AMERICAN SOCIETY FOR MICROBIOLOGY | JMBE Journal of Microbiology & Biology Education

Check for updates

| Curriculum

A course-based undergraduate research experience (CURE) designed for modern genetics and biodiversity courses using a classroom benchtop sequencer

Erika T. Ebbs,[1] Allyson Barredo,[1] Amelia Whitehurst,[1] Stephen E. Harris[1]

AUTHOR AFFILIATION See affiliation list on p. 13.

ABSTRACT Modern genetics increasingly relies on genomic data sets to address medical, ecological, and evolutionary questions. Investigating these questions requires a diverse set of core competencies in wet-lab techniques, data analysis, and bioinformatics. We describe the design and implementation of a nine-week course-based undergraduate research experience (CURE) embedded within an upper-level Genetics Laboratory course. This CURE exposed students to contemporary wet-lab and analytical techniques related to genomic sequencing and biodiversity analysis. Each student began with an intact biological specimen and independently completed DNA extraction, quantification, library preparation, and sequencing using Oxford Nanopore Technologies. Students were subsequently trained in basic bioinformatic and phylogenetic analyses. To simulate an authentic research experience, students worked with real research samples, engaged in iterative data evaluation, and were responsible for troubleshooting and planning next steps. Data analysis was student-driven; not all students participated in all aspects of the analysis, allowing for individual ownership and specialization. Modules on quantitative reasoning and scientific communication were integrated into the curriculum to ensure consistent development of key skills. This CURE focused on understudied groups of trematodes (Platyhelminthes), with uncertain taxonomic placement, providing an opportunity to explore species discovery, marker selection, intra- vs. interspecific variation and zoonotic disease transmission. "Teaching-Research Synergy" is central to this curriculum, which is adaptable to other study systems and can be leveraged to support a faculty's research program. We present findings from two iterations of this CURE and discuss how student-generated data can contribute meaningfully to faculty-led research.

KEYWORDS CURE, trematodes, biodiversity education, genetics education, mitogenomics, species discovery, nanopore sequencing

Course-based undergraduate research experiences (CUREs) are well-established in the biology education literature as effective tools for enhancing student learning outcomes (SLOs) (1), strengthening critical and quantitative reasoning skills (2), reducing affective barriers to learning (3, 4), and narrowing achievement gaps (5). CUREs can vary in scope, ranging from single laboratory modules to multi-week or semester-long research experiences (6, 7). When embedded within upper-division, discipline-specific STEM courses, CUREs can provide students with authentic, high-impact research experiences they may not otherwise access (8, 9) while also helping them develop transferable, marketable job skills (10).

This is particularly relevant in rapidly evolving fields such as biotechnology, genomics, and bioinformatics (11, 12), where demand for workforce-ready graduates continues to grow. CUREs have been developed to engage students across diverse subfields of

Editor David R. Wessner, Davidson College, Davidson, North Carolina, USA

Address correspondence to Erika T. Ebbs, erika.ebbs@purchase.edu.

The authors declare no conflict of interest.

Received 26 June 2025
Accepted 15 December 2025
Published 9 January 2026

genetics, genomics, and bioinformatics (6, 7, 13). Core competencies in these areas encompass a range of skills, including wet-lab techniques, next-generation sequencing, R programming, database mining, coding, and data visualization (12, 14).

One persistent challenge in achieving these competencies through CUREs is the substantial knowledge and technical training that must be scaffolded (15). Moreover, implementing multi-week or semester-long CUREs requires significant time and effort from faculty—a commonly cited barrier to adoption (16). In fields such as genetics and genomics, the rapid pace of bioinformatics tool development adds an additional "time-sink" barrier to curriculum design (6, 15). Designing CUREs with the intent of "Teaching-Research Synergy," such that the CURE aligns with the faculty research program, mitigates these challenges (17) as well as enhances the sustainability and overall impact (18, 19).

Here, we describe the design and implementation of an open-ended, 9-week (half-semester) CURE embedded within an upper-level Genetics Laboratory course at Purchase College, State University of New York. This CURE is intentionally aligned with the faculty's research program to (i) enhance sustainability and faculty engagement, (ii) maximize the use of available laboratory resources and faculty expertise, and (iii) fuel faculty research projects through data collection. The CURE aligns with a research program aimed at characterizing the trematode (Platyhelminthes: Trematoda) communities of North American waterbirds using whole mitochondrial genome data. Digenetic trematodes are a group rich in undescribed species diversity (20) and matched with a dearth of molecular genetic resources. This gap presents a low threshold for the detection of novel species, making trematodes ideal organisms for engaging students in species discovery and molecular identification (21). Additionally, trematodes are significant from both public health and economic perspectives, which grab student interest. Improved molecular tools are essential for accurately identifying infections to the species level (22), and for understanding life cycles and transmission dynamics (23). In this CURE, each student was provided with an "unknown" adult trematode specimen and tasked with applying next-generation sequencing (NGS), molecular genetics, and phylogenetic approaches to identify the species and investigate its life history using the published literature.

This curriculum is broadly adaptable for faculty engaged in biodiversity research, particularly involving invertebrate taxa, and can be modified to support diverse research programs while offering students meaningful, research-rich experiences.

Intended audience and prerequisite student knowledge

This CURE was developed for an upper-level undergraduate Genetics Laboratory course, typically enrolling sophomores through seniors. The framework is adaptable and could be implemented within Genomics, Biotechnology, or Integrative Taxonomy courses. It is expected that students have completed General Biology. Students enrolled in the CURE are either concurrently taking or have previously completed the Genetics lecture. Several quantitative reasoning activities (QRAs) are embedded within the CURE, utilizing R programming, which students begin learning during the first half of the Genetics Laboratory course. These QRAs can be readily adapted to other statistical software platforms, depending on instructor preference.

Learning objectives

The specific learning objectives for this CURE are outlined in Table 1, along with corresponding assessment strategies. The CURE was designed to develop transferable skills in three key areas: (i) wet-lab techniques, (ii) analytical skills relevant to modern genetics and genomics (14, 18), and (iii) science communication. In the sections that follow, we describe the overall CURE curriculum and present data and observations from implementation during the Spring 2023 and Spring 2024 semesters.

