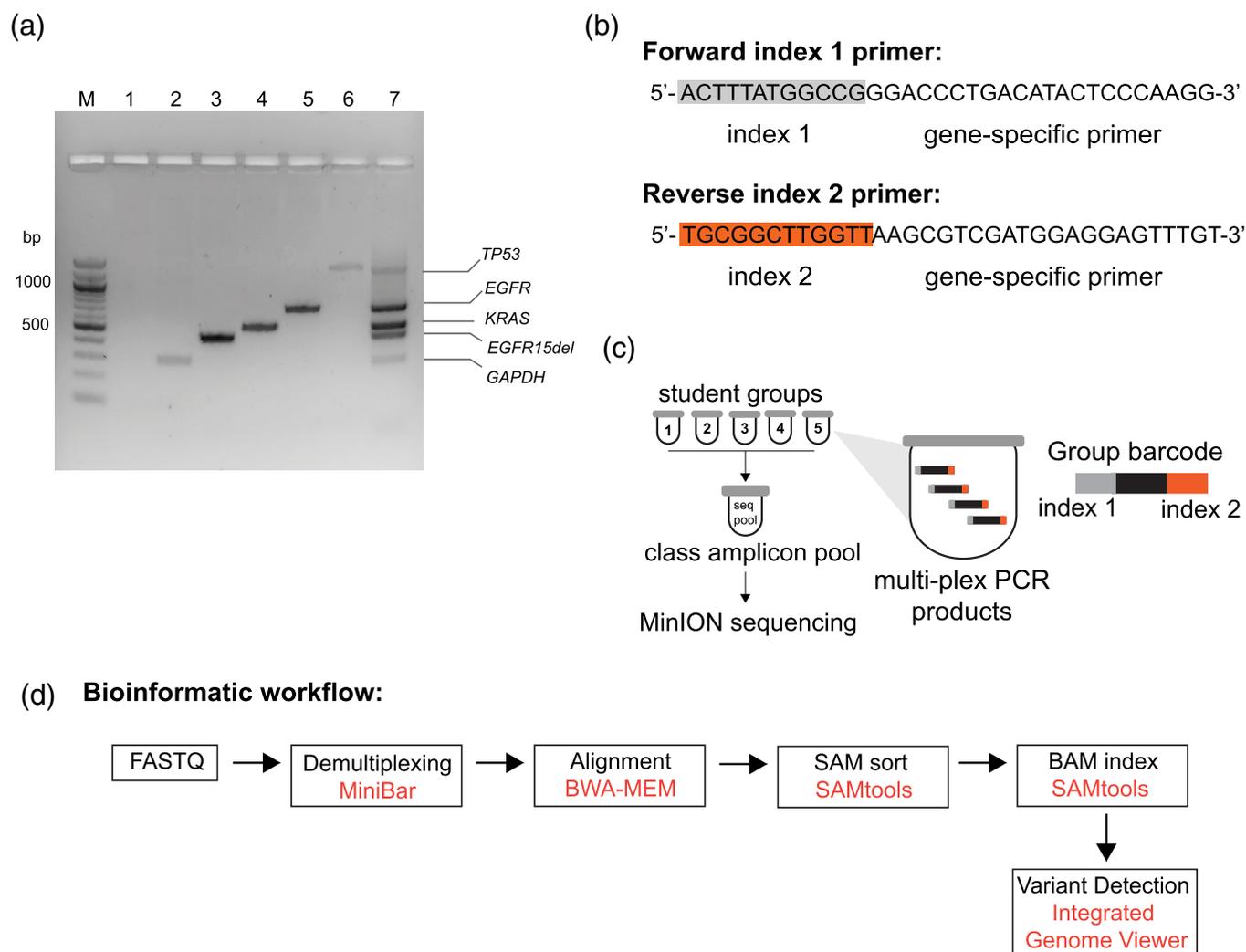


FIGURE 2 Class amplicon pool was sequenced on the MinION and the follow-up bioinformatic analysis performed. (a) Genotyping assay was performed to amplify regions encompassing genomic variants identified in the *TP53*, *EGFR*, and *KRAS* gene (lane 3–6). A multiplex PCR reaction containing all primers of interest are shown (lane 7). A reaction which lacked genomic DNA was included as a negative control (lane 1) while primers amplifying a region in *GAPDH* was included as the positive control (lane 2). (b) Each forward and reverse primer used in the PCR reactions contained a 12-nucleotide index at the 5' end. (c) The multiplex PCR reaction carried out by each student group were then pooled into a single tube (class amplicon pool) and sequenced on the MinION. (d) Data output from the MinION sequencing run was analyzed using the bioinformatic pipeline shown [Color figure can be viewed at wileyonlinelibrary.com]

