Microbiome education at under-resourced institutions: current status, barriers, and approaches to successful implementation

Supplementary Information

Contents

Vignettes of Microbiome Education Across Institutions

Skin Microbiome Extra Credit Assignment, Northern Virginia Community College

The Introductory General Biology course at Northern Virginia Community College is a popular laboratory science choice for students of all backgrounds. It fulfills a science laboratory course requirement for an associate's degree and transfer to a four-year college/university. This semester-long course focuses on cells and gene expression, with curriculum standardized by the Commonwealth of Virginia and established to create introductory course equivalency at all public higher education schools in Virginia. While genomics is only briefly introduced, the skin microbiome module based on NHGRI's Microbiome Virtual Lab Exploration [1] fits well into existing course material.

To bolster student participation, an extra credit assignment is offered as a group assignment and students collectively read a journal article, navigated NCBI's BLAST website [2], and ultimately identified the bacteria in a simulated unknown skin bacterium sample. Though the students did not perform isolation and sequencing, the activity allowed them for greater understanding of 16S bacterial signatures and the confidence with which they could identify a species. Ideally, future iterations of this assignment will include collecting skin samples and using Nanopore sequencing for students interested in the wet-lab components required to generate data.

Fruit fly gut microbiome module, Guttman Community College

The fruit fly gut microbiome module is an 8 week research module that bridges a 200-level microbiology class and a 200-level genetics course at Guttman Community College. The 200-level microbiology class is a capstone for the Science A.S. degree, with a writing-intensive requirement, and thus, lends itself well to independent research projects, as students have time to do sufficient literature searches and a culminating written journal article-style research report. Students use ethnographic family interviews to hypothesize which normal pantry staples might have probiotic effects. Using *Drosophila melanogaster* as a model organism, students replace yeast fly feed with an equivalent amount of the test food. From here the project is split into 2 avenues, the microbiology students smash the flies and plate the samples on MSR agar plates. The MSR plates allow them to quantify *Lactobacillus*, a beneficial bacterium in the fly gut. DNA extractions from the colonies on LB agar plates are purified and shared with the genetics research course. These samples will be analyzed by shotgun sequencing, using EPI2ME software to mine the data for *Lactobacillus*. Splitting the project between two classes, not only maximizes material and sample usage, but also introduces students to the real-life nature of science through collaboration, relying on previous research and making clear and concise data reports that can be followed by future researchers.

Monitoring Environmental Microbiome (MEM) workshop, Fort Lewis College

This week-long workshop offered in the summer engages undergraduate students on the effect of environmental pollution on sediment and water microbiomes [3]. Local EPA superfund sites are used to examine the effect of acid mine drainage, metal content and pH on the microbiome. Students develop an appreciation for the history of the sites and visit the collection sites. The Indigenous instructor uses a cultural integration and responsive curriculum as a driver for strong student engagement. Previously generated 16S amplicon data or newly generated sequence data is analyzed during the course of the week-long workshop which culminates with student presentations. While the key learning objectives focused on installation of tools or cloud computing (CyVerse VICE or Galaxy/ANVIL) and data analysis with QIIME 2 [4], different iterations of the workshop have emphasized data management, data types, algorithms, documenting (Jupyter Notebook), circling back to research question, interfaces of cloud computing, visualization or interpretation of results. Implementation of the workshop required access to computing power (memory intensive) freely available through CyVerse and molecular biology reagents and sequencing services whenever these were integrated in the design of the workshop. The workshop is currently being redeveloped as a Biochemistry course-based undergraduate research experience (CURE).

Introduction to Bioinformatics and Data Science, University of Puerto Rico–Ponce

The Introduction to Bioinformatics and Data Science course is offered as a regular semester long course or as a 4-week intensive course during winter break through the RISE program at the University of Puerto Rico-Ponce. The first section of the course provides an introduction to statistical analysis and programming using R. This foundational knowledge allows students to dive into bioinformatics and microbiome analysis. First, students construct microbiome taxonomy and abundance tables from cleaned data via QIIME 2 [4]. With these tables and metadata in hand, students use MicrobiomeAnalyst (https://www.microbiomeanalyst.ca/) to test hypotheses, including drawing connections between taxonomic abundance and variables of interest, such as BMI, smoking status, alcohol consumption, and family risk factors. Students work together to compare their results. To supplement expertise in microbiome data analysis, the course was developed with a collaborator at a different institution (George Washington University).

Microbiomes CURE, Spelman College

Bio334-Microbiomes is a research-based course designed to provide both hands-on experiences and computational experiences.The course is intended to be a gateway to upper-level undergraduate research courses offered by several faculty members in the department. The study site, the Spelman Gardens, was established over a hundred years ago and served as Victory Gardens during World War I and II. As students learn about the history of the site and acquire background knowledge on microbiomes, they start to develop and refine their research questions.

Typically, their investigations involve two phases. The first phase involves the isolation of bacteria followed by DNA extraction, library preparation and Nanopore sequencing. Bioinformatic pipelines are run on Galaxy (genome assembly with Flye, genome polishing with Racon, species identification with Kraken2, and search for Biosynthetic Gene Clusters with antiSMASH) [5];[6];[7];[8];[9]. The second phase involves the analysis of metagenomic data and assessment of bacterial interactions (alpha and beta diversities, taxonomic analysis and data visualization with R, and cross-streak assays). Thus, during the semester, students progress from 'think microbe' to 'think microbiome'. The lectures, journal club style presentations and hands-on activities were supplemented by external speakers.

The use of Nanopore sequencing required access to funds for reagents and supplies. In addition, the implementation of this CURE required some anticipation and planning in case of technical difficulties, especially during DNA extraction (e.g. lengthy protocol, low yield). These can be circumvented by either providing open lab hours or by providing back-up samples. Overall, the continuum between the research site, the research questions, the hands-on activities, the bioinformatic and statistical analysis provided for an integrated learning experience. Such an experience was particularly viewed by the faculty as important for re-engaging students post-covid.