TABLE 1 Student learning objectives (SLOs) across three specific areas are summarized with corresponding assessment methods[a]

Learning objectives	Assessment method(s)
Wet-lab skills	
1. Implement different approaches to extracting, quantifying, and evaluating DNA extractions a. Perform Qubit quantification b. Perform NanoDrop quantification c. Judge DNA quality data	Practical demonstration (all), QRA1 (LO 1), QRA2 (LO 2), Written Assessment (LO 1&2), Exit assessment (LO 1&2)
2. Perform library preparation and ONT sequencing a. Perform serial dilutions b. Prepare DNA libraries for ONT sequencing c. Demonstrate multiplexing samples d. Explain the process of ONT sequencing	
3. Develop micropipetting skills	
4. Maintain sterile technique	
Analytical and fundamental genetics skills	
1. Apply computational concepts used in bioinformatics a. Appraise a bioinformatic pipeline or workflow b. Create FASTA files	Practical demonstration (all), QRA1 (LO 2 and 14), QRA 3 (LO 2, 4, 5, and 14), QRA 4 (LO 11, 13, and 14), Written assessment (LO 4–13), Exit assessment (all)
2. Apply statistical correlations to sequence data a. Perform and interpret linear regression b. Perform and interpret Wilcoxon rank tests	
3. Compare structure/inheritance of the nuclear and mitochondrial genomes	
4. Articulate molecular genetic approach to species identification	
5. Interpret bioinformatics statistics (*e*-value, similarity scores)	
6. Perform a *de novo* assembly	
7. Differentiate between sequence "reads" and "contigs"	
8. Select an appropriate reference genome	
9. Use genomic databases to… a. Mine sequence data to test hypotheses b. Navigate a Genome Browser c. Build reference databases d. Perform and interpret BLAST searches	
10. Perform mitochondrial genome mapping a. Evaluate sequence coverage b. Evaluate sequence similarity	
11. Build a multi-sequence alignment	
12. Demonstrate understanding of general phylogenetic methods; a. Evaluate marker choice b. Describe basic phylogenetic algorithms c. Interpret phylogenetic trees and branch support values d. Explain phylogenetic trees	
13. Compare interspecific and intraspecific variation	
14. Demonstrate competence in multiple software packages to manipulate and analyze bioinformatics data (Geneious, R, MEGA X, and spreadsheets)	
Scientific communication skills	
1. Write the methods and results section of a scientific article	Written assessment and rubric scores over time
2. Presentation of scientific data a. Identify necessary background information b. Identify and articulate the implications of results c. Design a scientific poster	

[a]QRA refers to quantitative reasoning activity. LO refers to learning objectives.

PROCEDURE

Learning time

This CURE was designed to take place over a 9-week period during the second half of the Spring semester, with Spring Break falling in the middle of the CURE. The Genetics

Laboratory meets once a week for 170 min. Outside of class time, the students have access to the Genetics lab and computers, as well as faculty (2 h/week) and learning assistant (1 h/week) office hours. The CURE spans eight in-class sessions. It is intentionally aligned with Purchase College's end-of-semester Natural and Social Science Symposium, where a subset of CURE participants present their work in the form of a class poster or posters.

Safety issues

The trematode samples and reagents used within this CURE present a low risk to student safety. The adult (non-infective stage) trematode samples are preserved in ethanol and are not infectious to humans. Students were instructed to follow basic BSL-1 guidelines during all lab sessions, including wearing closed-toe shoes, wearing gloves when handling any laboratory chemicals, and proper handwashing.

DNA extraction and quantification (Weeks 1 and 2)

This CURE utilizes trematode specimens collected as part of the faculty's research program; accordingly, wet-lab procedures were optimized for processing trematode tissue. Each student was provided with a single trematode preserved in 95% ethanol and stored at –20°C. All host organisms were obtained through donations from the New York Department of Environmental Conservation (DEC) or local permitted hunters. Instructors working with larger organisms may need to section tissues prior to lysis, as input tissue volume should follow the recommendations of the DNA extraction kit used.

This CURE employed the QIAamp DNA Mini Kit (Qiagen), following the manufacturer's protocols. One day prior to the first lab session, preserved trematodes were transferred by the learning assistant to individual 1.5 mL tubes, and residual ethanol was removed by evaporation using a Vacufuge. Tissue lysis was initiated according to the manufacturer's instructions and incubated overnight at 56°C on a heat block. During the lab session, each student received a lysed sample along with pre-aliquoted reagents necessary for completing the extraction. Column-based extraction methods are preferable for trematodes due to their small size and low DNA yield; however, alternative protocols such as HotSHOT lysis (24) can be used to reduce the time and cost associated with this step.

In this CURE, we elected to use both NanoDrop and Qubit quantification (Week 2, Table 2). Qubit (Qubit dsDNA HS Assay Kit) quantification is optimal for downstream Oxford Nanopore Technologies (ONT) sequencing, while the low-cost NanoDrop provides multiple data points that allow students to conceptualize DNA quantity and quality. *Quantitative Reasoning Activity 1* (File S1) was developed using NanoDrop quantification data for 56 trematode individuals (data provided by the instructor). Students used this data and data visualization tools in R to identify patterns in DNA extraction outcomes and to make recommendations to improve extraction protocols. This curriculum could be streamlined by combining Week 1 (DNA extraction) and Week 2 (DNA quantification) into a single lab session. However, anecdotally, dedicating time in the first week to introducing and generating enthusiasm for the CURE is beneficial. A substantial portion of this initial session is devoted to discussing the overarching research goals, showcasing student work from previous cohorts, discussing trematode biology, and examining specimens from diverse trematode taxa.

All class data collected throughout the CURE was organized into a shared Google Sheets file where all students contributed their data (File S2). The instructor pre-populated column headings to guide students through the data collection process. As all students were granted editing privileges, the sheet was downloaded by the instructor before and after each lab session for version control.

