Methylothon: a versatile course-based high school research experience in microbiology and bioinformatics-- with pink bacteria

Peyton A. Jones[1], David Frischer[2], Shannon Mueller[3], Thomas Le[4], Anya Schwanes[5], Alekhya Govindaraju[6], Katie Shalvarjian[6], Jean-Baptiste Leducq[7], Christopher J. Marx[7], N. Cecilia Martinez-Gomez[8], and Jessica A. Lee[9,10]*

1 Integrative Biology Program, Harvard University, Cambridge, MA
2 Abraham Lincoln High School, San Francisco, CA
3 Berkeley High School, Berkeley, CA
4 Molecular and Cell Biology Program, University of California-Berkeley, Berkeley, CA
5 Galileo Academy of Science and Technology, San Francisco, CA
6 Plant and Microbial Biology Program, University of California-Berkeley, Berkeley, CA
7 Department of Biological Sciences, University of Idaho, Moscow, ID
8 Department of Plant and Microbial Biology, University of California-Berkeley, Berkeley, CA
9 Laboratory for Research in Complex Systems, San Francisco, CA
10 Current address: Space Biosciences Research Branch, NASA Ames Research Center, Moffett Field, CA
* jessica.audrey.lee@gmail.com

Support for the present work: National Science Foundation Division of Environmental Biology, Dimensions of Biodiversity program (award 1831838)

Running head: Methylothon: pink bacteria in education

Number of figures and Tables: 4 figures and 3 tables

Conflict of interest: The authors declare no conflict of interest.

ABSTRACT

Methylothon is an inquiry-based high school learning module in microbial ecology, molecular biology, and bioinformatics that centers around pink-pigmented plant-associated methylotrophic bacteria. Here we present an overview of the module's learning goals, describe course resources (available for public use on http://methylothon.com), and relate lessons learned from adapting Methylothon for remote learning during the pandemic in spring of 2021. This curriculum description is intended not only for instructors but also for microbial ecology researchers with an interest in conducting K-12 outreach. The original in-person version of the module allows students to isolate their own strains of methylotrophic bacteria from plants they sample from the environment, to identify these using PCR, sequencing, and phylogenetic analysis, and to contribute their strains to original research in a university lab. The adapted version strengthens the focus on bioinformatics and increases its flexibility and accessibility by making the lab portion optional and adopting free web-based tools. Student feedback and graded assignments from Spring 2021 revealed that the lesson was especially effective at introducing the concepts of BLAST and phylogenetic trees, and that students valued and felt inspired by the opportunity to conduct hands-on work and to participate in community science.

Keywords:

project-based learning
course-based research experience
microbial ecology
bioinformatics
methylotrophy
community science

INTRODUCTION

Biology education has recently embraced inquiry-based learning (1–3), due to its potential to improve enthusiasm for and learning retention in sciences, technology, engineering, and math (STEM), and to increase confidence among less-prepared students or those from underrepresented minorities (URM) (4, 5). In particular, Course-based Undergraduate Research Experiences (CUREs), allowing classes to engage in research questions of interest to the community, can achieve these advances (6–8). Much published work on CUREs focuses on college students, yet high school students may also benefit from CUREs. Moreover, interaction with STEM practitioners can help to widen high school students' understanding of who can be a scientist (9) and create bridges between classrooms and community STEM opportunities (10).

Intended audience

Here we present Methylothon, a high school learning module in microbial ecology and evolution, similar to a CURE but designed for 11th- and 12th-graders. This curriculum description is designed to help high school instructors teach Methylothon in their own classrooms, and to inspire research laboratories to conduct outreach with high schools. Methylothon was originally designed as a "nose-to-tail" microbial ecology module for high school bioscience classes, in which students isolate organisms from their local environment, conduct PCR and sequencing for identification, and learn bioinformatic tools to place their isolates on the tree of life, similar to SEA-PHAGES or PARE (11, 12). It includes a community science component, as isolates can contribute to original research by collaborating researchers (our co-authors). Optional final projects allow students to delve deeper into microbial diversity in the context of biotech or evolution, or focus on research design. Methylothon thus provides a scaffold for building diverse skills and experiences, including hypothesis formulation, field work, laboratory techniques in microbiology and molecular biology, bioinformatics, science communication, and literature review. It touches on several core concepts for biological literacy outlined in the AAAS 2011 Vision & Change report (Evolution; Information Flow; and Systems) and has potential to cover all six of the report's listed core competencies (1). It also incorporates 17 concepts and skills from the American Society for Microbiology (ASM) Curriculum Guidelines for Undergraduate Teaching (13) (Appendix 1).

Methylothon originated as part of a research grant to study the ecology and evolution of plant-associated methylotrophic bacteria. We recognized that we could share our science with students by involving them in isolating methylotrophs. In addition to being extensively studied in the context of single-carbon metabolism (14, 15), pink-pigmented facultative methylotrophs make ideal model organisms for introducing students to microbiology. They are: ubiquitous in the environment (16, 17); low risk to human and plant health; straightforward to isolate at room temperature; used in numerous biotechnological applications (17–21); and, importantly, pink. This pigmentation makes them easy to identify on agar plates.

To develop Methylothon, the university researchers in our group used an informal network to find high school teachers in the San Francisco Bay Area who were interested in using a methylotroph lab to replace other lessons on molecular biology and phylogeny. We discussed learning goals and the context of the existing curricula before developing course materials. The methods described here are intended to be accessible to most biology laboratories in universities or well-equipped high schools, and methylotrophs can be cultured from nearly any plants in any geographic location. We encourage high schools, undergraduate classrooms, and high school-university partnerships around the world to consider implementing and adapting Methylothon.

Redesigning the module in 2020-2021 to meet the needs of remote instruction during the COVID-19 pandemic unfortunately required breaking the link between the students' original data collection and data analysis, which are a valuable part of inquiry-based learning (22). However, product was a more versatile module that is accessible to classes without a molecular biology lab and/or with limited computing capacity. Here we describe the remote-learning version of Methylothon, but our online materials include resources for in-person laboratory instruction. Teachers may choose their own approach depending on learning goals and resources.

Learning time

The full Methylothon module is 7 sessions. The timing of delivery can easily be varied (Tables 1-2). Classes taught in 2021 took 1 - 3 weeks, with some classes adding a final project at the end.

Prerequisite student knowledge

Methylothon is designed for high school Juniors and Seniors; students should begin with awareness of microorganisms, DNA, and the central dogma. Some of our partner classes already had experience with PCR, culturing microorganisms, or conducting BLAST; for such classes, Methylothon can be a review/synthesis. For classes where all the concepts are new, teachers may use Methylothon as a scaffold for just-in-time learning, and supplement with additional material to reinforce particular learning objectives. The objectives listed here correlate with 17 key concepts and skills from the ASM Recommended Curriculum Guidelines for Undergraduate Microbiology Education (13), notated by their numbers in brackets, and described in Appendix 1. Methylothon offers flexibility to emphasize the learning objectives appropriate for the class; for example, those marked with an asterisk were taught in some classes but not all in 2021.

Learning objectives

By the end of Methylothon, students should be able to:

a) Describe the ubiquity and diversity of microbes in the environment [ASM 20, 27]

b) List 3 applications for microbes in biotechnology [ASM 23, 26]

c) Describe what methylotrophs are and what distinguishes them from other microbes [ASM 11, 12, 13]

d) Define what rare earth elements (lanthanides) are and identify the role they play in microbiological processes* [ASM 11, 12, 13]
e) Describe 3 methods used to identify microbes [ASM 34]
f) Explain how selective medium is used to culture methylotrophs [ASM 33]
g) Demonstrate standard methods for culturing microorganisms on agar plates and analyze the effects of methodological changes (the roles of temperature and moisture, etc.)* [ASM 33, 36, 37]
h) List the steps of PCR and explain the function of each ingredient [ASM 36]
i) Describe the role of the 16S rRNA gene in the study of evolutionary relationships [ASM 4, 5]
j) Run an NCBI BLAST analysis [ASM 34]
k) Describe what Multiple Sequence Alignment is and how it is used for understanding relatedness [ASM 5]
l) Analyze a phylogenetic tree in context of sample metadata to explore scientific questions [ASM 5]
m) Formulate and evaluate hypotheses, describe experimental procedures, and discuss sources of uncertainty for a microbial ecology experiment* [ASM 28, 29, 30, 38]

PROCEDURE

Methylothon consists of lectures, homework assignments, and virtual labs, and, optionally, an in-person leaf press lab (Fig. 1); all materials are available at http://methylothon.com. Also optional, but included in all our 2021 lessons, is an initial guest lecture by a practicing microbiologist. Students prepare for the guest lecture by reading a blog post (23), profile, or a scientific article written by the researcher (24), and formulating questions. Guest lectures discuss current research in methylotrophy (e.g. the role of rare earth elements), or microbial ecology and biotechnological applications more generally. Material covered in the guest lecture is not required for the remainder of the module; rather, the primary goal is as an entry event to spark interest and provide an opportunity to interact with a practicing scientist.

After the guest lecture, the first lesson of the module covers the motivation for working with methylotrophs and an overview of the process of sampling, isolating, and sequencing isolates, presented either as preparation for the leaf-press lab (for classes that include the lab), or as background information on the origin of the sequences that students would soon analyze (for classes that omit the lab). A "how to make a leaf press" video (available at http://methylothon.com) is provided to students as a guide for techniques such as the use of gloves, handling leaves, and parafilming a culture plate.

For partner schools that include the leaf-press lab (Appendix 2), it is assigned as asynchronous work due by the following day. In the lab, students collect plant leaves from their neighborhood, press them briefly onto selective media (recipe provided in Appendix 3), and incubate the plates, wrapped in parafilm, at room temperature in their homes. In 2021, our team provided teachers with culture plates, gloves, and parafilm to distribute to students with instructions for handling. To

explore the finding that some methylotrophs depend on rare earth elements for growth (24), each student used two plates: one containing lanthanum and one without. Incubation periods may last 5-10 days, depending on class schedule, though better results are obtained after >1 week. At the end of each lesson, plates with colonies can be returned to a partner lab (in 2021, the Martinez-Gomez lab at UC Berkeley), where isolates can be further characterized; or classes carrying out in-person instruction may conduct colony PCR (Appendix 4) and Sanger sequencing. Students are required to complete online survey forms to submit photographs of plants and culture plates (Fig. 2) and record observations: location, host plant identity (we encourage use of the iNaturalist Seek app, https://www.inaturalist.org/pages/seek_app), and any colony growth observed.

The second lecture covers the 16S rRNA gene and its role in microbial phylogeny. Students read two brief articles from the microBEnet website before lecture (25, 26), and review the homework at via an interactive comment board, Google Jamboard (https://jamboard.google.com/). The lecture then discusses the tree of life and the use of 16s rRNA as a standard for measuring the relatedness of bacteria and archaea, and reviews how to interpret relationships from phylogenetic trees, with practice problems and breakout rooms to check understanding. The third lecture covers DNA sequencing and Multiple Sequence Alignment (MSA). Pre-assignments include short readings and a video on PCR and sequencing, reviewed in class using a Jamboard. The lecture covers interpretation of Sanger sequencing chromatograms and MSA, and includes an in-class exercise on the underlying concepts of sequence alignment.

In place of the PCR and sequencing that would be done lab during in-person instruction, students next complete a virtual lab on bacterial identification produced by Howard Hughes Medical Institute Biointeractive (https://www.biointeractive.org/classroom-resources/bacterial-identification-virtual-lab). The simulation begins with collecting a bacterial colony from a culture plate and ends with a "mini-BLAST" of the sequenced DNA. We provide worksheets to reinforce learning for either asynchronous independent and synchronous group-work settings.

In the final lecture, we walk through the Bioinformatics Lab that the students will ultimately use to identify a mystery methylotroph. This includes NCBI BLAST and free online tools for sequence alignment and phylogeny. For the Bioinformatics Lab, students receive "mystery sequences": FASTA files of 16S rRNA gene sequences from methylotrophs isolated by previous classes. They also receive a file containing reference sequences of other bacteria for context. The students perform BLAST on their sequences for an initial identification, then perform MSA and construct a phylogenetic tree with the mystery sequences and reference sequences. Although each student has a unique sequence, they work in small groups in online breakout rooms for peer support. Students complete a worksheet during the lab, and each ultimately uploads an image file of their phylogenetic tree to a class slideshow. The final worksheet questions require students to interpret their own and their classmates' trees from the slideshow.

Suggestions for determining student learning

The main formative assessments for Methylothon are the handouts completed by students during the virtual labs, and a final project constitutes the summative assessment. Other components of the module, including submission of photos and metadata for the leaf press lab, can be used for judging participation. During virtual lab group work, instructors circulate among the breakout rooms, providing additional opportunity for informal assessment of learning and evaluation of the module itself.