Soil and termite gut microbiomes for capstone research course, Virginia State University

All biology major undergraduate students at the Virginia State University are required to complete one semester of research through the capstone Investigations and Research course (BIOL446). On average at least one section of this course each year has been developed for soil and termite gut microbiomes study. This course introduces the whole process from sample collections to data analysis. For the project, students start with termite collections in their natural habitats in easily accessible areas and soil sample collections at the same sites where termites are collected in order to study the relationship between termite gut microbiomes and their soil environment microbiomes. In the lab, students would dissect out termite guts and extract microbial genomic DNAs and in parallel, the soil genomic DNAs. After evaluating the quality of gDNA extractions, the students would continue with PCR amplification of the 16S rRNA marker and proceed with library preparation for paired-end sequencing with Illumina MiSeq. Upon the collection of the raw sequencing data, the students would demultiplex and denoise the data and characterize the taxonomic compositions, calculate both alpha and beta diversity, and conduct differential abundance analysis mostly using QIIME2 and RDPTools [4];[10] in a Linux command-line environment. Microbial taxonomy identification would also be visualized in Krona [11]. With metagenomic data generated from in-house Nanopore sequencers,

students would conduct taxonomic classification with kraken2 [9] and R packages, metagenomic assembly using metaFlye [12], upon which genes involved in cellulose degradation are identified using FragGeneScan [13], the CAZy database [14], hmmscan [15], and other programs.

BioDIGS miniCURE, Notre Dame of Maryland University

The BioDIGS miniCURE is a 5-week modular course-based undergraduate research experience embedded in the BIO340 Microbiology Lab at Notre Dame of Maryland University. The miniCURE introduces students to whole genome metagenomics data generated by the BioDIGS consortium (biodigs.org) which researches soil biodiversity and ecological factors that impact human health and ecological concerns. Students analyze the microbiome data for microbial diversity and screen for antibiotic resistance genes using Galaxy [7] with established tools, such as Kraken2 [9] and ABRicate [16]. A mixture of junior and senior students typically take this class as part of the Biology major curriculum with no coding experience required. Students who complete this work create an independent project poster in class that may use more BioDIGS data and/or other public metagenomics data from the NCBI Sequence Read Archive [17].

Basic Poll Information

This poll was launched in May 2024 to better understand microbiome education, specifically focusing on the challenges and opportunities faced by instructors. The purpose of this poll was to gather insights and data that will support our efforts to promote equal access to microbiome training and resources.

The first poll response was collected May 20, 2024. A blank version of the poll can be viewed here. All code associated with this project can be found on GitHub.

Response Rate

The poll had 75 respondents, for a response rate of 25.77%.

Institution Types

name	n	percent
Primarily undergraduate institution	27	36%
R1 institution	16	21%
Community college	15	20%
R2 institution	7	9%
Master's degree granting institution	6	8%
Unknown	4	5%

| name	n	percent
Hispanic Serving Institution	11	15%
Historically Black College or University	8	11%
Other under-resourced institution	6	8%
Asian American and Native American Pacific Islander Serving Institution	5	7%
Tribal College or University	2	3%

Responses: General Poll Questions

Are you familiar with microbiome research and its potential applications in public health, medicine, the environment, and industry?
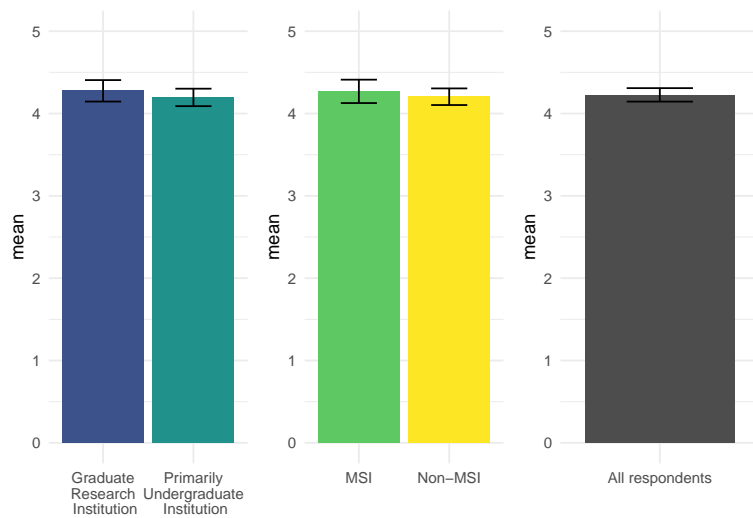
Have you conducted microbiome research, previously or currently