Library preparation, sample barcoding, and ONT sequencing (Week 3)

DNA libraries were prepared using the Ligation Sequencing Kit V14 (Oxford Nanopore). The Rapid Barcoding Kit 24 V14 (Oxford Nanopore) was used to multiplex samples, which

TABLE 2 Timeline of the 8-week CURE, including pre- and post-lab work[a]

Description	Time	Notes
Week 1 - Introduction & DNA Extraction		
Pre-Lab	20-30m	Students are provided with the manufacturer's protocol and asked to make a graphical step-by-step protocol
• Outline DNA extraction protocol		
Lab	30m 20m 100m	12-24 hr before the lab meeting the instructor lyses tissue samples following the manufacturer's protocol (15m)
• Introduction to the CURE lecture		
• Introduction to your organism		
• DNA extraction (*MiniAmp Kit Qiagen*)		
Post-Lab	20-30m	See section *Science Communication and Writing*
• DNA extraction protocol paragraph		
Week 2 - Evaluating DNA Extractions & Planning Next Steps		
Pre-Lab	10m	
• DNA quantification video		
Lab	20m 60m 45m 20m	
• Introductory lecture		
• DNA quantification (Qubit & NanoDrop analysis)		
• QR activity 1: Analyzing NanoDrop data in R		
• Class discussion: *Are the samples sufficient to move forward*?		
Post-Lab	10-20m	See section *Science Communication and Writing*
• DNA quantification protocol paragraph		
Week 3 - ONT Library Preparation & DNA Sequencing		
Pre-Lab	30-40m	Students are provided with the manufacturer's protocol and asked to make a graphical step-by-step protocol
• ONT Sequencing video & protocol outline		
• QR activity 2: Start sample dilution calculator (File S1)		
Lab	15m 30m 100m	Students are split into two groups and alternate performing ONT library preparation (with instructor) and QRA 3 (independently, File S1)
• QR activity 2: Complete sample dilution calculator		
• Class discussion: *Are the samples sufficient to move forward II?*		
• ONT library preparation and sample barcoding		
• Loading the flow cell (demonstration)		
• QR activity 3: Analyzing mitochondrial SNVs		
Post-Lab	20-30m	See section *Science Communication and Writing*
• ONT library preparation and barcoding protocol paragraph		
Week 4 - Initial Bioinformatic Analysis		
Pre-Lab	10m 10m	Students learned how to use NCBI and BLAST during the first half of the semester
• Introduction to Geneious Prime video		
• Genomic database refresher video		
Lab	30m 20m 30m 20m 20m 20m	These time estimates will vary based on the size and quality of the sequence data students obtain
• Introductory Genomics lecture		
• *De-novo* assembly (preliminary data)		
• Reference genome selection		
• Mapping to the mitochondrial genome (preliminary data)		
• Class discussion: *What data do we have to work with?*		
• Paper assignment & independent analysis discussion		
Post-Lab	30m 60m	
• Summary of ONT sequencing results paragraph		
• Begin paper assignment		
SPRING BREAK - Paper Assignment & Identify Independent Analysis		
Week 5		
Pre-Lab		Students determine three different analyses they could do with the data they collected, based on their paper assignments
• Submit a ranked list of potential analyses		

(*Continued on next page*)

TABLE 2 Timeline of the 8-week CURE, including pre- and post-lab work[a] (*Continued*)

Description	Time	Notes
Lab	60m 110m	Instructor will work with each group individually to plan their analyses
• Class discussion of paper assignment & student lists - refine analyses as needed		
• Break up into groups (2-3 students), begin independent analyses		
Post-Lab	20m	See section *Science Communication and Writing*
• Submit Summary of Independent Analysis		
Week 6		
Pre-Lab		Have students attend office hours to 'check-in' on their independent analyses
• Work on independent analyses as needed		
Lab	110m 60m	
• Small groups work on independent analyses		
• Class Discussion: *Small groups update & begin poster planning*		
Post-Lab		
• Work on independent analyses as needed		
Week 7		
Pre-Lab		Have students attend office hours to 'check-in' on their independent analyses
• Work on independent analyses as needed		
Lab	125m 45m	
• Small groups work on independent analyses		
• Class Discussion: *Small groups update & poster preparation*		
Post-Lab		See section *Science Communication and Writing*
• Work on final presentations & write up		
Week 8		
Pre-Lab		Students attend office hours to 'check-in' on their independent analyses
• Work on independent analyses as needed		
Lab	170m	See section *Science Communication and Writing*
• Final Presentation and Poster Symposium		
Post-Lab		

[a]The Genetics Laboratory meets one day per week for 170 min. Students have access to computer stations throughout the week to work independently as needed.

enabled upstream sequencing data to be linked to individual specimens. Multiplexed samples were run on a single flow cell (R10.4.1). Sequencing was performed using MinKNOW, Oxford Nanopore's proprietary software, with live basecalling over a 24-h run. Reads with a quality score ≥ q10 were demultiplexed and exported as .fastq files. Sequencing runs were initiated during class time, allowing students to observe and engage with the process in real time, including monitoring the first 30–60 minutes of the run and evaluating initial output files. For instructors without access to a benchtop sequencer, this CURE can be modified to utilize other NGS platforms, incorporate Sanger sequencing, or leverage publicly available genomic databases for sequence analysis.

Bioinformatic analysis and molecular phylogenetics (Weeks 4–8)

All bioinformatic and phylogenetic analyses (Fig. 1) were conducted using Geneious Prime (Version 2024). Temporary licenses were provided to students by Geneious upon request, and the software also offers a free trial. Many of the software discussed below and implemented in Geneious Prime (Velvet, Bowtie, minimap2, and BLASTn) are freely available for use in terminal without the graphical user interface.

Initial bioinformatic analysis (Week 4)

Students began by analyzing sequencing reads from their samples. Their first task was to summarize the raw sequencing data (number of reads and average read length) within Google Sheets (File S2), followed by performing a *de novo* assembly to organize total