Each group then made primer dilutions and used the genomic DNA sample as a template for both single target and multiplex PCR reactions (Figure 2(a)). A positive control primer set was provided to the entire class which targeted the *GAPDH* gene and routinely produced a distinctive 296 bp PCR amplicon. A negative control reaction which lacked genomic DNA was also included by each student group. Each student group had a distinct, non-redundant, index sequence added to the 5' end of the PCR primers used (Figure 2(b)). This dual indexing strategy was utilized for the purpose of identifying student group samples during the downstream bioinformatic analyses.

3.4 | Unit 3: PCR analysis, clean-up, and MinION sequencing

The third unit focuses on the generation of the pooled amplicon library, derived from the classes' multiplex PCR products, to be sequenced on the MinION platform in real-time. The prelab for this unit provided background information and instructions on how to perform DNA gel electrophoresis, PCR purification, and amplicon quantitation for pooling (see Appendix S1). Following interpretation of the DNA agarose gel, students performed PCR purification with magnetic beads and DNA quantification with the Qubit fluorometer. Group amplicons were pooled with all other samples and used for loading a single-use MinION Flongle flowcell as a live in-class demonstration (Figure 2(c)). To maximize student engagement, student groups were asked to submit questions to be answered during the question and answer period of the demonstration.

3.5 | Unit 4: Bioinformatic analysis of sample genotypes

The final unit began with an overview and detailed explanation of the bioinformatic analysis process and file conversion steps performed by instructors after receiving sequencing data (Figure 2(d)). In total, the MinION sequencing run produced 310,500 sequencing reads covering 232,217,962 bases (Table 2). Following demultiplexing of sequence reads based on dual-indexes, we found that each amplicon was overall proportionately represented in the sequencing run (Figure 3(a)). The read length distribution showed reads spanning from 350 to 850 bp with an average read length of 747.9 base pairs and PHRED score of 8.5 (Figure 3(b)). All amplicons mapped to their expected regions. The scripts used for file conversions, alignments, and quality control are provided in Appendix S1.

Each student group was provided with the sequencing results in the form of a BAM and BAI index files which were then used to perform variant analysis with the freely available Integrated Genome Viewer (IGV) software.¹⁷ The Unit 4 prelab distributed contained a tutorial on how to load the human genome and BAM files for visualization in IGV (see Appendix S1). Students were asked to identify and annotate if mutations at the specified genomic locations occurred when compared to the hg38 reference sequence. Upon examining gene alignments, student groups identified their assigned NSCLC cell-line based on the genomic profile. For example, groups that began with an HCC827 cells were able to visualize a characteristic 15-bp deletion found in exon 19 of the *EGFR* gene (Figure 4(a)), while groups that analyzed samples derived from H1975 cells were able to visualize the R273H mutation in *TP53* gene (Figure 4(b)).