In 2021, teachers developed final projects based on their classes' specific learning goals; examples are on our website. For instance, students in one Biotechnology class wrote a 2-page individual lab report summarizing the Leaf Press Lab, including a literature search on a question of their choice related to microbes in biotech (Appendix 5). In the International Baccalaureate (IB) Biology class, students wrote group reports describing what they could and could not conclude from their Leaf Press Lab results, practicing IB International Assessment (IA) rubric guidelines (Appendix 6). That class subsequently assigned students to write 5-6-page individual essays combining Methylothon and a previous Evolution unit, to make an evidence-based argument and design an experiment demonstrating evolutionary principles. Another Biotechnology class combined Methylothon with a human ancestry unit and asked students for a report in their chosen format (video, slideshow, essay, graphic novel) on the process of identifying biological samples using PCR, sequencing, and phylogenetics (Appendix 7). Pairing modules thus can emphasize the universality of these methods and principles.

Sample data

Methylothon generates both work that can be used for evaluating learning, and sample metadata and observations that can feed into the community-science component. Summaries of sample metadata from 2021 are shown in Fig. 3A and Appendix 8, and examples of students' observations are provided in Table 3. Appendices 9-10 show examples of students' individual/group work for evaluation.

Safety issues

Methylothon carries relatively few safety risks. For class-based plant sampling trips, typical field trip safety issues apply. For in-person instruction, instructors should provide lab safety training as necessary; please see the ASM Guidelines for Biosafety in Teaching Laboratories (27), and the "Addendum for biosafety considerations regarding at-home or DIY microbiology kits," with the caveat that there is no commercial kit provider to assume liability. Students at home do not use ethanol, flame, or other components of sterile technique that typically introduce laboratory hazards. Moreover, students do not re-open parafilmed leaf press plates, so the risk of contact with bacterial cultures is low.

One unavoidable hazard is the cycloheximide in the culture medium; we have found its inclusion to be necessary for inhibiting fungal growth. No occupational exposure limit has been established for cycloheximide, but it is mutagenic and teratogenic. Exposure for students is low because they do not handle concentrated chemical, but they must wear gloves when handling culture plates. Safety information on cycloheximide is included in the lesson materials, and instructors should emphasize to students not to touch the agar even with gloves, to leave plates always sealed with parafilm, and to wash hands after the experiment. For remote instruction, students must return culture plates and gloves to instructors for appropriate disposal in lab.

Because sample metadata uploaded to methylothon.com will be made publicly available alongside sequence data, student privacy should also be taken into account. Student names are collected so instructors can track participation, but are never published. If students sample at their homes, they are allowed to list GPS coordinates from nearby that do not identify their home addresses.

DISCUSSION

Field testing

The in-person version of Methylothon was delivered two consecutive years (fall 2018, 2019) to an Advanced Biotechnology class (seniors) at Abraham Lincoln High School (ALHS) in San Francisco, CA. , When many schools moved instruction online during the COVID-19 pandemic in 2020-2021, we modified Methylothon for remote learning, and simultaneously expanded to new schools. Our spring 2021 partners included ALHS, three sections of IB Biology (seniors) at Berkeley High School in Berkeley, and four sections of Biotechnology (juniors and seniors) at Galileo Academy of Science and Technology in San Francisco. Together, this entailed approximately 220 students and 55 hours of synchronous class time, with substantial variation in learning goals and student experience level. COVID-related institutional policies also varied, such as the balance of synchronous versus asynchronous learning, and the regulation enforced by some schools that students not be required to download software or register online user accounts.

Evidence of student learning

Each of the 2021 Methylothon classes assessed students differently. One of the Biotechnology classes (3-week schedule) used formative assessments during the sequence on PCR and bioinformatics, and summative assessments (analysis of phylogenetic trees and a final "abstract" writing assignment) at the end (Appendix 5). Students were evaluated for their understanding of methylotroph biology, communication of methods, interpretation of results, and elements of the scientific method such as research question formulation and use of references (scoring guide in Appendix 5). More than 70% of students showed early mastery of concepts related to microbial culturing, PCR, DNA sequencing, and DNA databases. Fewer demonstrated mastery of bioinformatics core concepts, particularly the use of BLAST and phylogenetic trees, with most difficulties relating to the purpose of intermediate steps of the analysis pipeline. We attribute this

to the fact that it was the students' first exposure to these concepts; future implementations might include additional background learning and/or practice opportunities to supplement Methylothon.

The IB Biology class (1-week schedule) assigned a summative assessment in which teams of 4 synthesized their work in a collaborative essay, including a hypothesis they had formulated and tested during the leaf press lab. Students were evaluated on their ability to incorporate background research to develop their research question and hypothesis, summarize their procedure, analyze their data quantitatively and qualitatively, develop a conclusion supported by evidence, and evaluate their experimental design (assignment and grading rubric in Appendix 6). The area that proved most challenging was the analysis of non-procedural sources of error and limitations of the experimental designs. Furthermore, because students were given only 1-2 days to develop their hypothesis, many did not have time to find and analyze scientific articles, but rather developed hypotheses based on background knowledge and information from the unit. However, all teams demonstrated achievement of core objectives, with 100% of the 21 teams receiving a grade of a 3.5 (out of 4) or higher, and 16 teams receiving a 4. The same class completed an additional assessment asking students to design an experiment using MSA, PCR, and BLAST to demonstrate microbial evolution. Over 90% of students successfully utilized the information from Methylothon, and many even featured bacteria outside of the methylotrophs. We recommend that classes consider incorporating experimental design in summative assessments for Methylothon, so students may continue developing these essential skills.

In addition to school-specific assessments, we gathered feedback from all classes via an anonymous survey sent 5-8 weeks after the final lesson. We asked two questions: what students remembered most from the lesson, and what they considered the most valuable takeaway. We received 79 responses, which we grouped into 7 categories based on main topic (Fig. 4). Students reported remembering and valuing concepts relating to phylogenetic trees most often, followed closely by BLAST. We are encouraged by this result — most students remembered the lesson's central goals! Additionally, several students connected BLAST and building phylogenetic trees in their answers. To a lesser extent, students remembered intermediate steps such as DNA sequencing and the use of databases to retrieve information.

In addition, students who participated in the Leaf Press Lab (48 of 79 respondents) responded positively to conducting in-person lab work; 25% of those students reported that they most remembered the opportunity to do a non-virtual lab, and nearly 20% reported that the lab was what they found most valuable about Methylothon. Finally, several of the students who did the lab mentioned that *Methylobacterium* are pink, highlighting the memorable nature of this phenotype. Some other facets highlighted by students as particularly helpful included the opportunity to speak to scientists about their background and research interests, and the chance to "experience what it is like to be a real scientist" (Box 1).

Possible modifications

Teaching Methylothon to several partner schools helped us identify multiple areas for improvement. One major pivot we made in later lessons was to change the file extensions of the DNA sequence FASTA files. FASTA is a format for molecular sequence data in a plain-text file; the .FASTA file extension can remind users of the format and help sequence analysis programs recognize the file. However, most students did not have the necessary underlying knowledge of file formats, extensions, and how to use files without viewing them, leading multiple classes to lose lesson time. We therefore changed the file extensions to txt.

Additionally, we found some students struggled to translate concepts from our initial phylogeny lecture to their interpretation of the trees they generated, including the significance of the reference sequences. We therefore added in a walkthrough of tree interpretation immediately after the students had shared their phylogenetic trees in the class slideshow and before they began interpretation. We found this review helped students approach interpreting their own results with more confidence. Relatedly, we found that some classes might do better with a lesson structure based on "just-in-time" learning. When Methylothon was stretched across multiple weeks, the elapsed time between lectures and hands-on practice in the bioinformatics lab allowed some concepts to be forgotten before they were practiced. One partner teacher related that they would consider breaking up the bioinformatics lab so individual steps followed immediately after the relevant lecture.

The unique challenges of executing lab work in a remote-learning environment included the question of how to provide a valuable scientific experience for the students, without being able to identify the methylotroph isolates. In one class, students formed hypotheses that could be evaluated based on colony growth alone, and then discussed limitations in the experimental setup (e.g. the number of plates available, lack of controls, short incubation time) in their written report. For students who had experience with fast-growing organisms such as *Escherichia coli*, we had to moderate expectations by emphasizing methylotrophs' slower growth rate.

Adoption by other programs

With Methylothon we aim to provide a framework for a microbiology unit that is adaptable and accessible to diverse high schools and microbiology research laboratories. For most programs, the most challenging implementation issue will likely be making the leaf press lab culture medium. Interested groups are welcome to contact the authors for help with the C7 metals mix, the most complex component of the medium.

With schools returning to in-person instruction, two of our participating 2021 schools chose to repeat Methylothon in spring 2022 and continued to adapt and expand it. One school that had not previously carried out the leaf press lab added it, and the other added a new ecological statistics exercise based on methylotroph biogeography in order to incorporate Methylothon into the

Ecology unit (where it had previously been part of Evolution). We foresee ample opportunity for further adaptations. As the name Methylothon implies, it was originally intended to incorporate an element of competition. For classes able to devote more time, students' isolates could be cultured by partnering research laboratories and "competed" in a wide variety of phenotypic assays to introduce concepts of ecological niche and evolutionary adaptation, while providing a chance for every isolate to "win" at something. As more lessons are developed and more methylotroph data are gathered each year, we will continue to expand the offerings on the Methylothon website.

ACKNOWLEDGEMENTS

We are grateful to the members of the Martinez-Gomez Lab (UC Berkeley), particularly Nathan Good for assistance with handling methylotroph isolates in 2021. We thank the laboratory of José de la Torre (San Francisco State University) for assistance with the in-person version of Methylothon in 2018-2019, especially Brittany Baker for providing help with bioinformatics instruction. Methylothon is funded by a grant from the National Science Foundation Division of Environmental Biology, Dimensions of Biodiversity program (award 1831838), to C. J. Marx, N. C. Martinez-Gomez, and J. A. Lee.

This project was reviewed by the Office of Research Assurances on behalf of the University of Idaho Institutional Review Board (Project 22-044). It was determined not to meet the definition of Human Subjects Research and therefore not to require IRB oversight.

REFERENCES

1. Bauerle C, DePass A, Lynn D, O'Connor C, Singer S, Withers M. 2009. Vision and Change in Undergraduate Biology Education: A call to action. American Association for the Advancement of Science, Washington, D.C. https://www.visionandchange.org/. Retrieved 17 May 2021.

2. NASEM (National Academies of Sciences, Engineering, and Medicine). 2015. Integrating discovery-based research into the undergraduate curriculum: Report of a convocation. The National Academies Press, Washington, DC. https://www.nap.edu/catalog/21851/integrating-discovery-based-research-into-the-undergraduate-curriculum-report-of. Retrieved 17 May 2021.

3. NASEM (National Academies of Sciences, Engineering, and Medicine). 2017. Undergraduate Research experiences for STEM students: Successes, challenges, and opportunities. The National Academies Press, Washington, D.C. https://www.nap.edu/catalog/24622/undergraduate-research-experiences-for-stem-students-successes-challenges-and-opportunities. Retrieved 17 May 2021.

4. Blumer LS, Beck CW. 2019. Laboratory courses with guided-inquiry modules improve scientific reasoning and experimental design skills for the least-prepared undergraduate students. LSE 18:ar2.

5. Estrada M, Hernandez PR, Schultz PW. 2018. A Longitudinal Study of How Quality Mentorship and Research Experience Integrate Underrepresented Minorities into STEM Careers. LSE 17:ar9.

6. Bangera G, Brownell SE. 2014. Course-based undergraduate research experiences can make scientific research more inclusive. LSE 13:602–606.

7. Auchincloss LC, Laursen SL, Branchaw JL, Eagan K, Graham M, Hanauer DI, Lawrie G, McLinn CM, Pelaez N, Rowland S, Towns M, Trautmann NM, Varma-Nelson P, Weston TJ, Dolan EL. 2014. Assessment of course-based undergraduate research experiences: A meeting report. LSE 13:29–40.

8. Wang JTH. 2017. Course-based undergraduate research experiences in molecular biosciences—patterns, trends, and faculty support. FEMS Microbiology Letters 364:fnx157.

9. Lescak EA, O'Neill KM, Collu GM, Das S. 2019. Ten simple rules for providing a meaningful research experience to high school students. PLoS Comput Biol 15:e1006920.

10. PCAST (President's Council of Advisors on Science and Technology). 2010. Prepare and Inspire: K-12 education in science, technology, engineering, and math (STEM) for America's future. Executive Office of the President.

11. Jordan TC, Burnett SH, Carson S, Caruso SM, Clase K, DeJong RJ, Dennehy JJ, Denver DR, Dunbar D, Elgin SCR, Findley AM, Gissendanner CR, Golebiewska UP, Guild N, Hartzog GA, Grillo WH, Hollowell GP, Hughes LE, Johnson A, King RA, Lewis LO, Li W, Rosenzweig F, Rubin MR, Saha MS, Sandoz J, Shaffer CD, Taylor B, Temple L, Vazquez E, Ware VC, Barker LP, Bradley KW, Jacobs-Sera D, Pope WH, Russell DA, Cresawn SG, Lopatto D, Bailey CP, Hatfull GF. 2014. A broadly implementable research

course in phage discovery and genomics for first-year undergraduate students. mBio 5:e01051-13.

12. Genné-Bacon EA, Bascom-Slack CA. The PARE Project: A short course-based research project for national surveillance of antibiotic-resistant microbes in environmental samples. Journal of Microbiology & Biology Education 19:19.3.40.