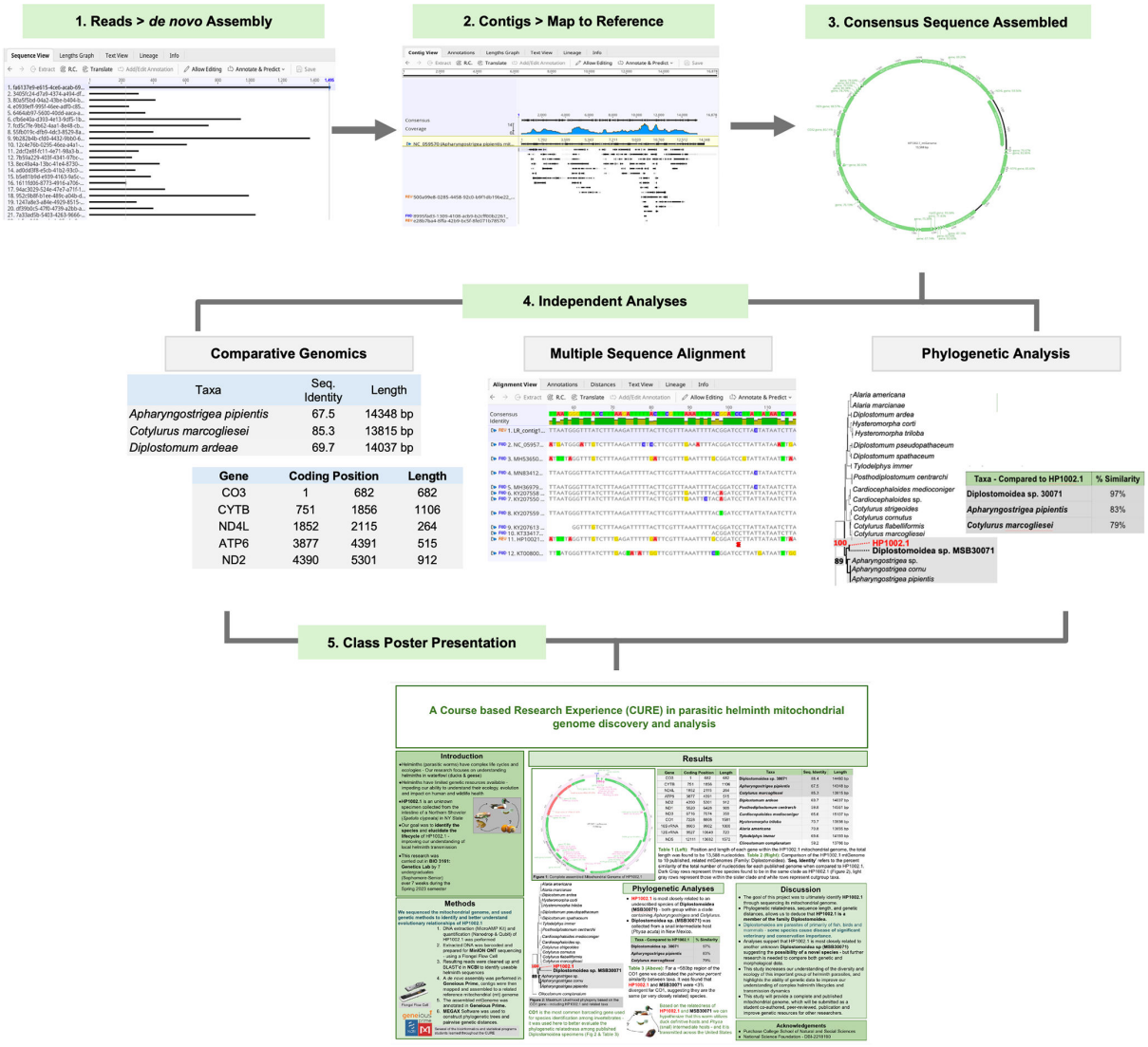

FIG 1 Workflow of CURE. All students use the sequence data that they generated from their individual specimens to execute steps 1–3. Students then break up into small groups (two to three students) to execute independent analyses (4) and generate products to contribute to the final class poster presented at the Purchase College Natural and Social Science Symposium (5).

reads into contigs (Fig. 1). They then assessed the quality of their contigs, exploring key concepts such as sequencing depth, coverage, and contig formation. Students compared different assembly algorithms implemented in Geneious (Geneious, Velvet, and Bowtie). Within Geneious, students conducted BLAST searches of their contigs to obtain tentative taxonomic identifications. In several cases, BLAST results revealed the presence of host, bacterial, viral, or other contaminant DNA. If contamination was extensive, the instructor had the option to pre-filter reads (e.g., removing bacterial sequences) before providing the remaining .fastq files to students. BLAST results were compiled into a shared Google Sheet, and the class collaboratively discussed the most likely taxonomic assignments. If faculty do not have access to Geneious, batch BLAST searching can be performed in terminal, or on the NCBI website (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

Mapping to reference genomes and annotating reads

Given that species identification was a primary objective, the class next discussed the use of informative phylogenetic markers and strategies to locate them within

their contigs. Students were first tasked with identifying mitochondrial genome reads. Using the NCBI Organelle Database (https://www.ncbi.nlm.nih.gov/datasets/organelle/?taxon=2759), they downloaded mitochondrial genomes closely related to their sample (based on BLAST results) to create a reference database. They then used the "Map to Reference" tool in Geneious to align their reads to the mitochondrial reference genome. If faculty do not have access to Geneious, the minimap2 software can be utilized for free in terminal. While some samples yielded no mitochondrial reads, others returned partial coverage across different regions of the mitogenome. Students with no recovered reads were able to view and analyze mapping results from peers, and outcomes were discussed as a group.

Similarly, students explored the nuclear rRNA gene cluster—specifically the 18S, 5.8S, 28S, ITS1, and ITS2 regions—which are commonly used phylogenetic markers in trematodes. They compiled relevant rRNA sequences into a reference database and performed mapping of their reads to these markers using Geneious.

Independent analyses and phylogenetic workflows

After Week 4, students began developing independent analysis projects using the cumulative class data set, allowing each student' experience within the CURE to become more individualized. These projects took multiple forms (Table 3), and the associated analytical approaches varied accordingly (Fig. 1). Over spring break, students were assigned two to five peer-reviewed articles, depending on class size, and asked to identify specific analyses, tables, or figures they would like to replicate using their own data. Upon returning from break, students submitted a ranked list of potential analyses to pursue. The first portion of the subsequent class was dedicated to refining and finalizing project ideas, resulting in groups of two to three students each contributing a unique analysis to the overall class project.

Building alignments

To construct phylogenetic data sets, students mined NCBI for sequences of related taxa and built multiple sequence alignments (MSAs) using Clustal Omega within Geneious. If faculty do not have access to Geneious, free software such as MEGAX can be used. Students were taught to visually inspect alignments for quality, reading frame consistency, and proper sequence overlap. These alignments served as the foundation for phylogenetic analyses and the calculation of various diversity metrics.

Calculating intra- and inter-specific distances

Understanding molecular variation within and between species is essential for phylogenetic inference and taxonomic placement. Students were instructed on how to calculate sequence identity (% identity) and to convert this into the number of polymorphisms between aligned sequences. *Quantitative Reasoning Activity 4* (QRA4, File S1) was developed to ensure competency across the cohort, regardless of their chosen project. This activity also included practice with interpreting and manipulating data matrices (File S1).