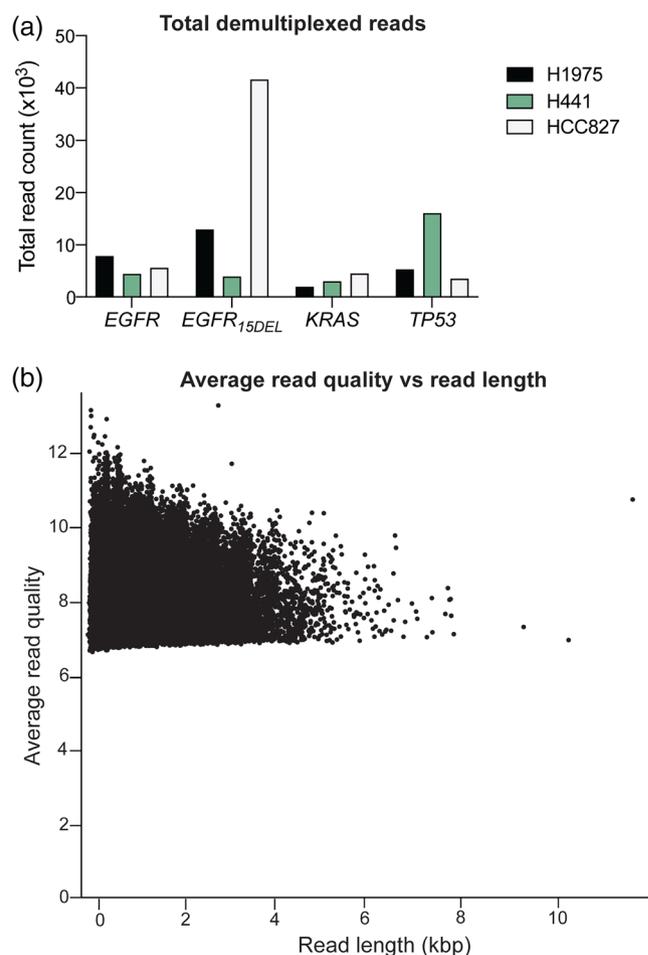


FIGURE 3 Quality of MinION sequencing reads. (a) The total demultiplexed reads for each amplicon sequenced is shown. (b) Plotting the average read quality versus read lengths (kbp). Each black dot represents a sequenced read that passed quality control cutoffs and were then used for downstream analysis [Color figure can be viewed at wileyonlinelibrary.com]