13. Merkel S, Reynolds J, Hung K, Smith H, Siegesmund A, Smith A, Baker N, Chang A. 2012. Recommended curriculum guidelines for undergraduate microbiology education. American Society for Microbiology.

14. Chistoserdova L, Lidstrom PME. 2013. Aerobic methylotrophic prokaryotes, p. 267–285. *In* Rosenberg, E, DeLong, EF, Lory, S, Stackebrandt, E, Thompson, F (eds.), The Prokaryotes. Springer Berlin Heidelberg.

15. Kelly DP, McDonald IR, Wood AP. 2014. The family *Methylobacteriaceae*, p. 313–340. *In* Rosenberg, E, DeLong, EF, Lory, S, Stackebrandt, E, Thompson, F (eds.), The Prokaryotes, 4th ed. Springer Berlin Heidelberg.

16. Iguchi H, Yurimoto H, Sakai Y. 2015. Interactions of methylotrophs with plants and other heterotrophic bacteria. Microorganisms 3:137–151.

17. Kumar M, Kour D, Yadav AN, Saxena R, Rai PK, Jyoti A, Tomar RS. 2019. Biodiversity of methylotrophic microbial communities and their potential role in mitigation of abiotic stresses in plants. Biologia 74:287–308.

18. Ochsner AM, Sonntag F, Buchhaupt M, Schrader J, Vorholt JA. 2014. *Methylobacterium extorquens*: methylotrophy and biotechnological applications. Appl Microbiol Biotechnol 99:517–534.

19. Tlusty M, Rhyne A, Szczebak JT, Bourque B, Bowen JL, Burr G, Marx CJ, Feinberg L. 2017. A transdisciplinary approach to the initial validation of a single cell protein as an alternative protein source for use in aquafeeds. PeerJ 5:e3170.

20. Hardy RW, Patro B, Pujol-Baxley C, Marx CJ, Feinberg L. 2018. Partial replacement of soybean meal with *Methylobacterium extorquens* single-cell protein in feeds for rainbow trout (*Oncorhynchus mykiss* Walbaum). Aquaculture Research 49:2218–2224.

21. Skovran E, Raghuraman C, Martinez-Gomez NC. 2019. Lanthanides in methylotrophy. Curr Issues Mol Biol 33:101–116.

22. Cooper KM, Knope ML, Munstermann MJ, Brownell SE. 2020. Students who analyze their own data in a course-based undergraduate research experience (CURE) show gains in scientific identity and emotional ownership of research. Journal of Microbiology & Biology Education 21.

23. Lee JA. 2015. What ice cream and biofuels have in common: vanillin and the microbes that eat it. BEACON Researchers at Work. https://www3.beacon-center.org/blog/2015/06/01/beacon-researchers-at-work-what-ice-cream-and-biofuels-have-in-common-vanillin-and-the-microbes-that-eat-it/.

24. Skovran E, Martinez-Gomez NC. 2015. Just add lanthanides. Science 348:862–863.

25. 2011. Fact Sheet: Ribosomal RNA (rRNA), the details. microBEnet: the microbiology of the Built Environment network https://microbe.net/simple-guides/fact-sheet-ribosomal-rna-rrna-the-details/. https://microbe.net/simple-guides/fact-sheet-ribosomal-rna-rrna-the-details/. Retrieved 19 May 2021.

26. 2011. Fact Sheet: rRNA in Evolutionary Studies and Environmental Sampling. microBEnet: the microbiology of the Built Environment network https://microbe.net/simple-guides/fact-sheet-rrna-in-evolutionary-studies-and-environmental-sampling/. https://microbe.net/simple-guides/fact-sheet-rrna-in-evolutionary-studies-and-environmental-sampling/. Retrieved 19 May 2021.

27. Byrd, Jeffrey J., Maxwell RA, Townsend HM, Emmert EAB. 2019. ASM guidelines for biosafety in teaching laboratories. American Society for Microbiology.

FIGURE LEGENDS

Figure 1. The components of Methylothon.

Figure 2. Selected photos of host plants and leaf press plates from Methylothon 2021. Both were collected in students' backyards in San Francisco. A-C) Plant, leaf press, and colony growth from host identified as Brambles (likely Himalayan Blackberry, *Rubus armeniacus*). D-F) Plant, leaf press, and colony growth from host identified as Cape Ivy, *Delairea odorata*.

Figure 3. A) Sites sampled for plants by San Francisco and Berkeley high school students during 2021 Methylothon; each pink dot represents one sample site. Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under ODbL. Map generated using R v4.0.2 with RStudio v1.3.959. B) Example of a student's phylogenetic tree, in which the name of the student's mystery sequence is in green text. Image taken from the class phylogenetic tree slide show; student's name is omitted.

Figure 4. Topics mentioned in student responses collected from an anonymous survey that asked "What's the one thing you remember most about the lesson?" and "What did you find most valuable about the lesson?". Some responses incorporated multiple topics and were therefore counted more than once. Responses not fitting a predefined category were classified as "Other."

TABLES

Table 1. The Methylothon curriculum can be taught over various timetables. Shown here is one of two example lesson plans executed by our partner schools. In this class, the entire module was delivered in the course of a single week, as students met daily for lecture and additionally twice a week for lab sessions.

1-Week Example Schedule		
Day	Period	Lesson
1	morning	Guest Lecture; begin Leaf Press Lab
2	morning	Intro to Methylotrophs Lecture
2	afternoon	Tree of Life and Phylogenetics Lecture
3	morning	Bacterial Identification Lab (asynchronous)
4	morning	DNA Sequencing and Multiple Sequence Alignment Lecture;

		Bioinformatics Lab Walkthrough
4	afternoon	Bioinformatics Lab
5	morning	Finish Bioinformatics Lab; assign Final Project; conclude Leaf Press Lab

Table 2. The second of two example lesson plans executed by our partner schools. This class was able to devote only 2-3 periods per week to the module, so the full duration was three weeks.

3-Week Example Schedule		
Week	Day	Lesson
1	1	Guest Lecture
	2	Intro to Methylobacterium Lecture; begin Leaf Press Lab
2	3	Tree of Life and Phylogenetics Lecture
	4	DNA Sequencing and Multiple Sequence Alignment Lecture; Bacterial Identification Lab (asynchronous)
3	5	Bioinformatics Lab Walkthrough
	6	Bioinformatics Lab
	7	Complete Bioinformatics Lab; assign Final Project; conclude Leaf Press Lab

Table 3. Subset of the metadata accompanying the samples shown in Fig. 2, entered by students in the online form as part of the Leaf Press Lab. School name, sample date, and GPS coordinates are omitted.

Sample code	Description of sample location	Other observations	Where in your house did you incubate your plate?	What is the general appearance of the colonies?	Plant identificatio

CH 2242021	The sample was taken in my backyard, which faces the ocean. My plant was surrounded by a grassy area with minimal flowers. Today was a relatively sunny day and the area where I sampled my leaves from gets a lot of sunlight.	The plant leaves have many thorns on the back and stems so when placing the leaves on the agar plate, the thorns pierced the agar. The rest of the process went smoothly.	On my desk in my bedroom.	The colonies are light pink, salmon-colored, and are crowded around the veins of the leaves. Overall, they are about the same size, but there are some that stick out more than others. On the plate without lanthanum, the colonies are hard to distinguish from each other because they are so packed together.	Brambles
KN, 2/24/21	The sample location was my backyard, which resides in a considerably chilly area of San Francisco, in the Excelsior district. The terrain is mountainous, and we grow chili peppers as well as numerous other things. There are many weeds and leaves of sorts. The amount of sunlight received within the area is decent, especially early afternoon. The sample grew from the ground and not in a pot.	Might have cracked the agar, but not entirely sure. Can clearly identify the outline of the plant on the agar, however. Process went smooth, in terms of sampling and identifying the plant One of the leaves was a bit too big for the plate but somewhat worked.	I incubated my plate in my room, besides the window. It protects the plates from excessive sunlight but still allows some to shine in, therefore the agar would not dry out. The area is bit chilly and moist.	On the 19 No La plate, the colonies are all pink, some lighter in color than others. They take the form and shape of the leaf, and do not spread out too much. The colonies are very tightly packed and it is difficult to distinguish among them. On the other hand, the 19 + La plate has a few white colonies that are opaque, and the rest are light pink, much lighter than the colonies on the 19 No La plate. They are very spread out, compared to the 19 No La plate.	Cape Ivy Delairea Genus: Delairea F Asteraceae Classi tribe Senecioneae as German ivy in o the world.

Box 1. Selected responses from the anonymous student survey.

- "What I found most valuable about the lesson was honestly just being back in a lab oriented setting to a certain extent. Since COVID-19 has affected us in not doing many labs, I believe this lab was something that made me feel a bit more interested and motivated again."
- "[What I valued most was] seeing the other student's [phylogenetic tree] results to compare."
- "[What I valued most was] getting to experience what it is like to be a real scientist."
- "The lesson I found most valuable was being able to conduct the experiment, even if it was virtual, I think that really fun and engaging even over [Z]oom"
- "[What I valued most was] the chance to watch as pink bacteria grew on my petri dishes

in the shape of the leaves, knowing that my data may potentially be used in a real scientific paper!"

- "I found the most valuable part of the lesson being learning about what is actually being researched right now in the field of biology."
- "I found the introduction of BLAST the most valuable. It was my first time hearing of this term, and being taught how to utilize this program is a great skill I could carry with me in the future."
- "The thing I found most valuable about the lesson was how easy it is to go outside and collect a sample to learn more about. "

Supplemental materials for Jones et al. (2022) Methylothon: a versatile course-based high school research experience in microbiology and bioinformatics-- with pink bacteria

Contents

Appendix 1. Components of ASM Curriculum Guidelines (Merkel et al. 2012) covered in Methylothon. 2

Appendix 2. The Leaf Press Lab handout for students. 4

Appendix 3. MP medium for Methylothon leaf press plates 9

Appendix 4. Colony PCR protocol for amplification of 16S rRNA and *rpoB* marker genes to identify methylotroph isolates 12

Appendix 5. An abstract-style writeup assignment as a summative assessment for Methylothon, given in a Biotechnology class 16

Appendix 6. A Team Quiz as a summative assessment for Methylothon, assigned in an International Baccalaureate Biology class. 18

Appendix 7. A flexible-format final assignment combining Methylothon with human ancestry, given as a summative assessment in a Biotechnology class. 22

Appendix 8. Plants sampled by students for leaf presses during Methylothon 2021. 25

Appendix 9. Example of student work on Bacterial Identification Virtual Lab worksheet. 26

Appendix 10. Example of student work in Bioinformatics Virtual Lab worksheet. 34

Appendix 1. Components of ASM Curriculum Guidelines covered in Methylothon.

Reference: Merkel S, Reynolds J, Hung K, Smith H, Siegesmund A, Smith A, Baker N, Chang A. 2012. Recommended curriculum guidelines for undergraduate microbiology education. American Society for Microbiology.

Concepts and Statements		
	Evolution	
	4	The traditional concept of species is not readily applicable to microbes due to asexual reproduction and the frequent occurrence of horizontal gene transfer.
	5	The evolutionary relatedness of organisms is best reflected in phylogenetic trees
	Metabolic Pathways	
	11	Bacteria and Archaea exhibit extensive, and often unique, metabolic diversity
	12	The interactions of microorganisms among themselves and with their environment are determined by their metabolic abilities
	13	The survival and growth of any microorganism in a given environment depends on its metabolic characteristics.
	Microbial Systems	
	20	Microorganisms are ubiquitous and live in diverse and dynamic ecosystems.
	23	Microorganisms, cellular and viral, can interact with both human and nonhuman hosts in beneficial, neutral or detrimental ways.
	Impact of Microorganisms	
	26	Humans utilize and harness microorganisms and their products.
	27	Because the true diversity of microbial life is largely unknown, its effects and potential benefits have not been fully explored.
Competencies and Skills		
	28	Ability to apply the process of science
	29	Ability to use quantitative reasoning
	30	Ability to communicate and collaborate with other disciplines
Microbiology Laboratory Skills		

	33	Use pure culture and selective techniques to enrich for and isolate microorganisms.
	34	Use appropriate methods to identify microorganisms (media-based, molecular and serological).
	36	Use appropriate microbiological and molecular lab equipment and methods.
	37	Practice safe microbiology, using appropriate protective and emergency procedures.
	38	Document and report on experimental protocols, results and conclusions.

Appendix 2. The Leaf Press Lab handout for students.

Name: **Date:** **Period:**

Aims

1. To isolate novel methylotrophic bacteria from plants in your geographic area
2. To compare microbial growth with and without the rare earth element lanthanum

1. FIELD SAMPLING

Materials

- latex/nitrile gloves (provided by school)
- brand-new, clean plastic sandwich bag
- scissors
- this protocol
- pencil or pen
- a phone or camera for taking digital photographs
- optional: phone with the Seek app installed
- optional: field guide to regional plants

Procedure

1. Gather materials and walk to your field sampling location.

2. Choose a plant to sample. Any plant will work! But you need to find a leaf (or cluster of leaves, if the leaves are very small) that is small enough to fit fully within the culture plate, and large enough to have some area on both halves of the plate.