Phylogenetic methods

MSAs were used to construct phylogenetic trees within Geneious. Students explored both built-in tools and external plug-ins such as MrBayes v.3.2.6 (25) and RAxML (26) to expand their analytical capabilities. This software, as well as others such as MEGAX, is freely available for download with the use of Geneious Prime. The CURE emphasized comparing distance-based methods (e.g., neighbor-joining) with likelihood-based approaches (e.g., RAxML), encouraging students to consider the contexts in which each method is most appropriate. All phylogenetic analyses were evaluated statistically, and

TABLE 3 Example of independent analyses[a]

Examples of independent analyses	Skills and products developed
SPRING 2023	
Characterization of the mitochondrial genome Comparative mitochondrial genomics of related taxa	Genome assembly, gene annotation, and calculate summary statistics relevant to genomics
Molecular phylogeny based on mitochondrial genes Molecular phylogeny based on nuclear genes—typically nuclear rRNA genes (28S, 18S, and ITS1&2) Intraspecific variation among class samples	Utilizing genomic databases, building a multiple sequence alignment, construct phylogenetic trees, and generate pairwise genetic distance tables
SPRING 2024	
Evaluating the correlation between starting DNA quantity and quality and ONT results	Building a multivariable quantitative data set, linear regression, and generate regression plot
Molecular phylogeny based on nuclear genes—typically nuclear rRNA genes (28S, 18S, and ITS1&2) Measures of genetic relatedness of mitochondrial and rRNA genes	Utilizing genomic databases, building a multiple sequence alignment, construct phylogenetic trees, and generate pairwise genetic distance tables
Summarize ONT sequence data	Descriptive statistics, knowledge of measures of sequence quality, and create summary tables for presentation

[a]Following data collection, students were broken up into groups of two to three individuals to plan and execute data analysis. Each small group's project developed skills in line with a specific SLO and was assessed by a specific product.

students learned to interpret and communicate the significance of bootstrap support values.

Scientific communication and writing (Weeks 1–8)

Scientific communication and writing are core SLOs of this CURE (Table 1) and are integrated throughout the curriculum in multiple ways. To develop technical writing skills, students are asked during Weeks 1–4 to write a detailed paragraph after each lab session summarizing the wet-lab work they performed. Instructors provide targeted feedback on each submission, which students are expected to incorporate into their next weekly entry. Students continually revise their work in a shared Google Doc, treated as a "manuscript in progress," allowing for iterative development. By the end of Week 4, each student has a polished Methods section documenting their laboratory work, which they will later expand with Results and Discussion sections as their analyses progress.

To scaffold scientific reading and presentation skills, student pairs are assigned a relevant peer-reviewed article to use as a model for their writing and data analysis. In Spring 2023, students led informal 10-min discussions of their assigned paper during weekly class meetings. In Spring 2024, due to a larger class size, one class session (~60 min) was dedicated to a structured group discussion of two selected papers, which students read over spring break. In both iterations, we employed the "Figure Facts" approach (27), a data-centered pedagogy that helps undergraduates dissect primary literature by focusing on reproducing figures or tables with their own data. This method has been shown to improve student engagement, critical thinking, and confidence in interpreting scientific data.

Beginning in Week 5, students shift focus to their independent group projects, using the sequencing and bioinformatic data generated in Weeks 1–4. Each group selects a figure or table from the peer-reviewed literature to replicate or adapt using their own data set. As a culminating experience, students choose to present their work either as a poster at the Purchase College Natural Science Symposium or as an oral presentation during the final class session. Regardless of the format, all students contribute to the class-wide synthesis of data and the development of the final project delivera-

bles, ensuring that each independent project feeds into the comprehensive poster or presentation.

Quantitative reasoning activities

A commonly cited barrier to CURE implementation is concern over reduced content coverage, with some faculty fearing that essential course material may be sacrificed (15, 28). To combat this, four QRAs were developed and implemented throughout the CURE (File S1) to ensure that foundational quantitative and analytical skills were explicitly taught and assessed. These QRAs are designed to integrate seamlessly within the research framework of the course while reinforcing key concepts.

RESULTS

Outcomes of Weeks 1–3 wet-lab activities

In Spring 2023 (S23, n = 7) and Spring 2024 (S24, n = 14), undergraduate students (sophomore-senior) enrolled in the Genetics Laboratory course participated in the CURE described above. In both cohorts, each student performed MinION Oxford Nanopore DNA Sequencing on an individual adult trematode. Within each semester's cohort, all specimens were presumed to be conspecifics. In S23, students worked with trematodes from the family Diplostomatidae, collected from a single infected Northern Shoveler (*Spatula clypeata*). In S24, specimens were from the family Echinostomatidae, collected from a single infected Double-crested Cormorant (*Nannopterum auritum*). Students were provided only the family-level identification and were tasked with identifying their specimens to the lowest possible taxonomic level using the molecular genetic techniques described above. A trained learning assistant supported wet-lab procedures in both semesters.

In S23, all trematode samples (n = 7) yielded DNA suitable for sequencing and were successfully sequenced on a MinION ONT flow cell. This run generated a total of 103,469 reads, with an average of 14,766 reads per barcode (range: 1,090–43,831). The average read length was 803 bases.

In S24, of the 14 samples processed, 8 yielded sufficient DNA (≥200 ng in 10 µL) for library preparation and ONT sequencing. Sequencing of these samples produced an average of 58,286 reads per barcode (range: 32,000–68,000), with an average read length of 236 bases and an average of 6,779 assembled contigs per sample. Student-led analyses (Table 3) identified a statistically significant relationship between the initial DNA concentration and the average contig length produced (linear regression, P = 0.0279), reinforcing the importance of high-quality extractions in downstream sequencing success. Comparing the two semesters, the smaller student-to-faculty ratio in S23 likely contributed to the higher quality DNA extractions, and ultimately better sequence coverage.

Outcomes of Weeks 4–8 bioinformatics and molecular phylogenetics activities

During Weeks 4–8, students were guided through a *de novo* assembly of their individual, demultiplexed sequence data using Geneious. Each student generated hundreds to thousands of contigs, which were then queried using BLASTn to attempt sample identification. First, students were encouraged to BLAST individual contigs. Often, the reads with the best coverage were bacterial in origin. This often led to confusion, while we intentionally allowed students to experience and troubleshoot this as part of the learning process, instructors with limited time may choose to pre-filter contaminant reads before distributing data sets to students. Students were also shown how to perform a batch BLASTn search to obtain tentative identifications for a set of contigs.

Next, students were introduced to mining genomic data using appropriate reference genomes. Given the primary objective of identifying their trematode specimens to the lowest possible taxonomic level, students focused on extracting phylogenetically

informative sequences, including the mitochondrial genome and the nuclear ribosomal RNA (rRNA) cassette. To construct a mitochondrial reference database, students accessed the NCBI Organelle Database (https://www.ncbi.nlm.nih.gov/datasets/organelle/?taxon=2759) and downloaded mitochondrial genomes from related taxa. For the rRNA cassette, they used NCBI's Taxonomy Browser to identify and download complete rRNA sequences (5.8S, 18S, 28S, ITS1, and ITS2) from relevant taxa. Students then mapped their reads to these reference sequences using the "Map to Reference" tool in Geneious and recorded their results in a shared Google Sheet (percent identity and query coverage).