FALSE   TRUE
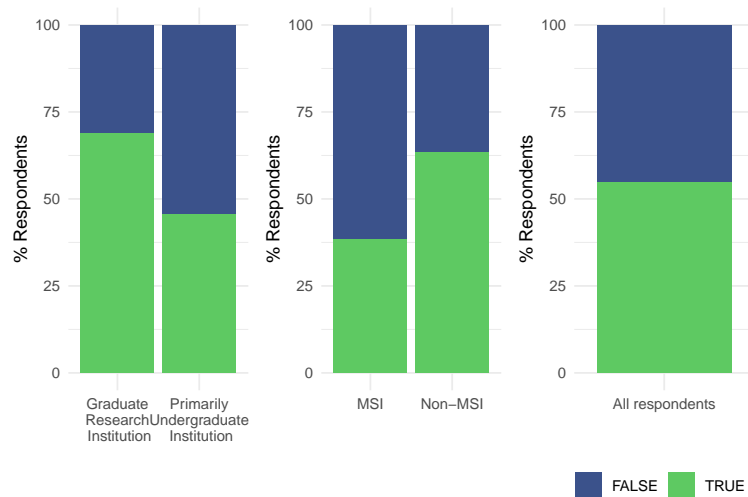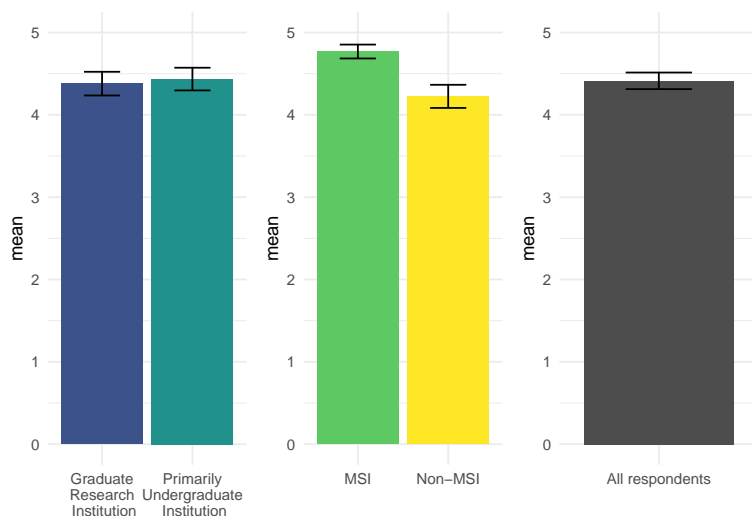
Do you, or others at your institution, offer specific courses or educational programs for students related to microbiomes?
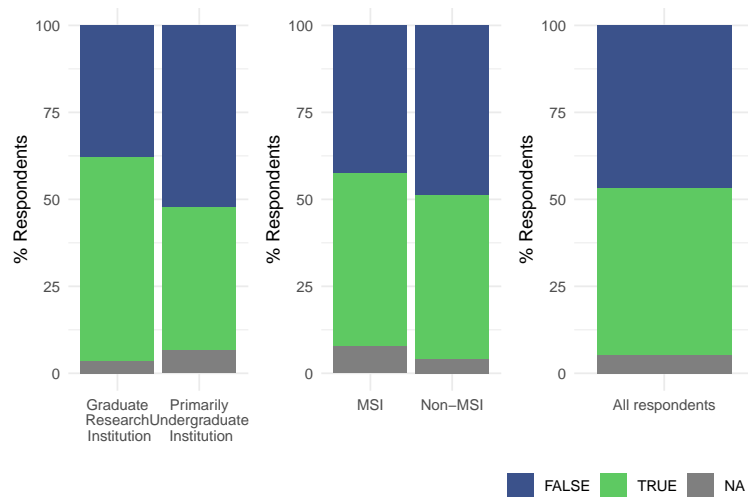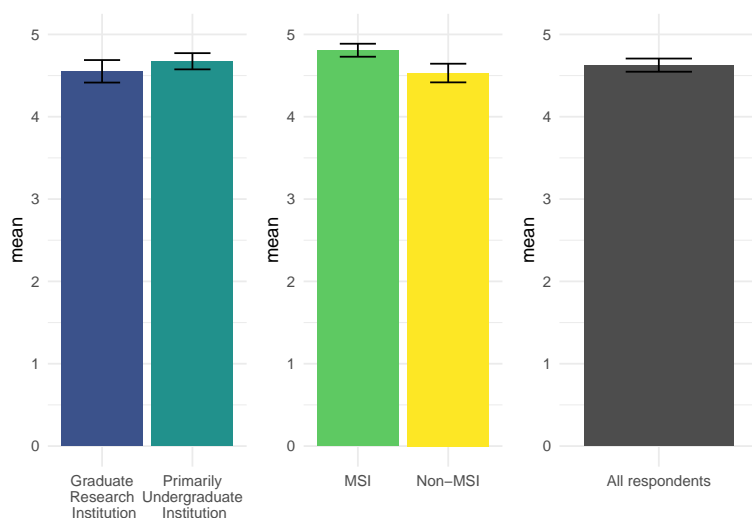

FALSE   TRUE   NA

Would you be interested in collaborating with other institutions, organizations, or faculty to

enhance microbiome education or research opportunities at your institution?

Do you believe that incorporating microbiome topics into the curriculum can provide students with interdisciplinary skills to address real-world problems?

Responses: Institutional Barriers

Institutional Barriers (separated)

The following are visualizations of institution types by each barrier. Note that research institution status and MSI status are not mutually exclusive. For example, a research institution could be either non-MSI or MSI.

When implementing or considering implementing a module or course, please indicate the extent of each of the following challenges, where 1 is "not a challenge" and 5 is "a strong challenge".

Confidence and/or expertise for teaching the material
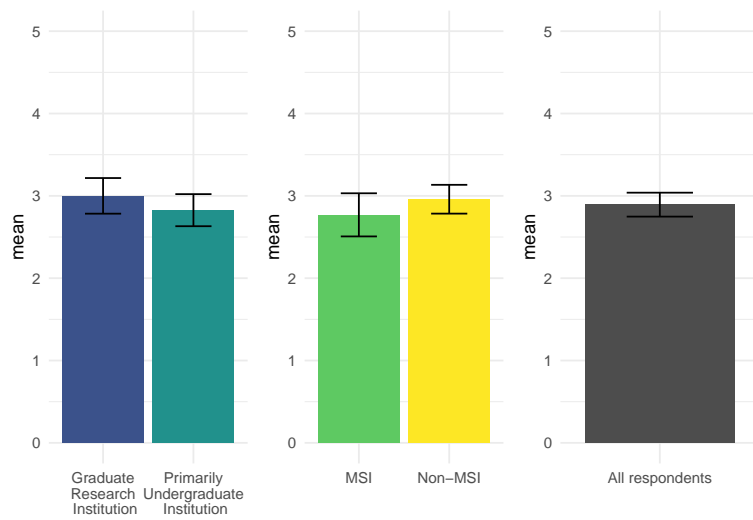
Opportunities for development and/or access to a community of practice
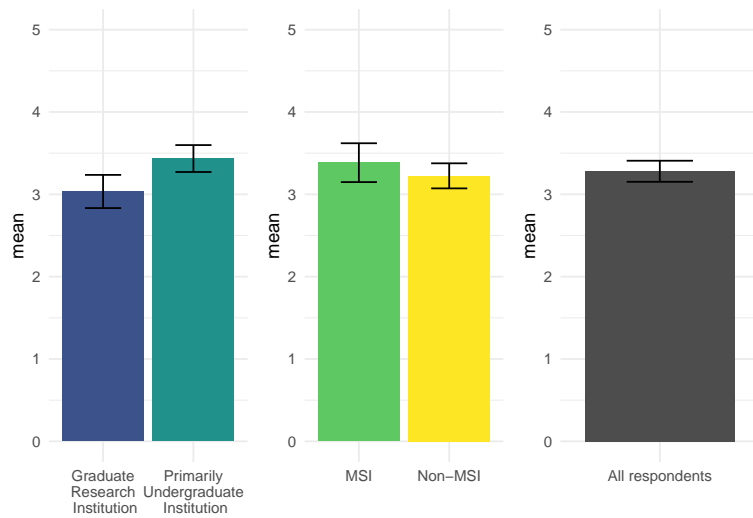
Access to computers and/or computational time