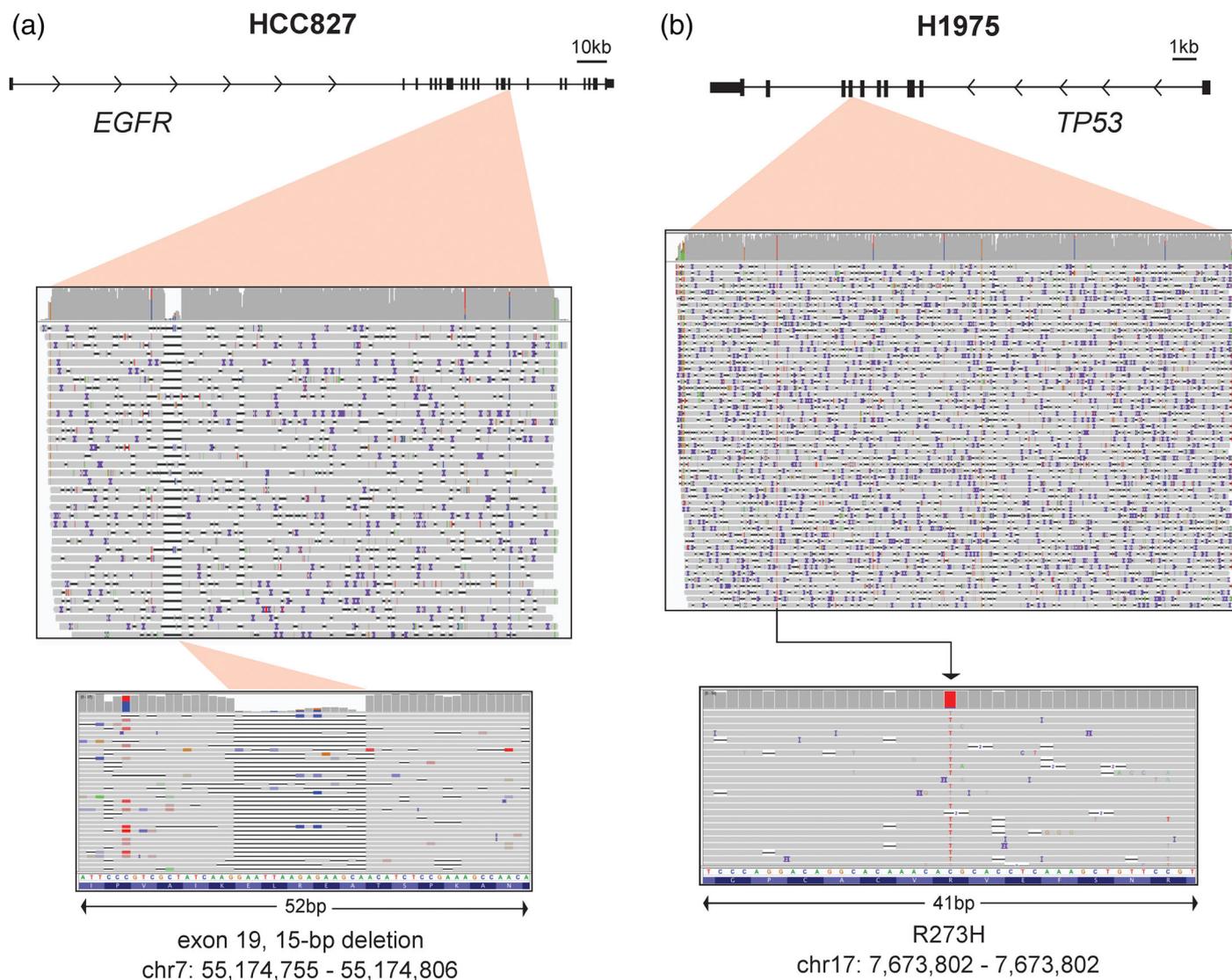


FIGURE 4 The integrated genomics viewer (IGV) was used by students to assess genomic variants in lung cancer cells. (a) Displaying the alignment of the genomic region amplified in the *EGFR* gene in HCC827 cells. A snapshot centered on the exon 19 deletion (15 bp) is shown. The gray horizontal bar represents a mapped read, while the lack of the gray bar indicates no reads mapped for that region. (b) Displaying the alignment of the genomic region amplified in the *TP53* gene in H1975 cells. A snapshot centered on the single nucleotide polymorphism R273H is shown. Bases that map to the reference genome are shown in gray, while a base mismatch is colored in red [Color figure can be viewed at wileyonlinelibrary.com]

3.6 | Assessment of student work

3.6.1 | Learning assessment

As individuals working in the biomedical laboratory fields require a strong background in scientific communication, we chose to evaluate students using both oral and written assessments. Therefore, in addition to the pre-laboratory assignments that accompanied each unit, this 4-week module culminated to a final project accounting for 25% of the course grade. The project continued to foster group work, through a group oral presentation to be presented to

the entire class. In addition, individual student learning was evaluated through a 3-page written summary.

For the oral portion, each student group was allotted a total of 15 min (12 min for the presentation and 3 min for questions). All student groups were instructed to incorporate the following during their presentations:

1. A broad overview of the analysis and data used for genotyping the NSCLC samples.
2. Describe the normal function of the genes that were mutated and why these specific genes are considered diagnostic in cancer.

- Identify and describe if these two genes are altered in other cancers.
- Read and describe a primary reviewed journal article that incorporated a NSCLC cell-lines to study one of the genetic alterations identified.

For the 3-page written summary, in addition to answering the questions above, students were asked to include the purpose of the research undertaken in the journal article selected and to summarize the results and major findings of the study.

Individual student participation was assessed by instructors as follows:

- Each student was asked to submit a peer evaluation form for each group presentation.
- Question(s) queried by the student during the question portion of the group presentation.

The rubric used to evaluate the group presentations and a summary of select written feedback provided by students is shown (see Table S3). Finally, to accommodate students with documented learning disabilities, students were permitted to pre-record their 12-min presentation to be viewed by the entire class, as well as, submit questions to be answered at a later time by writing them down on an index card and submitting them to the instructor after the session ended.

3.7 | Student perception survey

During the final week of classes, students were given a 5-question survey to determine student perception after completing this module. Out of the 24 students enrolled

in the course, 18 completed the evaluation. In total, 83% and 94% of students reported this was their first time learning about the MinION sequencing platform and IGV, respectively (Table 3). After participating in the laboratory, a total of 38% and 60% of students reported feeling more comfortable than not to prepare samples for MinION and to independently use IGV for the purpose of genetic variant analysis, respectively. Finally, 50% of students reported to improve upon their knowledge in NGS technologies after participating in this 4-week module.

Some students also provided general comments about their experience participating in this 4-week module. Students commented it was an overall 'positive lab experience' and 'loved that the module had a purpose'. However, some students did comment that at times the instruction felt 'rushed'. As a suggestion, a couple of students recommended spending more time, before starting 'hands-on' work, on the background or with the assignment tutorials provided.