3. Take a few photos of your plant and its surroundings.
Write a description of the location in the Field Notes section below. If desired, you may also use your phone to take GPS coordinates. If you don't take your coordinates now, you'll need to estimate them later using a digital map.

4. Identify the plant. For this you may want to use the Seek app, or any other field plant guide you choose to bring. If you can't figure it out now, you will need to do a more thorough search when you get home using the photos you've taken and what you observed of its location and growth habits. Try to identify the taxonomic (scientific) name of the plant to the level of genus, if possible. If you have a common name, include that too. Record on your sample notes sheet.

Collect leaf samples as follows:

5. Put on gloves, to prevent microbes on your hands from jumping to your leaf sample.

6. Open your sample bag (without sticking your hand in!) and put the bag around the leaf that you want to sample. Use your fingers or a pair of scissors to snip the leaf so that it falls into the bag. Harvest at least two leaves from the same plant.

Field observations

Description of sample location:

GPS coordinates [approximate is fine]

Plant identification

Other observations

Online data entry #1

Please use our project website to share all the information about the plant sample you used for this experiment. This information is absolutely necessary for the Martinez-Gomez lab to understand the *Methylobacterium* strains they will isolate from your samples.

1. Navigate to http://methylothon.com.
2. From the main menu, go to "Methylothon 2021: student page"
3. Sign in with the password [*insert this year's password here*]
4. Follow the instructions on the website under "After you've collected your sample and made a leaf press."

You will be asked to enter all the information you have recorded on this spreadsheet. You will also be asked to upload photos of the following:
1. the plant you sampled
2. the leaf press you made, showing the leaf on the culture plate
Before uploading, please rename your photos so that they contain your unique sample code!

2. LEAF PRESS (see the "How to Make a Leaf Press" video)

Materials
- your leaf sample
- 2 methylotroph culture plates, one with lanthanum (La) and without (provided by school)
- latex or nitrile gloves (provided by school)
- strips of parafilm (provided by school)
- a permanent pen
- a phone or camera for taking digital photographs

IMPORTANT SAFETY NOTE: The culture plates contain cycloheximide, an anti-fungal compound that can also be harmful to humans. Always wear gloves when handling your culture plate if the lid is open, try not to contact the agar directly with your gloves, and wash your hands thoroughly after the experiment. Keep plates closed and sealed with parafilm when possible.

Procedure
1. Prepare your culture plates by writing your name and the date on the back using a permanent pen. Try to write in small lettering near the edge of the plate so that you'll still be able to see most of the agar through the bottom of the plate.
In addition, **please come up with a sample code** consisting of 4-10 letters and/or numbers that we can use to easily connect your sample, your data, and your photos. Write this sample code on your plate, as well.

2. Make a leaf press. Wearing gloves, press a leaf directly onto the surface of a culture plate to allow microbes to transfer to the agar. If you're working a cluster of very small

leaves, make sure some leaves contact each side of the plate. Make sure the ***bottom side of the leaf*** contacts the plate. Do not touch the agar itself with your gloved fingers. ***Just a little pressure is sufficient-- try not to gouge the agar too deeply!*** We want microbes growing on the surface of the agar, not inside it.
Repeat with the second leaf on the second plate.

3. Take a photo of your leaf presses (ideally, both plates in one photo).

4. Carefully remove the leaves without disturbing the agar, and discard the leaves.

5. Replace the lid on the plate, and wrap a strip of parafilm around the edge of the plate to hold the lid on and prevent evaporation while allowing gas exchange. Store the plate ***upside-down*** at room temperature; if you store it right-way up, condensation may accumulate on the lid and then drip back down onto the agar, forming puddles. If possible, keep the place at warm (but not hot!) room temperature to encourage faster microbial growth. But keep the plate out of direct sunlight.

3. LABORATORY OBSERVATIONS

At least 1 week after plating, or on the date chosen by your class, check on your plates. You can observe the microbial growth through the bottom of the plates-- do not open the lid. Record the following.

Date: _____ Time: _____

Where in your house did you incubate your plate?

Approximately how many colonies are there on each of the two plates?

What is the general appearance of the colonies? (Color, sheen, shape, size. Are they uniform or diverse? Are they so packed together that it's hard to distinguish among them?)

Take a photo of your plates.

Online data entry #2

We want to collect your observations of the microbial growth on your plate, as well as the photo you took. You can reach the data entry form and the folder for photo uploads in the same place that you did at the beginning of the study.

1. Navigate to http://methylothon.com.
2. From the main menu, go to "Methylothon 2021: student page"
3. Sign in with the password [*insert this year's password here*]
4. Follow the instructions on the website under "After you've incubated your plates for ~1 week"

Before uploading, please rename your photos so that they contain your unique sample code!

What to do with your materials

Store used gloves in a plastic bag and return to your teacher for disposal in lab.
Return culture plates to your teacher for further study or disposal in lab, as appropriate.

Appendix 3. MP medium for Methylothon leaf press plates

Yield: 1000 ml (approximately 30-40 plates)
Overview: Make MP medium, a defined minimal medium optimized for methylotrophs, with agar. Just before pouring plates, add the appropriate carbon source, lanthanides, vitamins, and fungal inhibitors for leaf press plates.

Safety: Leaf press plates require supplementation with cycloheximide, a hazardous compound. Cycloheximide is fatal if swallowed, can cause skin and eye irritation on contact, and may cause genetic or developmental defects. Always wear gloves when handling cycloheximide and leaf-press plates, and wash hands thoroughly afterward.

For MP (Modified PIPES) medium
From Delaney et al., 2013. doi: 10.1371/journal.pone.0062957
(http://www.ncbi.nlm.nih.gov/pubmed/23646164)
Each batch of MP medium is composed from the following stock solutions. Full recipes and instructions for all stock solutions is in "Preparation," below.

Name	**Stock concentration**	**Final concentration in MP medium**	**Volume to add for 1 L**
PIPES (10x)	300 mM	30 mM	100.0 mL
P-Solution (100x)	145 mM	1.45 mM	10.0 mL
$MgCl_2$ (4000x)	2 M	0.5 mM	250 µL
$(NH_4)_2SO_4$ (250X)	2 M	8 mM	4.0 mL
C7-metals (1000X)	1.2 mM	0.0012 mM	1000 µL
$CaCl_2$ (100X)	2 M	0.02 mM	10 µL
Bacto Agar	n/a	15 g/L	15 g
ddH_2O			add to reach final volume

MP preparation

1	Prepare stock solutions as described in steps 2-8. If stock solutions are already available, skip to step 9.
2	To prepare PIPES stock solution, dissolve 90.711 g of PIPES free acid ($C_8H_8N_2O_6S_2$) in a final volume of 1 L of dH_2O. Suggestion: add some KOH to ~700 mL of water before adding any PIPES. Measure out all the PIPES you will need. Add a small amount of PIPES at a time; the solution will turn milky-white until the PIPES dissolves. If all the PIPES dissolves, add more until it doesn't. Then add more KOH and repeat. Going back and forth between KOH and PIPES, ensure that all the PIPES dissolves; JUST BE CERTAIN YOU DON'T OVERDO THE KOH. After you've added and dissolved

	all the PIPES the pH should still be acidic. Carefully adjust the pH to 6.75 by adding KOH. Bring the volume up to 1 L and check pH again; adjust with more KOH if needed.
3	To prepare P stock solution, add 33.1 g of K_2HPO_4* 3H_2O and 25.9 g of NaH_2PO_4* H_2O to 1 L of dH_2O.
4	To prepare the stock solution for $MgCl_2$ (magnesium chloride), add 81.32 g of $MgCl_2$* 6H_2O to 200 mL of dH_2O.
5	To prepare the stock solution for $(NH_4)_2SO_4$ (ammonium sulfate), add 52.9 g of $(NH_4)SO_4$ to 200 mL of dH_2O.
6	To prepare the stock solution for $CaCl_2$ (calcium chloride), add 58.8 g of $CaCl_2$ *2H_2O to 200 mL of dH_2O.
7	To prepare the stock solution for C7 metals, prepare a container with 100 mL of dH_2O, and add each of the following. Add them in the order listed, being sure to dissolve each completely before adding the next. 1) 1341.1 mg of $Na_3C_6H_5O_7$ (sodium citrate) 2) 34.5 mg of $ZnSO_4$* 7H_2O (zinc sulfate heptahydrate) 3) 19.8 mg of $MnCl_2$* 4H_2O (manganese chloride tetrahydrate) 4) 500.4 mg of $FeSO_4$* 7H_2O (ferrous sulfate heptahydrate) 5) 247.1 mg of $(NH_4)_6Mo_7O_{24}$* 4H_2O (ammonium molybdate tetrahydrate) 6) 24.96 mg of $CuSO_4$* 5H_2O (copper sulfate pentahydrate) 7) 47.58 mg of $CoCl_2$* 6H_2O (cobalt chloride hexahydrate) 8) 10.88 mg of Na_2WO_4* 2H_2O (sodium tungstate dihydrate)
8	Autoclave all separate solutions.
9	To make the media, mix all the components together. First, add the listed volumes from the stock solutions into a new bottle. Then, in a volume of ddH_2O sufficient to bring the medium to the final desired volume (calculate this based on the volumes of the supplements you will add after autoclaving - see below), melt the appropriate amount of Bacto Agar (15 g/L) using a stir bar and hot plate. DO NOT boil the agar. Combine all ingredients.
10	Distribute the media to smaller bottles (if desired) and autoclave.

Supplements should be added to agar medium while cool but still molten (approximately 50 °C), and the medium mixed thoroughly before being poured into plates.

Supplements for leaf press plates include:

- methanol (MeOH) [carbon substrate]: add 5 mL of 100% methanol per L of MP to reach a final concentration 125 mM
- cycloheximide [inhibits fungal growth]: purchase or make stock solution; add sufficient stock to reach a final concentration in medium of 50 μg/mL
- RPMI 1640 Vitamins Solution [facilitates growth of diverse organisms]: May be purchased commercially as 100x stock, sold as an ingredient for Roswell Park Memorial Institute (RPMI) 1640 medium. Add 10 mL of 100x stock per 1 L of medium.

- $LaCl_3$ (lanthanum chloride) [selective for some methylotrophs; used in only some plates]: make stock solution in ddH_2O; add sufficient stock to reach a final concentration in medium of 2 μM

Appendix 4. Colony PCR protocol for amplification of 16S rRNA and *rpoB* marker genes to identify methylotroph isolates

Colony PCR of two genes for sequencing for identification of your methylotroph isolate

Background

The sequence of 16S rRNA gene is commonly used as a phylogenetic "barcode" to identify unknown microorganisms. Similarly, the gene *rpoB* (which codes for the RNA polymerase enzyme) can also be used to identify organisms, and it is sometimes more useful than the 16S gene for distinguishing closely-related organisms.

In order to sequence these genes, they must first be amplified by PCR (polymerase chain reaction) to generate enough DNA of the specific gene to be sequenced. Today you will conduct PCR to amplify both genes from one of the strains you have isolated from the environment.

- **PCR template** is the DNA from the organism of interest that is added to the reaction mixture for amplification.
- **Colony PCR** is a quick and easy type of PCR in which the template consists of whole bacterial cells picked directly from a colony, rather than extracted DNA. An extra-long denaturation step (10 minutes) is included at the beginning of the program to help the cells lyse and release their DNA.
- A **no-template (or negative) control** is a PCR reaction in which no template is added; if this reaction generates a PCR product, it indicates that DNA contamination was introduced at some point during the experiment.

General notes

- Work with your lab partner for this experiment.
- Each partner should prepare a cell suspension from their own isolated strain. This template will be used for both the 16S PCR assay and the *rpoB* assay.
- One partner should prepare PCR reaction mixture for amplifying the 16S rRNA gene; the other partner should prepare the mixture for amplifying *rpoB*. Prepare enough of your mixture for your sample, your partner's sample, and a negative control (3 reactions total).
- You will use 6 PCR tubes in total. Label each tube with your bench # and a letter (A-F) as shown in the sample ID table below. Fill in the names of the isolates you are sequencing.
- Keep all reagents on ice when not using them.

I. Template preparation

For each person:

1. From your re-streaked isolates, choose one that you would like to sequence. From that isolate, pick a single colony that is well-separated from the others. On the back of the plate, use a permanent pen to circle it and label it. Enter the isolate's name into the Sample ID table.

2. Pipet 20 uL of PCR-grade water into a microcentrifuge tube.
3. Using a sterile inoculating loop, pick up a single colony of your isolate from the culture plate and transfer it into the tube of water. Swirl the loop to dislodge the cells. Discard the loop.
4. Close the cap of the tube and vortex it at high speed for 20 seconds. Use this water-with-cells suspension as the **template** for both PCR reactions.

II. PCR reaction setup

For each assay (16S and *rpoB*):

1. Label 3 PCR tubes as indicated on the sample ID table.
2. Following the recipe for ***3 reactions*** (next page), combine water, 2x master mix, forward primer, and reverse primer in a microcentrifuge tube. Close the tube and flick several times to mix well without generating too many bubbles.
3. Aliquot 24 uL of reaction mixture into each of the 3 appropriate PCR tubes.
4. Into each PCR tube, pipet 1 uL of prepared template, following the sample ID table. For the no template control, add 1 uL of PCR-grade water.
5. Seal the tubes and place in the appropriate thermal cycler: there is one thermal cycler for the 16S assay and one for the *rpoB* assay. Then add your information to the grid sheet: bench #, tube letter, and your initials. Throw away all other tubes.