Access to laboratory supplies to incorporate wet lab activities
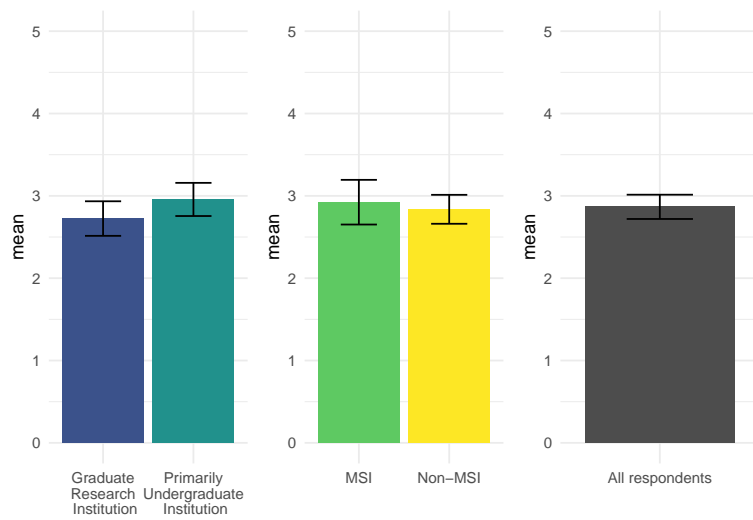
Time and/or financial support needed to design and/or implement courses
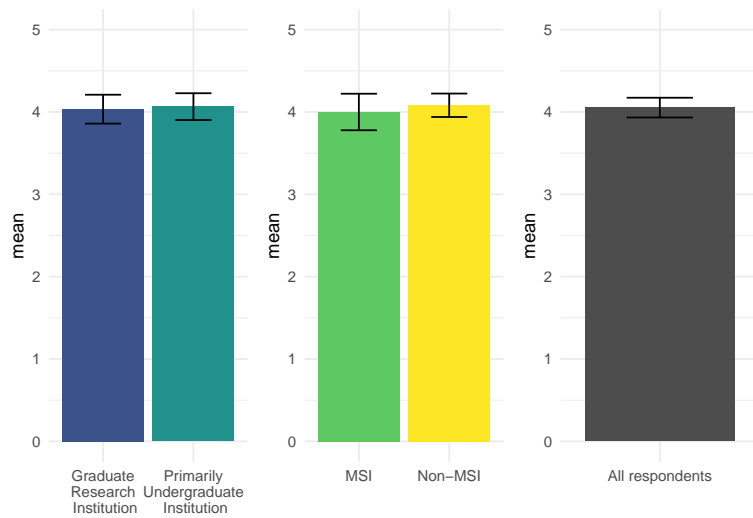
Access to existing course materials
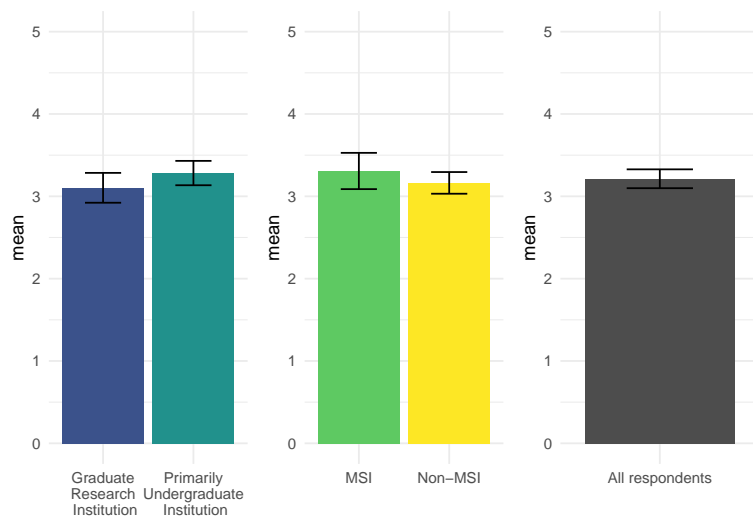
Freedom and/or flexibility to implement a course or module of your choosing
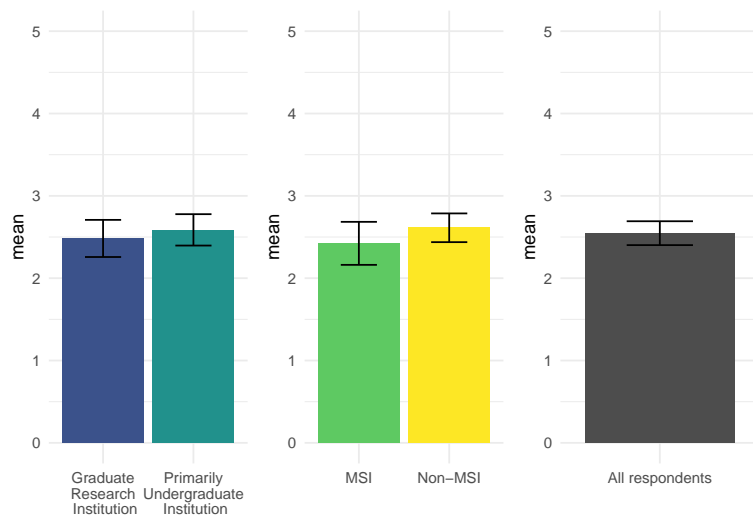
Student interest and/or engagement
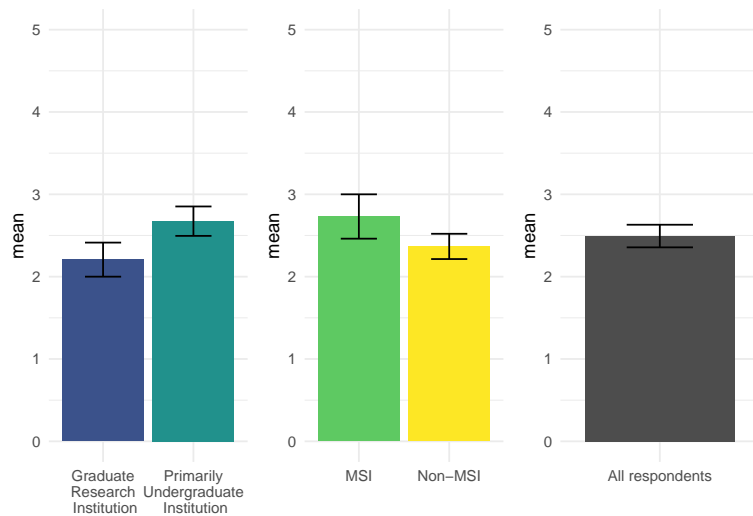
Institutional Barriers (aggregated)