4 | DISCUSSION

In this article, we present an accurate, specific and sensitive multiplex PCR reaction that can be sequenced in real-time on the MinION platform to provide students with inquiry-based hands-on learning of DNA sequence-based diagnostic tests. In the past, there were many barriers to overcome by instructors who sought to incorporate DNA sequencing into the teaching laboratory curriculum. With the advent of third generation sequencing platforms, DNA sequencing has become more accessible for a greatly reduced cost. The results reported here serve as proof-of-concept that MinION sequencing can be implemented in the teaching laboratory for the purpose

TABLE 3 Student evaluation of the 4-week module (n = 18)

	Definitely true		More true than false		In between		More false than true		Definitely false	
	%	n	%	n	%	n	%	n	%	n
This was my first time learning about MinION sequencing	83.3	15	5.6	1	0.0	0	5.6	1	5.6	1
I feel comfortable preparing samples to be sequenced on the MinION platform	11.1	2	27.8	5	38.9	7	11.1	2	11.1	2
This was my first time using Integrative Genome Viewer (IGV)	94.4	17	0.0	0	5.6	1	0.0	0	0.0	0
I feel comfortable using IGV to analyze genetic variants	38.9	7	22.2	4	22.2	4	5.6	1	11.1	2
I have improved upon my knowledge in NGS technologies due to participating in this module	50.0	9	22.2	4	11.1	2	16.7	3	0.0	0

of cultivating scientific inquiry and providing experience for methods used in contemporary molecular biology.

We provide a 4-week module with a week-to-week guide including assignments that cover primer design, multiplex PCR, sample preparation, and entry-level bioinformatics to be used in the molecular biology laboratory class setting. Using MinION sequencing, all student groups were able to correctly identify the NSCLC sample they were assigned, based on its unique genomic profile. The MinION system was able to produce the data in the time frame of a single class. The size and ability of the MinION to easily hook-up to a laptop allowed us to feature an in-class demonstration where students watched sequencing occur live. Students then used this data to detect both deletions and single nucleotide polymorphisms characteristic to lung cancer samples. To adapt to distance-learning, the MinION demonstration could be easily broadcasted and streamed online by instructors. Importantly, the raw data is available on NCBI and therefore, a majority of this lab exercise could be adapted for an entirely remote laboratory experience.

In conclusion, the MinION system allowed for the detection of genetics variants in non-small cell lung cancer cell-lines. By implementing this 4-week module students were able to take ownership in the design and sample preparation required for NGS sequencing. This module serves as a useful tool for instructors which could be formatted to accommodate a non-BLS2 (Biosafety Level 2) laboratory and modified for off-site learning. Furthermore, many outcomes were achieved as a result of this laboratory module. First, student independence and proficiency were enhanced by allowing them to design and test their own PCR primers. Second, students were introduced to advanced molecular assays including multiplex PCR and NGS sample preparation, in addition to computational experience in a practical introductory manner. Finally, this module can be effectively incorporated into a molecular biology curriculum as it could function to enhance genetic variant analysis and core concepts in molecular pathology and oncology. This NGS-based module is an easily implementable way to introduce a wide range of students to contemporary molecular biology techniques and bioinformatics, and also serves as a springboard for interested students to pursue further training in research and medical diagnostics.

ACKNOWLEDGMENTS

Special thanks to the students of the Applied Molecular Biology Laboratory, BHSC 282, fall of 2019 for their participation in this 4-week module. We would also like to thank Scott Tighe and Pheobe Kehoe of the Vermont Integrative Genomics Resource (VIGR). Materials used in the BHSC 282 course was funded by the University of

Vermont, Department of Biomedical and Health Sciences. We would also like to acknowledge the following funding sources, NSF OIA-1736097 (SF), T32AI055402 (PDR) and U54GM115516 from the National Institutes of Health for the Northern New England Clinical and Translational Research network.

DATA AVAILABILITY

The sequence data from this study have been submitted to the NCBI Short Read Archive (<https://www.ncbi.nlm.nih.gov/sra/>) under accession number PRJNA668750.