Sample ID table

Bench #	Letter	Template (isolate name)	Gene
	A		16S
	B		16S
	C	No Template	16S
	D		*rpoB*
	E		*rpoB*
	F	No Template	*rpoB*

PCR Set Up 1: V4-V5 region of 16s rRNA Gene

The primers used for this reaction will amplify the V4-V5 hypervariable region of the 16S rRNA gene, found in bacteria and archaea.

This protocol comes from the Earth Microbiome Project, which seeks to catalogue microbial diversity around the world: http://www.earthmicrobiome.org/protocols-and-standards/16s/

Primers:

Forward - Primer 515F: GTGYCAGCMGCCGCGGTAA

Reverse - Primer 926R: CCGYCAATTYMTTTRAGTTT

Product size: ~411 bp

Recipe for PCR reaction mixture

Reagent	**1 Reaction (uL)**	**3 Reactions (uL)**
PCR Grade Water	13.0	**39.0**
PCR Master Mix (2x)	10.0	**30.0**
Primer 515F (10 µM)	0.5	**1.5**
Primer 926R (10 µM)	0.5	**1.5**
Total	24.0	**72.0**

With the addition of 1 uL of template, the final reaction volume in each tube will be 25 uL.

Thermocycler conditions: Methylo-16s

	Step	**Temp (°C)**	**Time (min)**	**repeats**
1	Initial denature	94	10:00	x1
2	Denature	94	0:45	x35
3	Anneal	50	0:45	
4	Elongate	72	1:30	
5	Final Elongate	72	5:00	x1
6	Store	10	forever	

PCR Set Up 2: *rpoB* gene, encoding β subunit of bacterial RNA polymerase

The primers for this reaction are specific to the genera *Methylobacterium* and *Methylorubrum*. They will not amplify the gene in other organisms.

This protocol was developed by Jean-Baptiste Leduc, Université du Québec Montréal, specifically for studying methylotroph diversity on tree leaves.

Primers:

Forward - Primer Met02-352-F: AAGGACATCAAGGAGCAGGA
Reverse - Primer Met02-1121-R: ACSCGGTAKATGTCGAACAG

Product size: ~768 bp

Recipe for PCR reaction mixture

Reagent	**1 Reaction (uL)**	**3 Reactions (uL)**
PCR Grade Water	10.9	**32.7**
PCR Master Mix (2x)	12.5	**37.5**
Primer Met02-352-F (10 µM)	0.3	**0.9**
Primer Met02-1121-R (10 µM)	0.3	**0.9**
Total	24.0	**72.0**

With the addition of 1 uL of template, the final reaction volume in each tube will be 25 uL.

Thermocycler conditions: Methylo-rpoB

	Step	**Temp (°C)**	**Time (min)**	**Repeats**
1	Initial denature	94	10:00	x1
2	Denature	94	0:45	x35
3	Anneal	60	0:30	
4	Elongate	72	1:30	
5	Final Elongate	72	10:00	x1
6	Store	10	forever	

Appendix 5. An abstract-style writeup assignment as a summative assessment for Methylothon, given in a Biotechnology class

Methylobacterium Abstract

Your completed written work cannot exceed 2 pages.
This does not include the phylogenetic tree or reference page.

Your final work must include the following areas:

1. **Introduction:** An intro to *Methylobacterium* that addresses the following:
    a. What *Methylobacterium* are (identify characteristics that define the classification (see https://microbewiki.kenyon.edu/index.php/Methylobacterium for help).
    b. Major current work (research and/or application) involving *Methylobacterium*
    c. Purpose of our experiment
2. **Protocol/Methods:** A brief description of the experiment - include all major work involved in this lab, including stuff we did not directly do. Be brief but concise - a fellow scientist should know exactly what you did but you do not need to provide all the details.
    - **Note**: Although we did not do a lot of actual lab work, I would like you to describe the things we did cover, which will include the **sample collection and culturing** (leaf press and culturing after), **DNA extraction** (we did not fully cover this so very broadly what does DNA extraction mean here), **PCR** (which we did not do but you should be able to describe broadly how this should work, remember we would be looking to amplify certain sections of our microbes genomes to be able to compare them), and **sequencing** (which we did not do but you should be able to briefly explain what would happen in a sequencing reaction if you did complete it (not the lab part but how does sequencing work, broadly). **Just to recap -- you do not need to explain HOW these things would work in the lab but conceptually what is the purpose of each.**
3. **Results:** Attach a screenshot of your phylogenetic tree and then briefly describe your portion of the phylogenetic tree. This section is just you presenting what you found.
4. **Conclusions:** Provide a 2-3 sentence summary on what the results actually mean. **Address this idea in your conclusion:** Microbes are incredibly diverse and we know almost nothing about them. (How does your work relate to this statement and either support it or refute it?)
5. **Research Question**: Answer a question of your own that relates to the work we did. This should be related to the topics we covered in class but can be about anything of interest to you. This must involve outside research of primary or secondary sources. We will cover what this means in class. See this document for your opening question: Opening Questions (you can use one from here or a question of your choice). You should integrate your research into the work we did as best you can. Example: if your question deals with applications, then make sure to make the connection between applications and the work we did.)
6. **Citations/References**: Include a reference section that is set up using APA format.

Abstract Scoring Guide	

Introduction includes ● A thorough but concise overview of *Methylobacterium* ● Discussion of current work/research and use of *Methylobacterium* ● The purpose of our experiment	**/20**
Protocol includes ● A basic explanation for each major section of the experiment (that does not get bogged down in details).	**/5**
Results ● Includes observations from the parts of the experiment we did (from the sampling and plate observations; you should include pictures.) ● Discusses and cites specific data (from the sampling, the plate observations and bioinformatics)	**/20**
Conclusions ● Summarizes what your tree tells you and its implications to microbial diversity	**/25**
Research Question ● Question is interesting and related to the topic ● Answer is thorough and uses primary and secondary sources to support it (you do not have to fully answer the question if it is something that is still an ongoing area of discovery)	**/25**
References ● **Only 1 verified reference is required.**	**/5**
TOTAL:	**/100**

Appendix 6. A Team Quiz as a summative assessment for Methylothon, assigned in an International Baccalaureate Biology class.

QUIZ FOR THE METHYLOBACTERIUM WEEK

____/4 PROFICIENCY POINTS

From Dr. Martinez-Gomez, you have learned that some species of *Methylobacterium*, but not all, require rare earth elements to survive. We provided you with two culture plates designed for isolating methylotrophic bacteria; the plates were identical except that one contained the rare earth element lanthanum and the other did not. You used the plates to isolate microorganisms from a plant leaf. ***Your final assignment is to work with your team to complete a 1-2 page limit (double spaced with 1 inch margins) writeup with the following:***

1. Background information (considering citing Monday's reading here, Ceci's presentation, presentations from the week)

2. Describe a straightforward **research question/hypothesis** that can be tested through the experiment that you carried out. NOTE: Your team needs to coordinate this BEFORE you choose your leaves.

3. BRIEFLY summarize your **procedure**.

4. Describe your **results/observations** of microbial growth on your 4 culture plates at the end of the experiment. Were your observations sufficient to address your hypothesis?

5. If so, please explain your **conclusions** and the evidence supporting them. Consider mentioning **sources of error, limitations to the experiment, improvements, and/or an extension.**
If not, that's okay! It can be very difficult to conduct rigorous scientific research in the space of 1 week! In this case, please explain why your results were inconclusive, and what you think we would need to do to change the experiment (assuming we had more time, more resources, and the ability to work in a lab) in order to truly address the

hypothesis you posed about *Methylobacterium* and rare earth elements. You may choose to discuss some of the following concepts:
- the methods we use to identify microbes
- what we know microbial growth requirements
- the importance of replication
- qualitative versus quantitative observations
- the differences between getting results that you don't expect, getting insufficient information from an experiment, and experiment totally failing

Make a title and attach your <u>one-two page response</u> to this google classroom assignment. Please include the names of <u>all</u> team members.

Rubric:

	4	3	2	1
Introduction	Presents a clear summary of the aims of the study and its significance. Includes a straightforward **research question and hypothesis.** Briefly describes experimental design. Probably includes **one or more references to supporting sources*.**	Either lacks clarity or is missing one of the primary elements.	Weak or missing primary elements	No real introduction.
Materials and Methods	Gives the reader a clear picture of the **methods and materials** used. Does not use prescriptive language. Uses specific, not general, terminology. Detailed, step-by-step procedures are clearly referenced. Avoids long, redundant descriptions	Some methods are presented so briefly and/or vaguely that it is unclear how or why they were done. May be some written as a protocol rather than a description.	Some methods are omitted; others are presented in a piecemeal, vague form.	Methods barely mentioned.
Results	All **results** are clearly presented, with a logical sequence. Controls are	Some data may be missing.	Data is presented haphazardly. It is sometimes not	No logical connection between

	clearly indicated (*if applicable*).		possible to tell what material or procedure was used to obtain the data.	methods and data. Irrelevant data may be included, and relevant data left out.
Discussion	It is clear that the methods and results have been understood. The results (including controls) are related to the questions posed and **analyzed** for their effectiveness. Scientific reasoning is included. **Possible explanations for inconsistencies and/or unexpected results are given.**	There may be some lack of clarity. Did the writer understand why certain methods were used, and how the results could shed light on the questions asked? Incomplete analysis of inconsistencies and unexpected results.	Very little analysis of the results. Statements are vague and general. Inconsistencies are explained by 'human error' or something similar.	Mostly a restatement of results. No analysis given. No recognition of error sources. No understanding of controls.
Cohesiveness	It is clear that the report covers a group of related procedures with a clear set of goals.	Sometimes the goals are not clearly related to the report. Some fragmentation occurs, with methods and results apparently unrelated to each other.	Transitions are abrupt. Each day's work seems unrelated to the next's. Aims are not clearly present throughout.	Disjointed. No flow. Very little use of headings, or explanatory sentences.
Spelling and grammar	No spelling or grammatical errors	An occasional error.	Apparently not proofread for errors.	Frequent grammatical errors: incomplete sentences, tense changes, misspellings.

*Please cite any sources in APA. For example, if you use the Monday reading about lanthanides, please cite it in APA at the end.

Appendix 7. A flexible-format final assignment combining Methylothon with human ancestry, given as a summative assessment in a Biotechnology class.

Paleogenetics Project

You are assisting a team of anthropologists studying an ancient cave site. Nearby, burial grounds are discovered! These bones look different from known hominins (human species) in the area. A sample of bone containing DNA is given to you to analyze.

Because you are the team's expert in bioinformatics, it is your job to determine whether this sample comes from a known species of humans or an entirely undiscovered species. You'll also determine which hominins this group is most related to (where this group resides on the tree of life) and this group's likely migratory pattern (how these people came to this cave site).

Create a presentation* in which you answer these questions:

1. How will you obtain enough DNA for the analysis?
2. How will you prevent contamination?
3. How will you sequence the DNA?
4. How will you determine if this is a known species or a new species?
5. How is a multiple sequence alignment done, and how is a phylogenetic tree created?
6. How will you determine this group's migration route?

*You may choose to make a video, a slideshow, write an essay, or create a graphic novel. If you have another idea, run it by me, and we'll see if it could work.

Standards which apply to this assignment:

- I can describe the use of common lab equipment and sterile techniques.
- I can describe how DNA is isolated from a sample.
- I can describe the steps of PCR.
- I can explain how DNA is sequenced.
- I can explain what it means to BLAST a DNA sequence.
- I can describe what a computer does to create a multiple sequence alignment.
- I can create a phylogenetic tree using DNA sequences.

- I can describe what a molecular clock is and how it is used.
- I can explain how we determine human ancestry and migration.
- I can describe why mtDNA is a useful tool for studying ancestry.
- I communicate clearly and concisely about science.

Paleogenetics Project Rubric

	Exemplary	Proficient	Developing	Support Needed
I can describe the use of common lab equipment and sterile techniques.	Excellent descriptions of lab equipment and sterile techniques.	Correct descriptions of lab equipment and sterile techniques. There are a few errors or omissions.	Incomplete descriptions: just lab equipment or just sterile techniques.	Incorrect terms or descriptions of equipment or sterile techniques.
I can describe how DNA is isolated from a sample.	Accurate and complete description of DNA extraction.	Generally correct description of DNA extraction. There are a few errors or omissions.	Incomplete description of DNA extraction..	Incorrect description of DNA extraction.
I can describe the steps of PCR.	Accurate and complete description of PCR.	Generally correct description of PCR. There are a few errors or omissions.	Incomplete description of PCR.	Incorrect description of PCR.
I can explain how DNA is sequenced.	Accurate and complete description of DNA sequencing.	Generally correct description of DNA sequencing. There are a few errors or omissions.	Incomplete description of DNA sequencing.	Incorrect description of DNA sequencing.
I can explain what it means to BLAST a DNA sequence to create a multiple sequence alignment and a phylogenetic tree.	Accurate and complete description of DNA databases, the BLAST tool, MSA, and tree creation.	Generally correct description of DNA databases, the BLAST tool, MSA, and tree creation. There are a few errors or omissions.	Correctly describes two of the four.	Mentions DNA databases but not how they are used.
I can describe what a molecular clock is and how it is used to study human ancestry and migration.	Accurately explains how mtDNA functions as a molecular clock and how haplogroups are used to study human ancestry and migration.	Correctly describes how haplogroups are used to study human ancestry and migration.	Correct description of human ancestry or human migration.	Mentions mtDNA or haplogroups, but does not describe what they are or how they are used.