When implementing or considering implementing a module or course, please indicate the extent of each of the following challenges, where 1 is "not a challenge" and 5 is "a strong challenge".
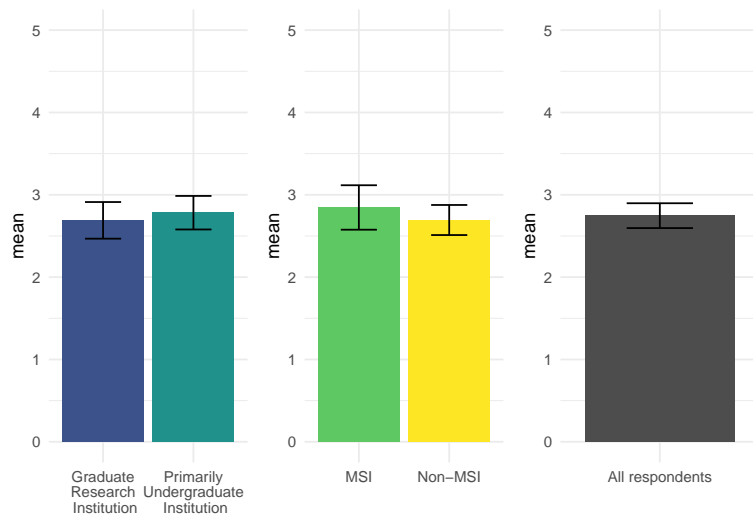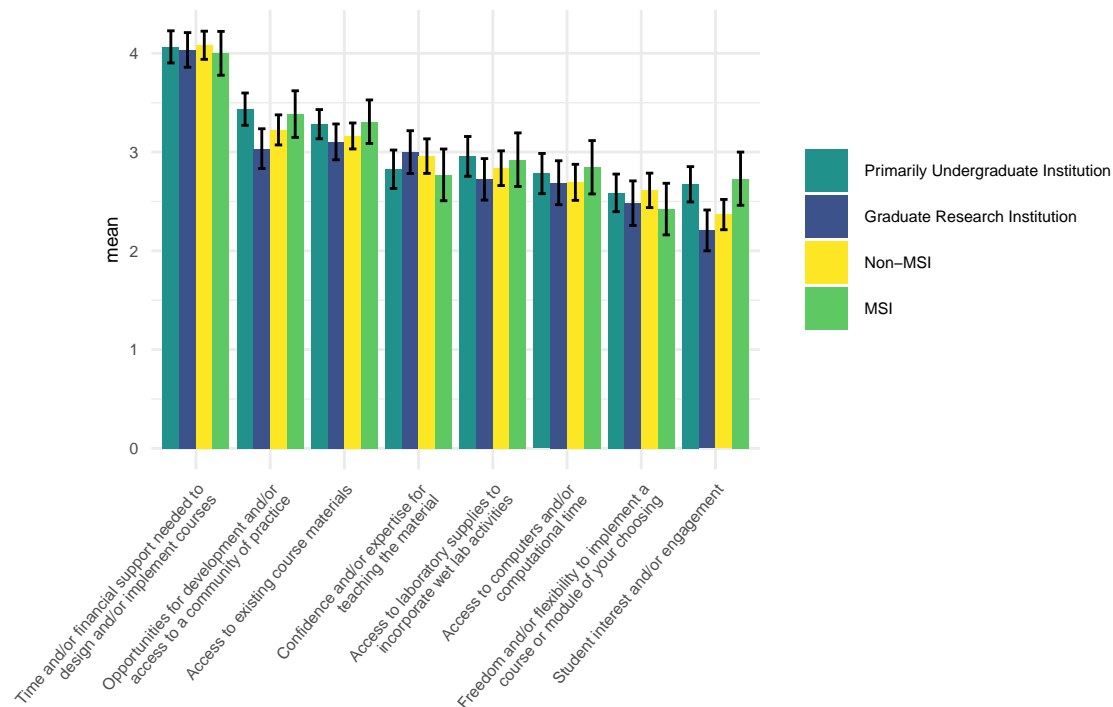
Statistics

Would you be interested in collaborating with other institutions, organizations, or faculty to enhance microbiome education or research opportunities at your institution?

Based on the outcome below, MSI faculty rank interest in collaborating higher.

```
##
##  Welch Two Sample t-test
##
## data:  t_msi and t_nonmsi
## t = 3.3223, df = 71.063, p-value = 0.001413
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.2178121 0.8716698
## sample estimates:
## mean of x mean of y
##  4.769231  4.224490
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  t_msi and t_nonmsi
## W = 832, p-value = 0.01326
## alternative hypothesis: true location shift is not equal to 0
```

Logistic Regression of Barriers

The following is a poisson regression model looking at how research type ("Primarily Undergraduate Institution" vs. "Graduate Research Institution") and is/is not an MSI affect the perception on different barriers to

implementing microbiome work.

Based on the outcome below, there isn't evidence that research type / MSI affects these perceived barriers. However, "Time and/or financial support needed to design and/or implement courses" is the biggest barrier for all institutions.

```
##
## Call:
## glm(formula = value ~ name + research_type + msi_type, family = poisson(link = "log"),
##     data = stats_dat)
##
## Coefficients:
##                        Estimate Std. Error z value Pr(>|z|)
## (Intercept)            1.029124   0.075317  13.664  < 2e-16 ***
## namecourse_freedom    -0.075603   0.100448  -0.753   0.4517
## namedev_opportunities  0.177455   0.094442   1.879   0.0602 .
## namelab_access         0.042762   0.097496   0.439   0.6609
## namematerials_access   0.156921   0.094888   1.654   0.0982 .
## nameself_confidence    0.052021   0.097276   0.535   0.5928
## namestudent_interest  -0.096768   0.101004  -0.958   0.3380
## nametime_financial     0.389152   0.090243   4.312 1.62e-05 ***
## research_typeTRUE     -0.054411   0.049987  -1.088   0.2764
## msi_typeTRUE           0.005577   0.050669   0.110   0.9123
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 345.51  on 599  degrees of freedom
## Residual deviance: 301.61  on 590  degrees of freedom
## AIC: 2060.7
##
## Number of Fisher Scoring iterations: 4
```

Free Text Thematic Analysis

We used a qualitative research approach called thematic analysis (Braun and Clarke 2006) to better understand free text responses. This approach has previously been used to better understand instructor perspectives. For example, it has been used to understand instructor perspectives on the use of LLMs in the classroom (link), blended learning (link), professional development opportunities (link), and threshold concepts (link). It has been used extensively to better understand STEM education (link, link, link).