In S23, students recovered a nearly complete mitochondrial genome (126 reads) and corresponding nuclear rRNA sequences suitable for downstream phylogenetic analysis. In contrast, the S24 cohort recovered only a small number of mitochondrial reads (15 reads totaling 4,247 bases), though they achieved adequate coverage of the nuclear rRNA cassette. This is likely due to the lower quality and quantity of sequence data generated by this cohort. The mitochondrial reads obtained in S24 mapped to genes such as 12S, 16S, CO_2, ND1, and ND3, but notably, the CO1 gene—commonly used in trematode systematics—was not recovered. This limitation prompted valuable class discussions on marker choice, database completeness, and the challenges of species identification in non-model taxa.

Guided by assigned peer-reviewed literature, students identified analyses, tables, and figures relevant to the project goals. With faculty support, students formed small groups (two to three students) to develop and execute original analyses contributing to the class-wide research effort. Representative examples of these independent projects and their outcomes are summarized in Table 3.

Final projects and exit assessments

In both iterations, a subset of students (2–4) led the preparation and presentation of the class poster presentation at the Purchase College Natural and Social Science Symposium, held at the end of the Spring semester. All other students prepared a final oral presentation of their independent analyses, following the provided guidance. All participants were eligible to present posters and were self-selected volunteers. In both cohorts, the students who led the poster presentations demonstrated the highest levels of understanding for all aspects of the projects, measured by their final assessments. In Spring 2024, instead of in-class presentations, we opted for one-on-one "*exit assessments*" (File S3), where individual student groups gave their 5- to 10-min presentation to the instructor, followed by 15–20 min of discussion and questions relating to the SLOs (Table 1) and an assessment rubric (File S3). This allowed the instructor to gain significantly more information on student understanding and attitudes about the work, relative to the in-class presentations. Regarding wet-lab and bioinformatics SLOs assessed, all student groups performed at an "Exceeds (3)" or "Excellent (4)" level of understanding of all SLOs, with an average of 3.4 (n = 8 groups). However, SLOs pertaining to "Analytical and Fundamental Genetics Skills" were more bimodal. Depending on what skills were emphasized in their independent projects, some student groups demonstrated "Excellent" (4) understanding of phylogenetic and species-level analyses, while most student groups demonstrated a "Developing (1)" or "Competent (2)" level of understanding. This indicates that some independent projects address some SLOs more adequately than others. Regardless of independent projects, all student groups scored between 3 and 4 (3.6 average) for "applying statistical concepts in bioinformatics," "articulating the difference between sequence reads and contigs," and "selecting an appropriate reference genome." Designing independent projects to more evenly address all SLOs is an area where this CURE could be strengthened moving forward.

DISCUSSION

Overall, the described CURE offers a flexible framework for applying modern evolutionary genetics to questions of biodiversity and has the potential for broad integration into faculty research programs across diverse biological systems. By aligning the CURE with the instructor's specific research agenda, a positive feedback loop is established: faculty expertise, resources, and the substantial effort required to run a CURE contribute directly to advancing faculty research, which in turn continues to sustain and enrich the course. The structure, research questions, and focus of the CURE shift notably after Week 4, as students begin to interpret their sequence data and develop independent lines of inquiry. A comparison of the Spring 2023 and Spring 2024 cohorts illustrates the diverse directions these research questions can take (Table 3).

Fueling faculty research

The synergy between teaching and research is central to the success and sustainability of CUREs (28). By involving students in meaningful aspects of faculty research, their contributions can directly support and advance ongoing research programs (18, 19). This model can be especially impactful at primarily undergraduate institutions (PUIs), where research activity may slow during the academic year due to heavy teaching responsibilities. A recent meta-analysis by Wolfe and Steed (19) found that the likelihood of publishing data generated through CUREs significantly increases when the course is closely aligned with a faculty member's research agenda. As demonstrated in this course, faculty may intentionally structure a CURE to facilitate the collection of publishable data.

In the S23 offering, students successfully assembled a nearly complete mitochondrial genome for their focal trematode species. Using phylogenetic methods, they determined that their sample was ~97% identical to a previously published mitochondrial genome (MN834120, *Diplostomoidea* sp. MSB:PARA:30071), recovered from its intermediate host (Gastropoda: *Physella* sp.) in New Mexico. *Diplostomoidea* sp. MSB:PARA:30071 had only been identified to the family level; to further resolve its identity, students incorporated CO1 sequence data from NCBI to build a more taxonomically inclusive phylogenetic data set. Their analysis revealed that both the unknown worm and Diplostomoidea sp. MSB:PARA:30071 were conspecifics with a sequence from Alberta, Canada, which had been identified as a novel species within the genus *Australapatemon* (MH369793, *Australapatemon* sp. LN10; [29]). Based on these analyses, students concluded that their specimen represented a novel trematode species within the genus *Australapatemon*, identified its intermediate host, and documented its broad geographic range. The S23 CURE laid the groundwork for a formal species description and a forthcoming manuscript.

In the S24 iteration, students used recovered rRNA sequences and molecular systematics to identify their unknown specimen as *Drepanocephalus auritus* (Trematoda: Echinostomatidae). By analyzing intraspecific variation in the ITS1 and ITS2 regions across eight individuals, students inferred that their specimens were genetically identical, suggesting clonal transmission. They further mined sequence data from NCBI to compare their samples with *Drepanocephalus* specimens from other regions of the United States and Central America, thereby evaluating interspecific diversity. *D. auritus* and several of its congeners are distributed throughout the Americas and pose a significant threat to aquaculture (30), as the larval stage (metacercaria) encysts in fish tissues, causing substantial losses in farmed species such as catfish. These findings prompted discussions on the application of genomic data in disease tracking, diagnosis, and outbreak monitoring.

Within the described CURE, there are multiple mechanisms for the faculty to steer the project towards success and ensure that it results in publishable data. For example, two weeks prior to the implementation of the CURE, the trained learning assistant (or faculty) prepared and sequenced a conspecific helminth to generate a data set by which to crosscheck student sequence data. This also allowed the instructors to prepopulate datasheets and assign directed readings (based on initial taxonomic identifications). In

Week 4, as a quality control measure, students who were unsuccessful in obtaining usable data, either low quality or contaminated, "removed" their data from the class pool and were given the instructor data set to work with.