ORCID

Seth Fretze  <https://orcid.org/0000-0003-4058-3661>

REFERENCES

1. Taylor S, Bennett KM, Deignan JL, Hendrix EC, Orton SM, Verma S, et al. Molecular pathology curriculum for medical laboratory scientists: a report of the association for molecular pathology training and education committee. *J Mol Diagn.* 2014;16(3):288–96.
2. Houwink EJ, Muijtjens AM, van Teeffelen SR, Henneman L, Rethans JJ, Jacobi F, et al. Effect of comprehensive oncogenetics training interventions for general practitioners, evaluated at multiple performance levels. *PLoS One.* 2015;10(4):e0122648.
3. McClaren BJ, Crellin E, Janinski M, Nisselle AE, Ng L, Metcalfe SA, et al. Preparing medical specialists for genomic medicine: continuing education should include opportunities for experiential learning. *Front Genet.* 2020;11:151.
4. Freeman S, Eddy SL, McDonough M, Smith MK, Okoroafor N, Jordt H, et al. Active learning increases student performance in science, engineering, and mathematics. *Proc Natl Acad Sci U S A.* 2014;111(23):8410–5.
5. Jain M, Olsen HE, Paten B, Akeson M. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome Biol.* 2016;17(1):239.
6. Laszlo AH, Derrington IM, Ross BC, Brinkerhoff H, Adey A, Nova IC, et al. Decoding long nanopore sequencing reads of natural DNA. *Nat Biotechnol.* 2014;32(8):829–33.
7. Kasianowicz JJ, Brandin E, Branton D, Deamer DW. Characterization of individual polynucleotide molecules using a membrane channel. *Proc Natl Acad Sci U S A.* 1996;93(24):13770–3.
8. Mikheyev AS, Tin MM. A first look at the Oxford Nanopore MinION sequencer. *Mol Ecol Resour.* 2014;14(6):1097–102.
9. Wongsurawat T, Nakagawa M, Atiq O, Coleman HN, Jenjaroenpun P, Allred JI, et al. An assessment of Oxford Nanopore sequencing for human gut metagenome profiling: a pilot study of head and neck cancer patients. *J Microbiol Methods.* 2019;166:105739.
10. Norris AL, Workman RE, Fan Y, Eshleman JR, Timp W. Nanopore sequencing detects structural variants in cancer. *Cancer Biol Ther.* 2016;17(3):246–53.
11. Minervini CF, Cumbo C, Orsini P, Brunetti C, Anelli L, Zagaria A, et al. TP53 gene mutation analysis in chronic lymphocytic leukemia by nanopore MinION sequencing. *Diagn Pathol.* 2016;11(1):96.

12. Amann J, Kalyankrishna S, Massion PP, Ohm JE, Girard L, Shigematsu H, et al. Aberrant epidermal growth factor receptor signaling and enhanced sensitivity to EGFR inhibitors in lung cancer. *Cancer Res.* 2005;65(1):226–35.
13. Bentivegna S, Zheng J, Namsaraev E, Carlton VE, Pavlicek A, Moorhead M, et al. Rapid identification of somatic mutations in colorectal and breast cancer tissues using mismatch repair detection (MRD). *Hum Mutat.* 2008;29(3):441–50.
14. Chou CW, Lin CH, Hsiao TH, Lo CC, Hsieh CY, Huang CC, et al. Therapeutic effects of statins against lung adenocarcinoma via p53 mutant-mediated apoptosis. *Sci Rep.* 2019;9(1):20403.
15. Iwakawa R, Kohno T, Enari M, Kiyono T, Yokota J. Prevalence of human papillomavirus 16/18/33 infection and p53 mutation in lung adenocarcinoma. *Cancer Sci.* 2010;101(8):1891–6.
16. Onitsuka T, Uramoto H, Tanaka F. Lack of direct association between EGFR mutations and ER beta expression in lung cancer. *Anticancer Res.* 2011;31(3):855–60.
17. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol.* 2011;29(1):24–6.
18. Comai L, Howell T. Barcode Generator. Comaiwiki. 2012. Retrieved October 1, 2019. http://comailab.genomecenter.ucdavis.edu/index.php/Barcode_generator.
19. Srivathsan A, Baloglu B, Wang W, Tan WX, Bertrand D, Ng A, et al. A MinION™-based pipeline for fast and cost-effective DNA barcoding. *Mol Ecol Resour.* 2018;18(5):1035–9. <https://doi.org/10.1111/1755-0998.12890> Advance online publication.
20. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics (Oxford, England).* 2018;34(15):2666–9.
21. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997 [q-bio.GN]. 2013.
22. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinformatics (Oxford, England).* 2009;25(16):2078–9.
23. Benchling [Biology Software]. 2019. Retrieved from <https://benchling.com>.
24. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3—new capabilities and interfaces. *Nucleic Acids Res.* 2012;40(15):e115.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Rodriguez PD, Mariani M, Gay J, et al. A guided-inquiry investigation of genetic variants using Oxford nanopore sequencing for an undergraduate molecular biology laboratory course. *Biochem Mol Biol Educ.* 2021;49:588–597. <https://doi.org/10.1002/bmb.21514>