Barrier Thematic Analysis

"Please provide any additional thoughts on barriers you have encountered with microbiome education."

Overall Barrier Themes

There might be some disagreement about what constitutes microbiome education at different institution types. For example, a CURE at a non-research institution might leverage existing data, while a research institution might feel that expensive original data acquisition is required.

- Sequencing was mentioned repeatedly for non-MSIs but not at MSIs
- Costs were mentioned at research institutions but not at undergraduate institutions.
- CUREs were mentioned 6 times at non-research MSI institutions, but not at all at research institutions / non-MSI institutions.

Themes for barriers at MSIs

1. Curriculum integration and time constraints

- "We are a two-year college offering lower-division undergraduate coursework, so it can be challenging to incorporate advanced topics and areas of active research into the curriculum."
- "Integrating the topic into existing learning outcomes of a course. For our institution, we won't be able to offer a specific course on microbiome, but I can see several courses into which microbiome lessons can be integrated."
- "Primarily the time to research it and the availability of teaching resources"
- "may need some training and lab equipment to conduct such lab module"
- "Access to researchers with correct expertise and TIME to collaborate"
- "Computational and statistical expertise to support microbiomes studies is lacking at our institution. We have tried reaching out to computer science and statistics faculty but there has been very limited interest and/or experience in this specific application to microbiomes data."

2. Student engagement and background

- "Student understanding of technologies; lack of resources to incorporate activities in student labs. Lack of student preparedness for lower-level microbiology course."
- "Students are interested when microbiome research is discussed but often lack knowledge"
- "Bioinformatic flows are difficult for students to learn. Thus, they become frustrated. Many of my students state they prefer the wet-bench portion of metagenomics."
- "It is hard to get students excited or engaged in anything. However, students were more interested when they were presented with examples, articles, social media clickbait of how it relates to them and their health."

Themes for barriers at non-MSIs

1. Cost barriers for new sequencing

- "Money and time and training are the big three"
- "Lack of funding for sequencing"
- "Lack of funds for expensive reagents and whole genome sequencing."
- "Cost per sample and turnaround time for receiving data if using an external vendor could be barriers for having students do the DNA sequencing for microbial communities in class."
- "Cost and resources are a really big factors. I can teach all day about the theory behind microbiome concepts and techniques, but doing and practicing the techniques can be quite difficult."
- "Cost for supplies and sequencing if generating data and updating computers for analysis are both challenges at my institution."
- "Paying for sequencing and knowing how to do the bioinformatics"
- "They are listed above - finances and equipment"
- "It is too expensive"

2. Instructor skills and confidence

- "My own ignorance and weaknesss is a big roadblock."
- "Some challenges are library prep and makes sense out [of] .. output."
- "I would love hands on training and a network. It would give me to confidence I need to create a whole class around it rather than just a module or assignment."
- "As an instructor, I lack some of the statistical background to explain in detail when community compositions are different. I would love to have (and contribute to) resources available for this."
- "Paying for sequencing and knowing how to do the bioinformatics"

Themes for barriers at non-research institutions

These themes overlap with MSIs (above) including lack of student background, flexibility to implement microbiome content, and time/expertise constraints.

Themes for barriers at research institutions

Costs and time, either from the instructor themselves or lack of personnel or collaborators.

Other Thoughts Thematic Analysis

"Please provide any additional thoughts on your experience with microbiome education."

Overall Thoughts Themes

There were fewer divisions among types of institutions. In general, respondents felt a lot of enthusiasm for the topic and for collaboration. There is some concern around how to implement microbiome content in the classroom, as well as expertise and time to make that happen.

Other Thoughts Themes at MSIs

1. Enthusiasm for the topic

- "Very important and ties into many content areas, there are lots of misconceptions on the part of the students."
- "I do think this is an outstanding topic with tons of potential and it should be part of the articulation agreements."
- "Microbiome studies are a particularly good subject for learning the interdisciplinary nature of biology and integrating molecular biology, microbiology, ecology and environmental sciences. It is an area rich in unanswered questions so it is perfect for CURE learning and teaching."
- "It is a huge passion of mine that I always include in my courses but never as a "module" so I would be very excited to see collaboration"
- "I do discuss in my courses, and would like to implement it more"

2. Enthusiasm for collaboration

- "My colleague (a chemist/environmental scientist who studies heavy-metal pollution) and I have been developing a collaboration to study microbes living in mine water drainage remediation systems. We are interested in incorporating this project into our courses."
- "I am excited to collaborate with researchers and educators at other institutions to receive training/support that will advance microbiome studies at our campus"
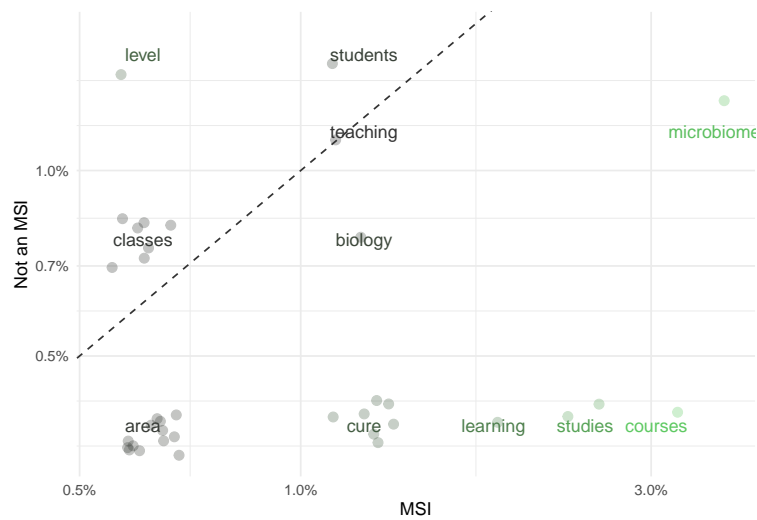- "The desire is there, finding Microbiome experts has been challenging"

Other Thoughts Themes at non-MSIs

1. Enthusiasm for the topic and partnering with others

- "My Introduction to Microbiology students are very interested in the complexities and presence of microbial communities! This is very much an interesting and important area of research and I already have plans to start developing a class focused in this direction."
- "I just briefly do some case studies and discussions about microbiome work, but I would love to do more"
- "I have been working on a collaborative project with my colleague in environmental to study microbial communities in acid mine drainage remediation systems, which is a project we are interested in building on."
- "This is a topic that is very important for me to teach and I incorporate it into my nutrition and biodiversity classes. Ideas for ways to teach it are always welcome."
- "REMnet and Microbiomes for all has been a big help."