I communicate clearly and concisely about science.	The work flows, is clear and concise, and articulately explains each element.	The work is clear and concise. A few incorrect terms or grammatical problems.	The presentation gets off topic. Distracting errors in terms and grammar.	The work is muddled and contains many errors.
I turn in my work on time.	Assignment is turned in early	Meets deadline	Within 2 hours after deadline	Turned in late

Appendix 8. Plants sampled by students for leaf presses during Methylothon 2021.

Abutilon sp. (Flowering Maple)
Abutilon sp. (Mallow)
Acacia melanoxylon (Black Acacia)
Acer palmatum (Green-leaf Japanese Maple)
Allamanda blanchetii (Purple Allamanda)
Anthurium andraeanum (Flamingo Flower)
Artemisia vulgaris (Common Mugwort)
Camiella japonica (Japanese Camellia)
Chrysanthemum sp.
Citrus limon (Lemon) 'Meyer,' 'Eureka'
Citrus x sinensis (Orange)
Claytonia perfoliata (Miner's lettuce)
Crassula multicava
Cyclamen persicum (Persian cyclamen)
Delairea odorata (Cape Ivy)
Fern, unknown species
Ficus benjamina (Weeping Fig)
Fortunella japonica (Kumquat)
Fragaria sp. (Strawberry)
Fragaria vesca (Wood strawberry)
Gaultheria shallon (Salal)
Geranium purpureum (Little-Robin)
Geranium sp.
Ginko biloba
Hedera helix (common English Ivy)
Hibiscus sp.
Hoya carnosa (Honey plant)
Impatiens sodenii (Poor Man's Rhododendron)
Ipomoea purpurea (Common morning glory)
Lilaceae sp. (Lily)
Lonicera japonica (Honeysuckle)
Loropetalum chinense (Fringe flower)

Macadamia sp.
Medicago lupulina L, (Black Medick)
Mentha (Mint Leaf)
Nepenthes x Miranda (Tropical pitcher plant)
Oxalis pes-caprae, (sour grass, Bermuda buttercup)
Oxalis stricta (common yellow woodsorrel)
Pachira aquatica (Provision Tree)
Parietaria sp.
Pelargonium peltatum (Ivy Geranium)
Persea americana (Winter Mexican Avocado)
Persicaria sp. (Knotweed, Smartweed)
Philodendron hederaceum (Heartleaf Philodendron)
Philodendron laciniatum
Physalis peruviana (Cape Gooseberry)
Pisum sativum (Common Pea)
Pleargonium zonale (L.) L'Her. ex Aiton " (Horseshoe geranium)
Quercus agrifolia (Coast Live Oak)
Quercus sp. (Oak)
Rosa sp. (Rose)
Rubus Armeniacus (Bramble, Himalayan Blackberry)
Spathiphyllum sp. (Peace Lily)
Tagetes erecta (Orange Marigold)
Trifolium sp. (Clover)
Tropaeolum sp. (Nasturtium)
Urtica urens (Dwarf Nettle)
Uvularia grandiflora (Yellow Bellflower)
Vicia faba (Fava bean)

Appendix 9. Example of student work on Bacterial Identification Virtual Lab worksheet.

Methylothon

Bacterial Identification Virtual Lab **Date: ____10 March 2021____Period: ___**

_____/4 POL Points

Team Member Names and Roles:
[redacted]

First, **document managers, make a copy** of this document and share it with your group so that all of you can add to the document at the same time.

Next put your name in the blank:

_____ **scientist:** shares screen and manipulates lab (you must have Flash)

_____ **recorder:** writes down the procedure (tells the story)

_____ **documents manager:** screen captures pictures of the tools, results

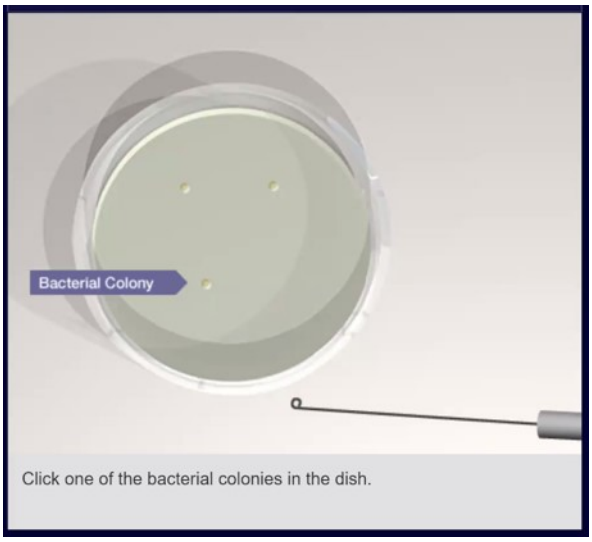
_____ **captain:** writes down what the tools are used to do

Scientist: access the virtual **Bacterial Identification Lab** and share your screen. Follow the instructions and click the prompts.

Recorder, documents manager, and captain: split your screens so that you can watch the lab AND edit this document at the same time.

A. Sample Preparation

1. --Documents manager: **upload** a picture of the sample plate with bacterial colonies.



2. --Captain: What is the purpose of centrifuging the sample? Which part of the centrifuged product are we collecting?

We centrifuge to separate the sample by weight in order to get rid of the proteolytic enzyme. By centrifuging the sample, the cellular debris is removed from the sample. We are collecting the DNA liquid that is left in the centrifuge.

Everyone:

3. --In Methylothon, we isolate our *Methylobacterium* with a specific kind of culture plate. What kind of medium do we use, and why?

We use a selective minimal culture medium in order to only allow certain organisms to grow. In the case of the lab, our specific medium is designed to allow for only the methylobacterium to grow.

4. --Why is it important to select a sample from one colony, rather than collecting from multiple colonies?

Different colonies could have different methylobacterium on them and we only want to see the methylobacterium from one colony, not many.

B. PCR Amplification

5. --Recorder: list the times and temperatures for each step of the PCR process

Denaturing-Temperature: 95°C, Time: 5 minutes on initial cycle. Annealing-Temperature: 5°C below Tm of primers; no lower than 40°C. Time: 30-45 seconds.

Extension- Temperature 72C, Time: ~1 min/kb of expected product; 5-10 min on last cycle.

6.--Captain: List the substance in the red, green, and blue vials and explain why they are used in the experiment.

Red: the red is the PCR master mix and is used to keep the pH constant Green: the green is the positive control DNA, and is used as a control Blue: the blue is deionized water, and it is used as a negative control

7.--Documents manager: capture images of the PCR animation

Everyone:

8. --Explain the significance of the temperature required by each step.

Denaturation - Temperature: 95°C is used to separate the double helix DNA strands Annealing - 60°C is used so the primer can bind to the single stranded DNA (annealing) Extension - Temperature 72C is used for the creation of new strands of DNA made using the original stands as templates.

9.--What is the purpose of using heat-stable DNA polymerase in PCR reactions?

Since 95°C is the denaturing temperature, and 5°C below Tm of primers is annealing, the polymerase must be able to not denature in these temperatures due to the role it plays in separating, amplification, and reconstruction.

10. --How many copies of DNA are generated at the end of 30 cycles?

There are 1 million copies at the end of 30 cycles.

C. PCR Purification

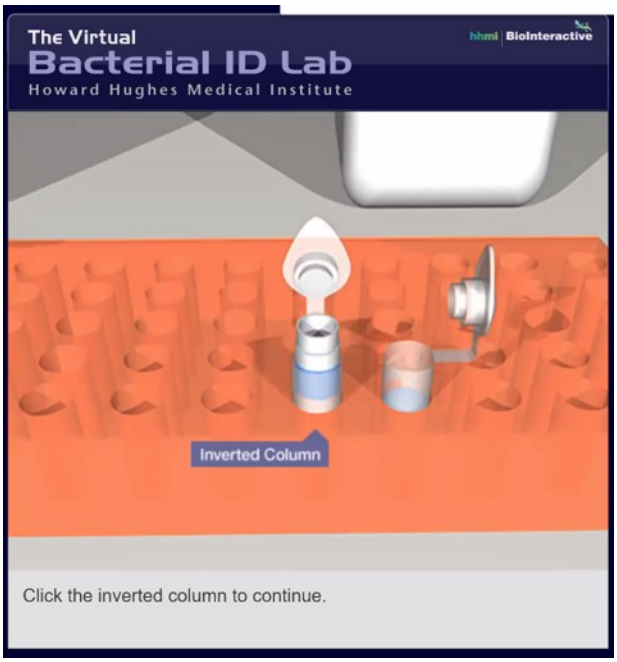
11. --Recorder: List the steps in PCR Purification

1. PCR Buffer PB (400ul) 2. Column binding 3. Washing 4. Drying 5. Elution 6. Pure DNA fragment

12. --Captain: What additional substance are we adding to the column? Why?

We are adding a buffer solution to the column because it makes sure that the DNA can separate from the column and go into the collection tube.

13. --Documents manager: take an image of the column after it is transferred to a new test test tube. Which tube contains the supernatant? What is in the supernatant?



The old test tube would contain the supernatant since the column was removed and placed into a new test tube, separating the DNA and leaving the supernatant at the bottom of the old test tube. The supernatant is everything but the DNA.

Everyone:

14. --Explain the results we expect to see in each of the three lanes of our gel electrophoresis and why these expected results would indicate a successful PCR reaction.

One lane would be a negative control only containing water.
The middle lane would be a positive control containing the PCR known product.
The last lane would be for the unknown sample.

The results that would indicate a successful PCR reaction would include a positive control reaction.

15. --What alternative to gel are we using to purify our product?

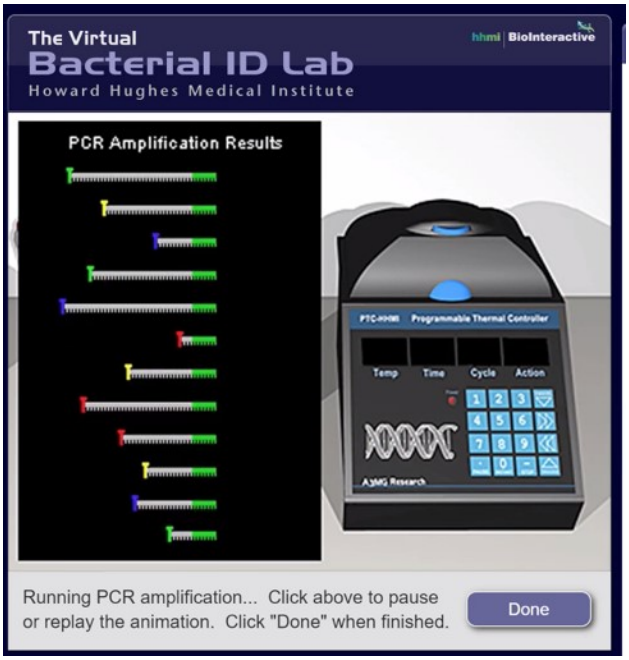
Compact microfilters is an alternative to gel.

D. Sequencing Prep

16. --Captain: Explain what is in the "Sequencing brew" in the blue and green tubes.

The green and blue tubes had different buffers and primers in each tube in addition to DNA polymerase and fluorescence tagged terminators.

17. --Documents manager: Take a picture of the different length sequences. **Paste** it below.



Everyone:

18. --Why do we use multiple primers in PCR cycle sequencing for long sequences?

Using multiple primers allows many short, overlapping pieces of DNA to be put together to find the complete sequence.

19. --What is the significance of using primers that bind to conserved regions of the 16s rDNA gene?

This allows them to bind to the sequence regardless of bacterial source.

20. --Describe what occurs in the tube containing the primer 651R.

In this tube, the DNA strands bind to the primer and have one fluorescence-tagged terminator at the end that they don't bind to the primer.

E. DNA Sequencing

21. --Documents manager: capture a picture of the chromatogram. **Paste** it below.

Everyone:

22. --Explain why DNA molecules move from one end of the gel to the other during gel electrophoresis.

Gel electrophoresis applies an electrical current to the tube, and since DNA molecules are negatively charged, they will move through the tube towards the positively charged end, smaller pieces moving faster.

23. --What is the purpose of fluorescent markers in DNA sequencing?
In what order do DNA fragments move through the sequences (i.e. why might one sequence travel faster than another)?

The fluorescent markers are used when the DNA fragments are pushed through a laser beam and this interaction sparksdecistors to recognize the fluorescent markets. The DNA fragments are pushed through the gel based on size.

F. Sequencing Analysis

24. --What species of bacteria does the DNA belong to?