Well-aligned CUREs can also serve as a pipeline for recruiting students into faculty research programs. For example, data generated by students in both the S23 and S24 cohorts directly contributed to faculty-sponsored senior thesis projects, demonstrating the dual pedagogical and research benefits of this model. As a PUI, Purchase College does not have graduate Teaching Assistants; however, advanced undergraduates can serve as Learning Assistants. In both S23 and S24, we were able to select undergraduates trained in MinION sequencing and associated analyses to serve as learning assistants.

Bioinformatics is foundational in modern genetics and genomics

Bioinformatics is now a cornerstone of modern genetics, genomics, evolutionary biology, evolutionary medicine, and related disciplines (14). The ability to handle, manage, and analyze large-scale sequence data sets represents a high-impact, transferable skill set for today's biology graduates (14, 18). In an era defined by big data and high-throughput sequencing, students must develop an understanding of how genomic data can be mined to drive hypothesis generation—a perspective distinct from the traditional hypothesis-first approach. This CURE is designed to immerse students in an authentic, data-first experience of the scientific process, preparing them to engage with contemporary challenges in biological research. Faculty in research labs that do not generate novel sequence data can still apply the "dry-lab" aspects (Weeks 4–8) of this CURE, by employing published genomic data.

Pros and cons of ONT sequencing

One of the core goals of CUREs is to engage students in authentic scientific research while building relevant, real-world skills (18). DNA sequencing fits this aim but has traditionally required significant funding and reliance on external facilities (31). When sequencing is outsourced, students miss out on an essential part of the scientific process. ONT makes sequencing more accessible through its portable MinION device, which runs via a laptop and allows students to complete every step—from library preparation to sequencing—in-house (32). The hands-on nature of ONT sequencing, particularly library preparation and flow cell loading, gives students a tangible understanding of the workflow. With ONT's live basecalling feature, students can analyze and BLAST reads in real time, prompting immediate discussions about data quality and interpretation. Additionally, ONT produces long reads that are easier for students to handle bioinformatically (33).

Despite these advantages, ONT has limitations. It has higher error rates than Illumina or PacBio, and although bioinformatics can improve accuracy (34), results may vary with student skill levels. ONT also requires more input DNA and has higher per-base costs and lower throughput compared to short-read platforms (33). For example, in our CURE, low read depth prevented us from assembling a complete mitochondrial genome. However, the educational benefits of real-time, hands-on sequencing often outweigh these challenges, making ONT a valuable tool in undergraduate research courses.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation award 2218190 to E.T.E. and S.E.H.

We acknowledge the Purchase College Natural and Social Science staff and students who have been invaluable in facilitating this work.

AUTHOR AFFILIATION

[1]Department of Biology, The State University of New York, Purchase College, Purchase, New York, USA

AUTHOR ORCIDs

Erika T. Ebbs http://orcid.org/0000-0001-5987-9484

AUTHOR CONTRIBUTIONS

Erika T. Ebbs, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review and editing | Allyson Barredo, Conceptualization, Data curation, Project administration | Amelia Whitehurst, Conceptualization, Data curation, Project administration | Stephen E. Harris, Funding acquisition, Project administration, Resources, Writing – review and editing

ADDITIONAL FILES

The following material is available online.

Supplemental Material

Files S1 and S3 (jmbe00147-25-S0001.docx). Description of QRAs and exit interview assessment rubric.
File S2 (jmbe00147-25-S0002.xlsx). Pre-populated Google Sheets tables.

REFERENCES

1. Corwin LA, Graham MJ, Dolan EL. 2015. Modeling course-based undergraduate research experiences: an agenda for future research and evaluation. CBE Life Sci Educ 14:es1. https://doi.org/10.1187/cbe.14-10-0167
2. Brownell SE, Hekmat-Scafe DS, Singla V, Chandler Seawell P, Conklin Imam JF, Eddy SL, Stearns T, Cyert MS. 2015. A high-enrollment course-based undergraduate research experience improves student conceptions of scientific thinking and ability to interpret data. CBE Life Sci Educ 14:14: https://doi.org/10.1187/cbe.14-05-0092
3. Martin A, Rechs A, Landerholm T, McDonald K. 2021. Course-based undergraduate research experiences spanning two semesters of biology impact student self-efficacy but not future goals. J Coll Sci Teach 50:33–47. https://doi.org/10.1080/0047231X.2021.12290516
4. Stanfield E, Slown CD, Sedlacek Q, Worcester SE. 2022. A course-based undergraduate research experience (CURE) in biology: developing systems thinking through field experiences in restoration ecology. CBE Life Sci Educ 21:ar20. https://doi.org/10.1187/cbe.20-12-0300
5. Theobald EJ, Hill MJ, Tran E, Agrawal S, Arroyo EN, Behling S, Chambwe N, Cintrón DL, Cooper JD, Dunster G, et al. 2020. Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. Proc Natl Acad Sci USA 117:6476–6483. https://doi.org/10.1073/pnas.1916903117
6. Forte NR, Veasey VN, Christie BJ, Carter A, Hanks MA, Holderfield A, Houston T, Challa AK, Turner AN. 2023. Engaging students in a genetics course-based undergraduate research experience utilizing *Caenorhabditis elegans* in hybrid learning to explore human disease gene variants. J Microbiol Biol Educ 24:e00078-23. https://doi.org/10.1128/jmbe.00078-23
7. Lopatto D, Silver Key SC, Van Stry M, Siders J, Leung W, Sandlin KM, Rele CP, Genomics Education Partnership, Reed LK. 2023. Supporting the democratization of science during a pandemic: genomics course-based undergraduate research experiences (CUREs) as an effective remote learning strategy. J Microbiol Biol Educ 24:e00039-23. https://doi.org/10.1128/jmbe.00039-23
8. Rowland SL, Lawrie GA, Behrendorff JBYH, Gillam EMJ. 2012. Is the undergraduate research experience (URE) always best?: the power of choice in a bifurcated practical stream for a large introductory biochemistry class. Biochem Molecular Bio Educ 40:46–62. https://doi.org/10.1002/bmb.20576
9. Smith KPW, Waddell EA, Dean AN, Anandan S, Gurney S, Kabnick K, Little J, McDonald M, Mohan J, Marenda DR, Stanford JS. 2023. Course-based undergraduate research experiences are a viable approach to increase access to research experiences in biology. J Biol Educ 57:618–632. https://doi.org/10.1080/00219266.2021.1933135
10. Gunnels CW, Chastain J, Brunelle S, Carlin A, Cimarusti TM, Crone-Romanovski M, Coughlin RW, Culbertson C, Elek J, Felton A, Felton S, Giambo DA, et al. 2024. Undergraduate research experiences grow career-ready transferable skills. SPUR 8:15–25. https://doi.org/10.18833/spur/8/1/1
11. Norman L. 2023. Development and implementation of a bioinnovation focused course-based research experience for undergraduate students. Biomed Eng Educ 3:225–233. https://doi.org/10.1007/s43683-022-00099-8
12. Gao L, Guo M. 2023. A course-based undergraduate research experience for bioinformatics education in undergraduate students. Biochem Mol Biol Educ 51:189–199. https://doi.org/10.1002/bmb.21710
13. Yang MA, Korsnack K. 2024. Pairing a bioinformatics-focused course-based undergraduate research experience with specifications grading in an introductory biology classroom. Biol Methods Protoc 9:bpae013. https://doi.org/10.1093/biomethods/bpae013
14. Wilson Sayres MA, Hauser C, Sierk M, Robic S, Rosenwald AG, Smith TM, Triplett EW, Williams JJ, Dinsdale E, Morgan WR, et al. 2018. Bioinformatics core competencies for undergraduate life sciences education. PLoS One 13:e0196878. https://doi.org/10.1371/journal.pone.0196878
15. Govindan B, Pickett S, Riggs B. 2020. Fear of the CURE: a beginner's guide to overcoming barriers in creating a course-based undergraduate research experience. J Microbiol Biol Educ 21:21. https://doi.org/10.1128/jmbe.v21i2.2109
16. Genné-Bacon EA, Wilks J, Bascom-Slack C. 2020. Uncovering factors influencing instructors' decision process when considering implementation of a course-based research experience. CBE Life Sci Educ 19:ar13. https://doi.org/10.1187/cbe.19-10-0208
17. DeChenne-Peters SE, Scheuermann NL. 2022. Faculty experiences during the implementation of an introductory biology course-based undergraduate research experience (CURE). CBE Life Sci Educ 21:ar70. https://doi.org/10.1187/cbe.21-06-0154
18. McLaughlin JS, Coyle MS. 2016. Increasing authenticity & inquiry in the cell & molecular biology laboratory. Am Biol Teach 78:492–500. https://doi.org/10.1525/abt.2016.78.6.492
19. Wolfe AL, Steed PR. 2023. Generating publishable data from course-based undergraduate research experiences in chemistry. J Chem Educ 100:3419–3424. https://doi.org/10.1021/acs.jchemed.3c00354
20. Cribb TH, Barton DP, Blair D, Bott NJ, Bray RA, Corner RD, Cutmore SC, De Silva MLI, Duong B, Faltýnková A, et al. 2025. Challenges in the recognition of trematode species: consideration of hypotheses in an inexact science. J Helminthol 99:e54. https://doi.org/10.1017/S0022149X25000367