2. More planning is needed in how to integrate microbiome topics appropriately

- "It's very powerful and interesting, but I believe it is better suited to higher level courses than first -year general biology"
- "I'm still unsure how to design enduring learning outcomes at the undergrad level regarding microbiomes."
- "It is tricky to know how to make space in the curriculum"

3. Lack of time to learn or implement

- "I'd love to jump back into iPython/Jupyter notebooks but that requires time to develop I don't really have in my heavy teaching load position."
- "time to teach something like R and computing power/software accessibility have prevented me from trying to implement this level of analysis in my classes."
- "If instructors weren't trained during grad or post-doc it is harder to find the right place for training"

Wordclouds

"Please provide any additional thoughts on barriers you have encountered with microbiome education."

"Please provide any additional thoughts on your experience with microbiome education."

Word Frequencies: MSIs versus Non-MSIs

Below shows the differences between frequency of certain words. Words farther away from the midpoint are more commonly found in that institution type.

The word pool is the following question.

"Please provide any additional thoughts on barriers you have encountered with microbiome education."
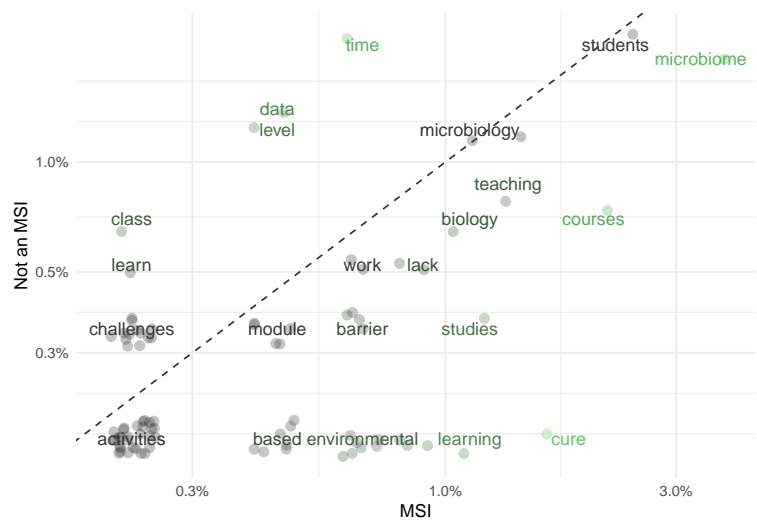
The word pool is the following question.

"Please provide any additional thoughts on your experience with microbiome education."

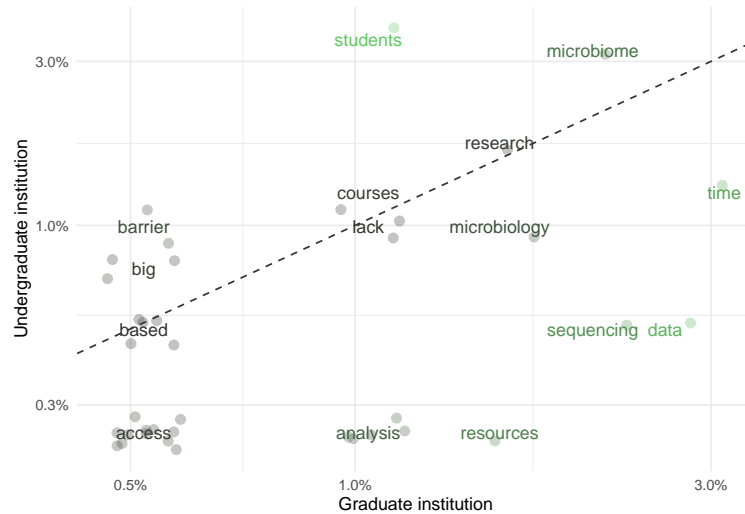The word pool is both questions above.

Word Frequencies: Graduate institutions vs. undergraduate institutions

Below shows the differences between frequency of certain words. Words farther away from the midpoint are more commonly found in that institution type.
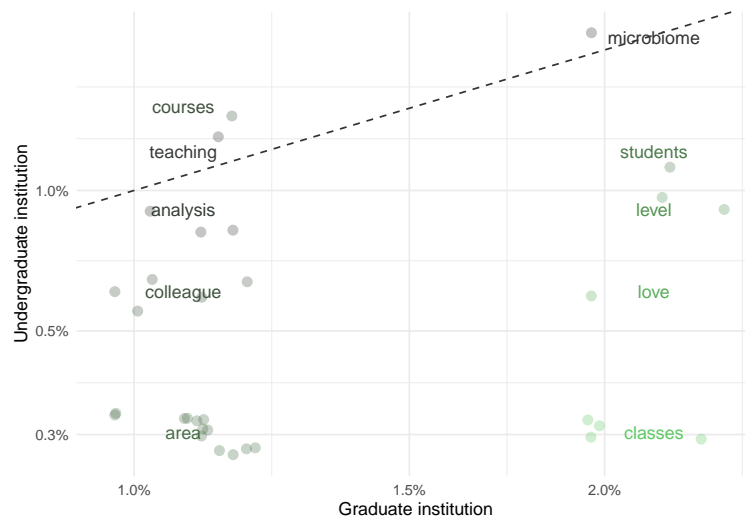
The word pool is the following question.

"Please provide any additional thoughts on barriers you have encountered with microbiome education."
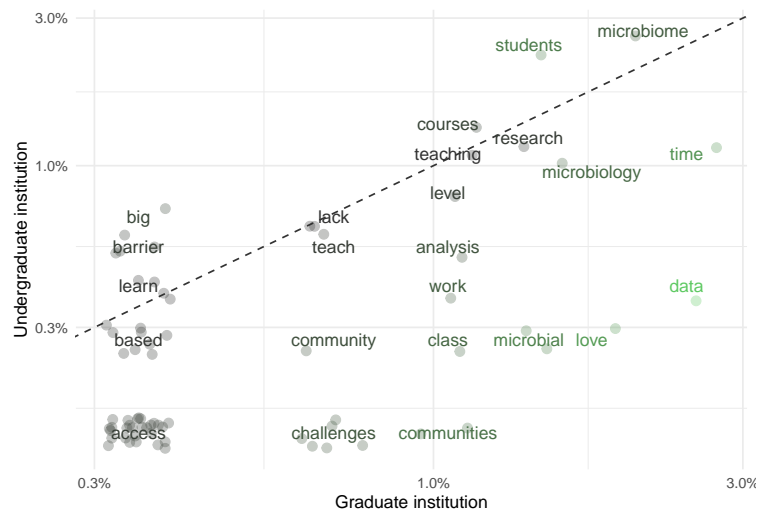

The word pool is the following question.