The DNA belongs to the Bartonella henselae species.

Documents managers, please turn in this assignment on google classroom. Everyone else, please mark this as done WITHOUT turning it in.

Appendix 10. Example of student work in Bioinformatics Virtual Lab worksheet.

Name: Date: 3-12-21 Period:

Bioinformatics Tutorial Worksheet (Thursday Class and Lab)

_____/4 POL Points

In this BLAST tutorial, we're going to be downloading files, uploading files, and moving from website to website frequently. There's a lot of moving parts here, so this worksheet (hopefully) will help make some of the nitty gritty website-wrangling steps clearer. Of course, feel free to stop by office hours if you have any questions!

1. Getting your sequences for BLASTing

In this lab, you'll be BLASTing an unknown sequence collected from a plant leaf by last year's Methylothon-ers! You'll also need reference sequences for building a phylogenetic tree in the final step.

1. Navigate to http://methylothon.com. Find your individual assigned DNA sequence here.
2. From the main menu, go to "Methylothon 2021: student page"
3. Sign in with the password PinkBacteria2021
4. Follow the instructions on the website under "DNA sequences you need for the Bioinformatics Tutorial".

2. BLASTing your unknown

What's happening in this step: In this step, we're using NCBI's BLAST tool to compare our unknown DNA to a database of known sequences. We're going to gather the top hits (known sequences in the database most similar to the unknown sequence) and construct an alignment.

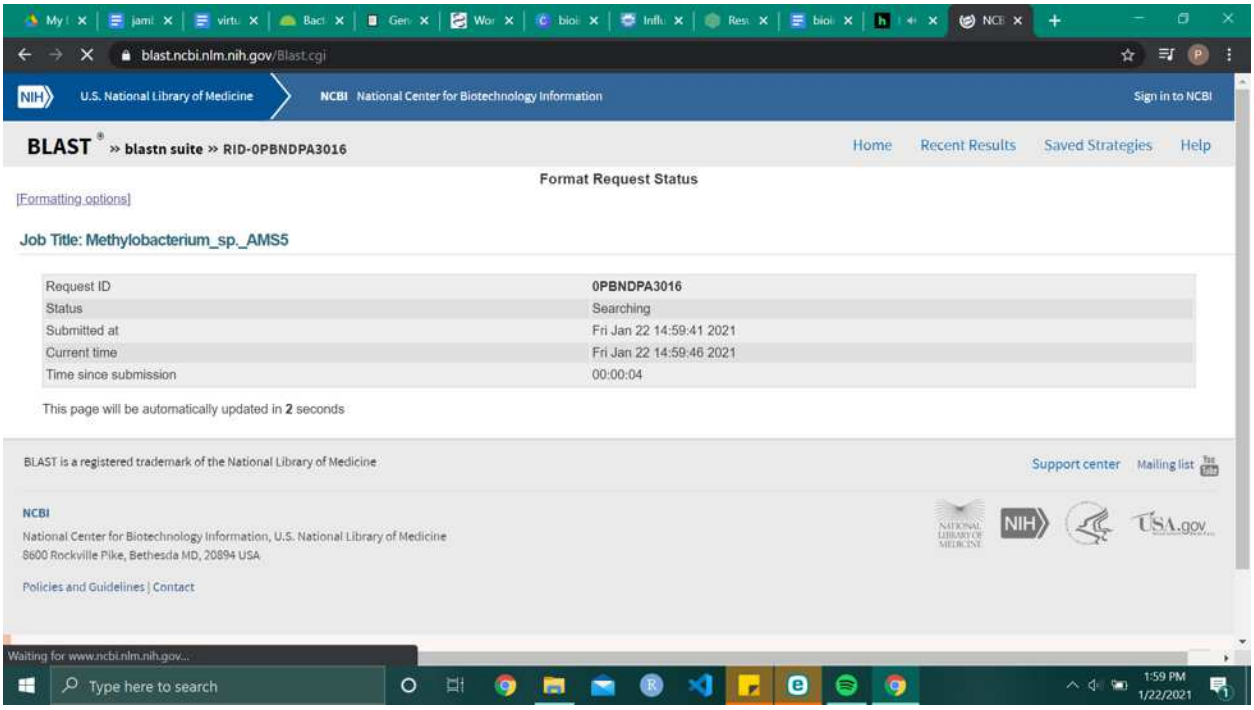
first, navigate to the BLAST homepage at https://blast.ncbi.nlm.nih.gov/Blast.cgi

scroll down and click the "Nucleotide BLAST" button.

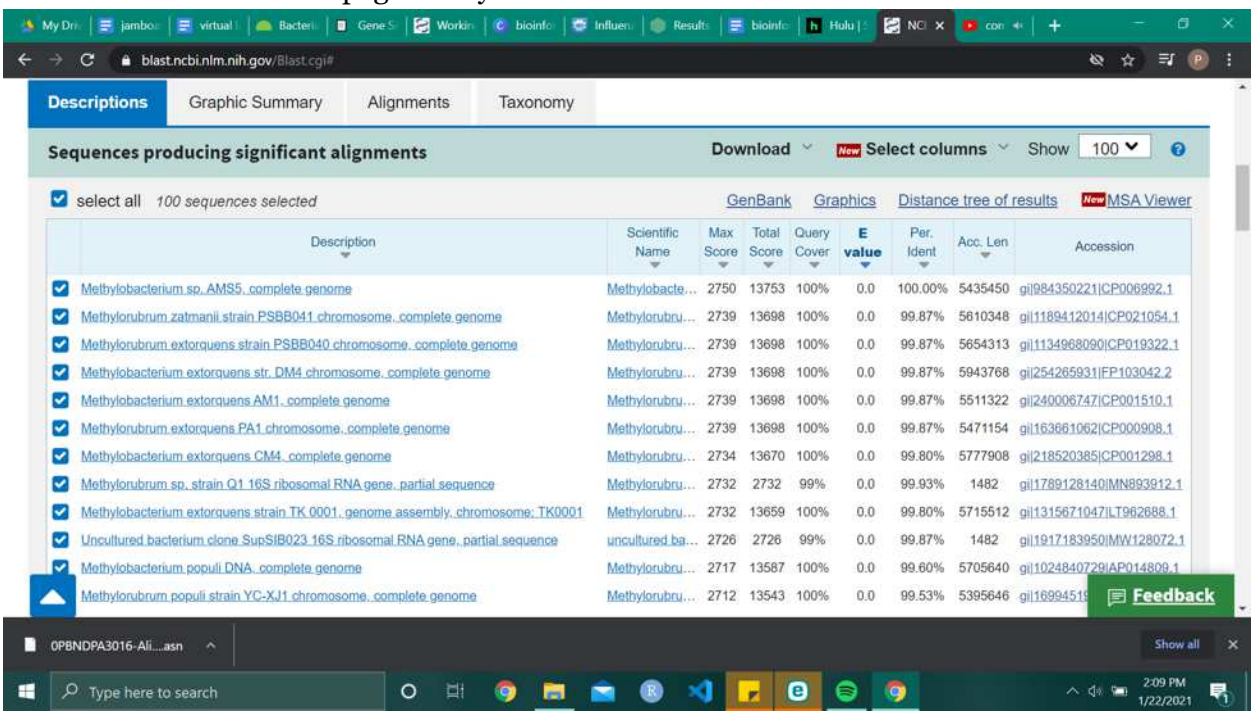
your screen should look like this:
click the "choose file" button (circled in red) and upload your unknown.

scroll down, and make sure that the "megablast" option is enabled.
then, click BLAST to see your results!

****sometimes, BLAST will take a while to search and you'll get stuck on this loading screen for a minute -- this is normal!****



scroll down on the results page until you see a list of match results:



uncheck the "select all" button, and manually check the first 4 results.

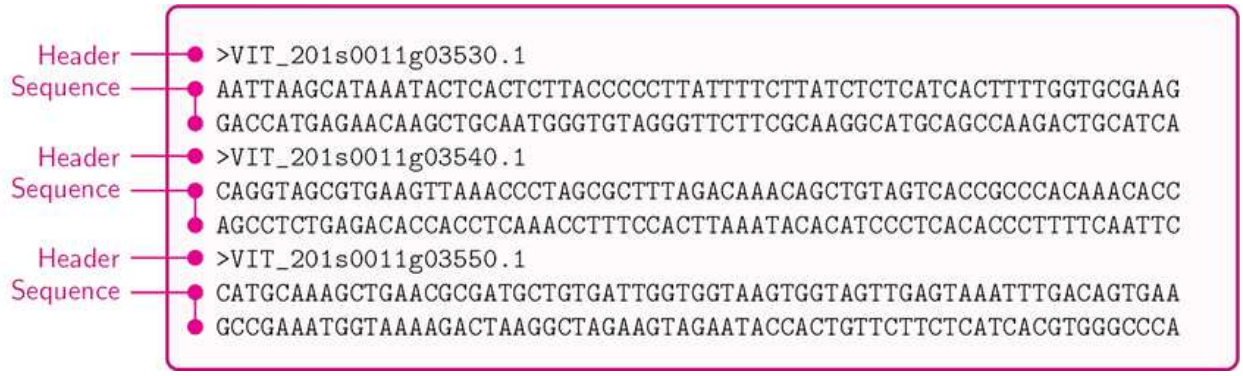
click "download" and download the *aligned* FASTA sequences. We choose this option so that we download the only portions of the sequences that match our query sequence in length.

A. STOP AND ANSWER: What kinds of organisms are the top hits for your sequence? Is this what you expected? Explain. If you're curious, you can click on any of those hits and follow the weblinks to the Nucleotide database entry. This entry sometimes contains information about where the sequence came from.

The organisms that were top hits for the sequence of my DNA sample were all sequences of the ribosomal RNA gene of methylobacterium. This is what I expected because the lab that was done in order to obtain this DNA was one that was intentionally designed to culture methylobacterium. The sample this student collected from their plates contains DNA of the methylobacterium that this student cultured from their leaf press.

2.5 Reformatting

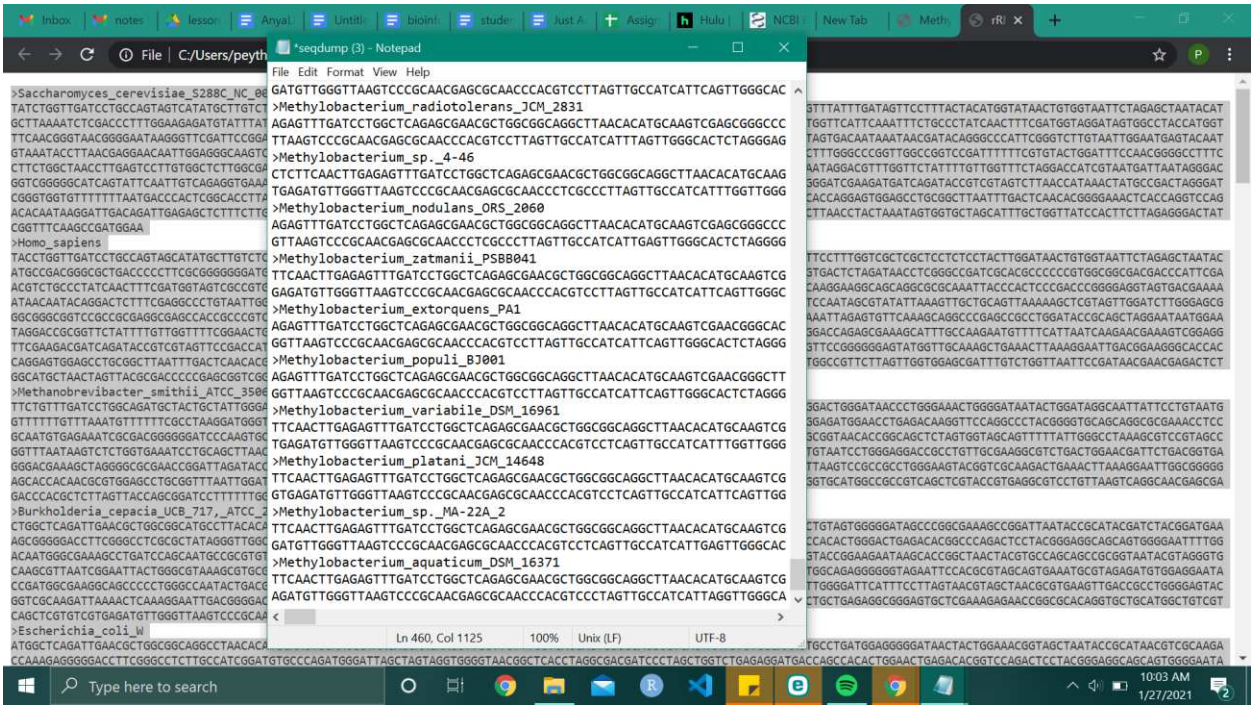
Before we visualize relatedness among our three different DNA sources (your original unknown, top BLAST hits, and reference sequences), we can make the process of managing computer programs much easier if we condense all of our sources into a single file. The files we'll be bundling together are all FASTA files. A FASTA file is a filetype that displays species name / information followed by genetic code. An example of the FASTA format is provided below. FASTA files are often saved with the file extension ".fasta" or ".fa" to help DNA-sequence-editing programs recognize what they are, but they can also be saved with the extension ".txt", which is what we will do here.