21. Carlson CJ, Dallas TA, Alexander LW, Phelan AL, Phillips AJ. 2020. What would it take to describe the global diversity of parasites? Proc Biol Sci 287:20201841. https://doi.org/10.1098/rspb.2020.1841
22. Douchet P, Boissier J, Mulero S, Ferté H, Doberva M, Allienne JF, Toulza E, Bethune K, Rey O. 2022. Make visible the invisible: optimized development of an environmental DNA metabarcoding tool for the characterization of trematode parasitic communities. Environmental DNA 4:627–641. https://doi.org/10.1002/edn3.273
23. Pinto HA. 2023. Describing formally larval trematodes: some reflections in the taxonomic integrative era. Trends Parasitol 39:889–890. https://doi.org/10.1016/j.pt.2023.07.006
24. Zieritz A, Yasaeng P, Razak NFA, Hongtrakul V, Kovitvadhi U, Kanchanaketu T. 2018. Development and evaluation of hotshot protocols for cost- and time-effective extraction of PCR-ready DNA from single freshwater mussel larvae (Bivalvia: Unionida). J Molluscan Stud 84:198–201. https://doi.org/10.1093/mollus/eyy011
25. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol 61:539–542. https://doi.org/10.1093/sysbio/sys029
26. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313. https://doi.org/10.1093/bioinformatics/btu033
27. Round JE, Campbell AM. 2013. Figure Facts: encouraging undergraduates to take a data-centered approach to reading primary literature. CBE Life Sci Educ 12:39–46. https://doi.org/10.1187/cbe.12-06-0085
28. Shortlidge EE, Bangera G, Brownell SE. 2016. Faculty perspectives on developing and teaching course-based undergraduate research experiences. Bioscience 66:54–62. https://doi.org/10.1093/biosci/biv167
29. Gordy MA, Hanington PC. 2019. A fine-scale phylogenetic assessment of digenean trematodes in central Alberta reveals we have yet to uncover their total diversity. Ecol Evol 9:3153–3238. https://doi.org/10.1002/ece3.4939
30. Alberson NR, Rosser TG, Buddenborg SK, Khoo LH, Loker ES, Richardson TD, Woodyard ET, Wise DJ, Pote LM, Griffin MJ. 2017. North and South American Haplotypes of *Drepanocephalus auritus* (Digenea: Echinostomatidae) are released from *Biomphalaria havanensis* (Mollusca: Planorbidae) inhabiting catfish aquaculture ponds in Mississippi, U.S.A. Comp Parasitol 84:87–101. https://doi.org/10.1654/1525-2647-84.2.87
31. Wray C. 2024. Classroom to career: implementation considerations for engaging students with meaningful DNA sequencing learning opportunities, p 137–155. In In Rigor and Reproducibility in Genetics and Genomics. Academic Press.
32. Williams J. 2025. There's never been a better time to be a STEM educator. Genetics 230:iyaf090. https://doi.org/10.1093/genetics/iyaf090
33. Zhang T, Li H, Jiang M, Hou H, Gao Y, Li Y, Wang F, Wang J, Peng K, Liu Y-X. 2024. Nanopore sequencing: flourishing in its teenage years. J Genet Genomics 51:1361–1374. https://doi.org/10.1016/j.jgg.2024.09.007
34. Lerminiaux N, Fakharuddin K, Mulvey MR, Mataseje L. 2024. Do we still need Illumina sequencing data? evaluating Oxford nanopore technologies R10.4.1 flow cells and the Rapid v14 library prep kit for Gram negative bacteria whole genome assemblies. Can J Microbiol 70:178–189. https://doi.org/10.1139/cjm-2023-0175