"Please provide any additional thoughts on your experience with microbiome education."

The word pool is both questions above.

Figures for the publication

Please see the repository: https://github.com/fhdsl/microbiome-poll for rendered figures.

System Settings

```
## R version 4.4.2 (2024-10-31)
## Platform: aarch64-apple-darwin20
## Running under: macOS Sonoma 14.4.1
##
## Matrix products: default
## BLAS:   /Library/Frameworks/R.framework/Versions/4.4-arm64/Resources/lib/libRblas.0.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/4.4-arm64/Resources/lib/libRlapack.dylib;  LAPACK ve
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## time zone: America/New_York
## tzcode source: internal
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
##  [1] lubridate_1.9.3    forcats_1.0.0      stringr_1.5.1
##  [4] dplyr_1.1.4        purrr_1.0.2        readr_2.1.5
##  [7] tidyr_1.3.1        tibble_3.2.1       ggplot2_4.0.0
## [10] tidyverse_2.0.0    scales_1.4.0       SnowballC_0.7.1
## [13] wordcloud_2.6      RColorBrewer_1.1-3 tidytext_0.4.2
## [16] patchwork_1.2.0    googlesheets4_1.1.1
##
## loaded via a namespace (and not attached):
##  [1] gtable_0.3.6      xfun_0.52         gargle_1.5.2       lattice_0.22-6
##  [5] tzdb_0.4.0        vctrs_0.6.5       tools_4.4.2        generics_0.1.3
##  [9] curl_5.2.1        fansi_1.0.6       janeaustenr_1.0.0  pkgconfig_2.0.3
## [13] tokenizers_0.3.0  Matrix_1.7-1      S7_0.2.0           lifecycle_1.0.4
```

```
## [17] compiler_4.4.2    farver_2.1.2     tinytex_0.51      htmltools_0.5.8.1
## [21] yaml_2.3.10       pillar_1.9.0     openssl_2.2.0     tidyselect_1.2.1
## [25] digest_0.6.35     stringi_1.8.4    labeling_0.4.3    fastmap_1.2.0
## [29] grid_4.4.2        cli_3.6.3        magrittr_2.0.3    dichromat_2.0-0.1
## [33] utf8_1.2.4        withr_3.0.2      rappdirs_0.3.3    googledrive_2.1.1
## [37] timechange_0.3.0  rmarkdown_2.27   httr_1.4.7        cellranger_1.1.0
## [41] askpass_1.2.0     hms_1.1.3        kableExtra_1.4.0  evaluate_1.0.5
## [45] knitr_1.47        viridisLite_0.4.2 rlang_1.1.4       Rcpp_1.0.12
## [49] glue_1.8.0        xml2_1.3.6       svglite_2.1.3     rstudioapi_0.16.0
## [53] jsonlite_1.8.8    R6_2.5.1         systemfonts_1.1.0 fs_1.6.4
```

References

1. Fuller KS, Lee A, Farr J, Conlan S, Kazmi W. Microbiome Lesson Plans. 2021. Available: https://www.genome.gov/careers-training/Professional-Development-Programs/NHGRI-Short-Course-in-Genomics/Microbiome-Lesson-Plans

2. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. Journal of Molecular Biology. 1990;215: 403–410. doi:10.1016/S0022-2836(05)80360-2

3. Lee JS, Lowell JL, Whitewater K, Roane TM, Miller CS, Chan AP, et al. Monitoring environmental microbiomes: Alignment of microbiology and computational biology competencies within a culturally integrated curriculum and research framework. Molecular Ecology Resources. 2025;25: e13867. doi:10.1111/1755-0998.13867

4. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature Biotechnology. 2019;37: 852–857. doi:10.1038/s41587-019-0209-9

5. Blin K, Shaw S, Vader L, Szenei J, Reitz ZL, Augustijn HE, et al. antiSMASH 8.0: Extended gene cluster detection capabilities and analyses of chemistry, enzymology, and regulation. Nucleic Acids Research. 2025; gkaf334. doi:10.1093/nar/gkaf334

6. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. Nature Biotechnology. 2019;37: 540–546. doi:10.1038/s41587-019-0072-8

7. The Galaxy Community, Abueg LAL, Afgan E, Allart O, Awan AH, Bacon WA, et al. The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update. Nucleic Acids Research. 2024;52: W83–W94. doi:10.1093/nar/gkae410

8. Vaser R, Sović I, Nagarajan N, Šikić M. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Research. 2017;27: 737–746. doi:10.1101/gr.214270.116

9. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. Genome Biology. 2019;20: 257. doi:10.1186/s13059-019-1891-0

10. Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, et al. Ribosomal Database Project: Data and tools for high throughput rRNA analysis. Nucleic Acids Research. 2014;42: D633–D642. doi:10.1093/nar/gkt1244

11. Ondov BD, Bergman NH, Phillippy AM. Interactive metagenomic visualization in a Web browser. BMC Bioinformatics. 2011;12: 385. doi:10.1186/1471-2105-12-385

12. Kolmogorov M, Bickhart DM, Behsaz B, Gurevich A, Rayko M, Shin SB, et al. metaFlye: Scalable long-read metagenome assembly using repeat graphs. Nature Methods. 2020;17: 1103–1110. doi:10.1038/s41592-020-00971-x

13. Rho M, Tang H, Ye Y. FragGeneScan: Predicting genes in short and error-prone reads. Nucleic Acids Research. 2010;38: e191–e191. doi:10.1093/nar/gkq747

14. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Research. 2014;42: D490–D495. doi:10.1093/nar/gkt1178

15. Finn RD, Clements J, Eddy SR. HMMER web server: Interactive sequence similarity searching. Nucleic Acids Research. 2011;39: W29–W37. doi:10.1093/nar/gkr367

16. Seemann T. ABRicate. 2020. Available: https://github.com/tseemann/abricate

17. Leinonen R, Sugawara H, Shumway M. The Sequence Read Archive. Nucleic Acids Research. 2011;39: D19–D21. doi:10.1093/nar/gkq1019