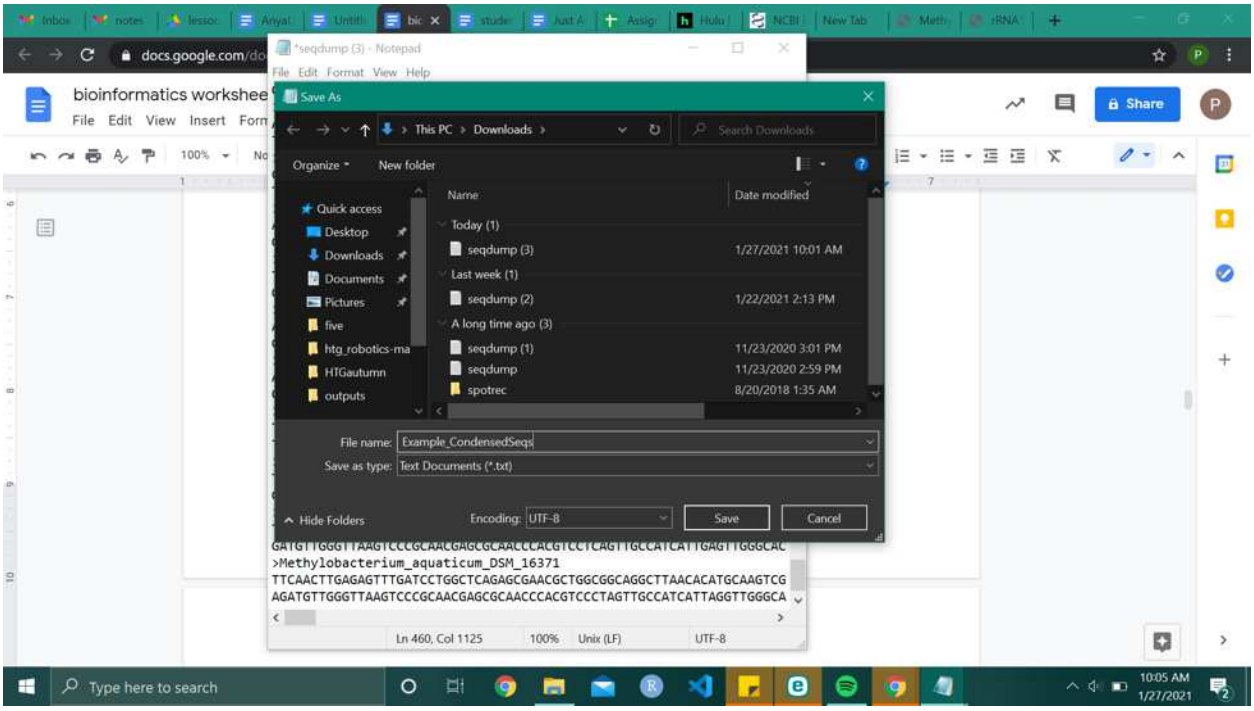
We're going to create a file that has all the DNA we want to investigate in a single file. To do this, open up a plain text editor (for instance, "Notepad" on Windows devices, or "TextEdit" on Mac). Then open up the following:

1. your BLAST top hits (called "seqdump.txt" unless you renamed it)
2. the reference sequences you downloaded from the project website (called "Methylothon_2018-2019_reference_sequences.txt")
3. your mystery sequence.

Paste in all of your sequences of interest from the three different files into a new document, one right after another as seen in the example image above (eg. no blank lines between sequences).



Save your new document as [LASTNAME_CondensedSeqs].txt.



3. Multiple Sequence Alignment

What's happening in this step: In this step, we're going to visualize the alignments of our unknown and the closest hits, as well as our reference sequences.

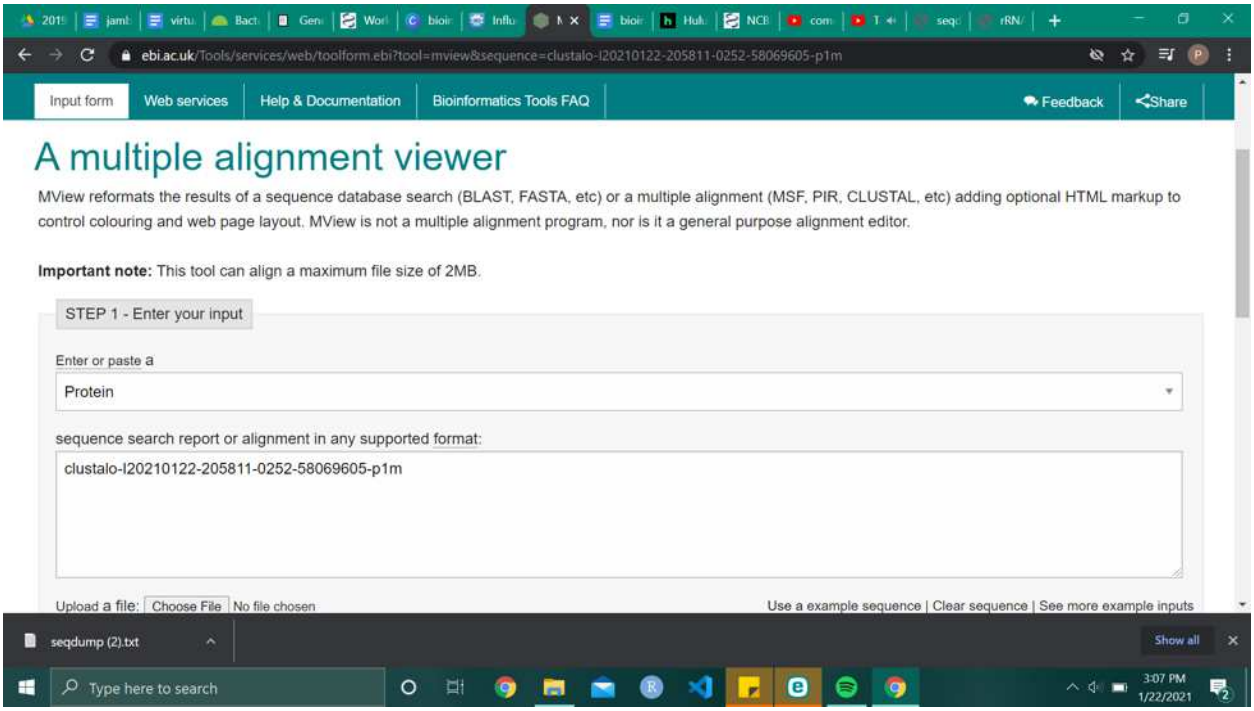
We're going to use EMBL's Clustal Omega tool at this link:
https://www.ebi.ac.uk/Tools/msa/clustalo/

First things first, we're going to change the set from "PROTEIN" to "DNA"
upload your combined FASTA file, scroll down and click submit!

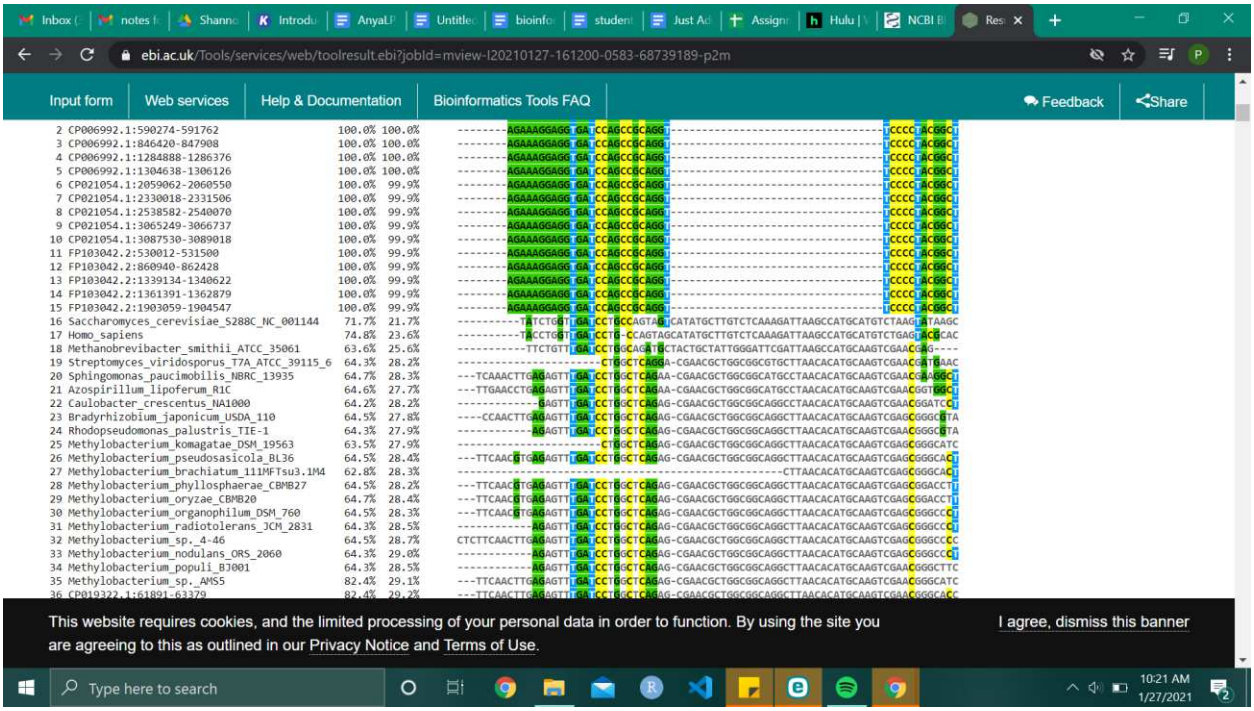
to view a MSA of your results, click the "results viewer" tab.

Scroll down to "View in MView" and click it!

the MSA generator will be autopopulated with the data from previous step...



so just scroll down and click submit!



B. STOP AND ANSWER: Inspect the alignment.
Do the sequences generally look pretty well aligned? How can you tell?

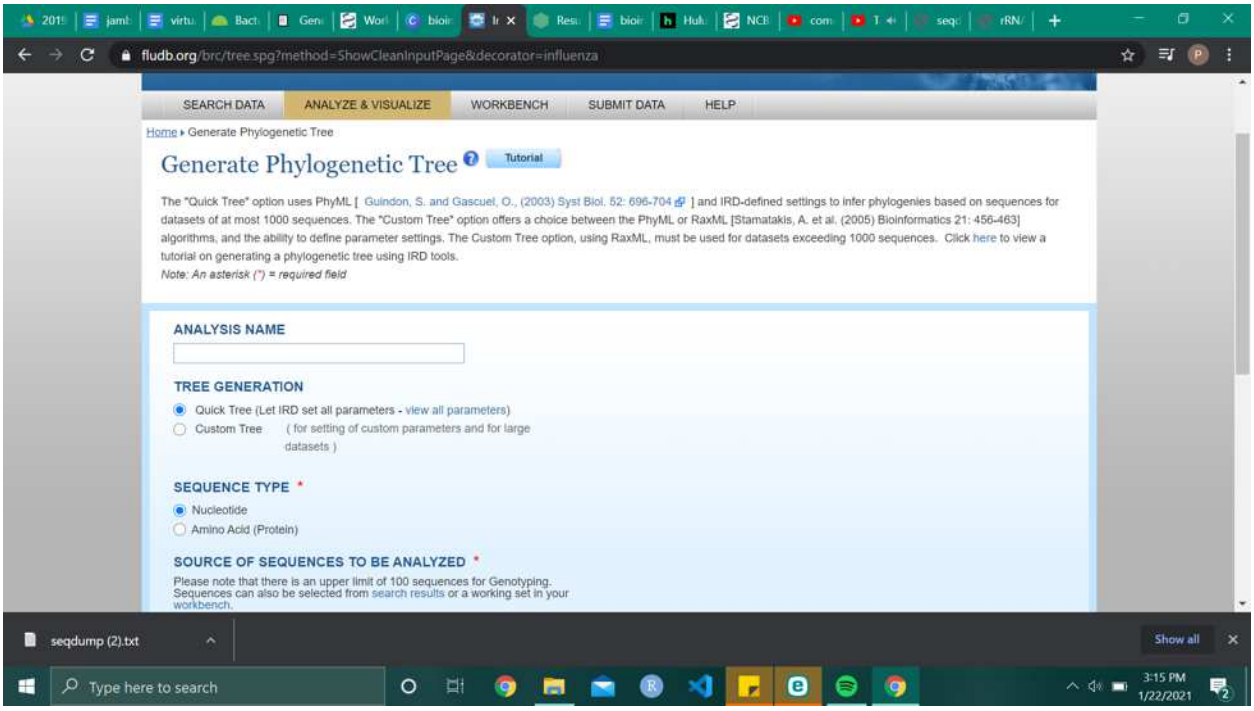
Yes, the sequences look well aligned, made evident by the highlighted columns of nucleotides. This is especially true for certain sets of samples compared to others, as some groups have less alignment than others.

C. STOP AND ANSWER: Just from looking at the alignment, can you guess which sequences are most distantly related from your mystery sequence?

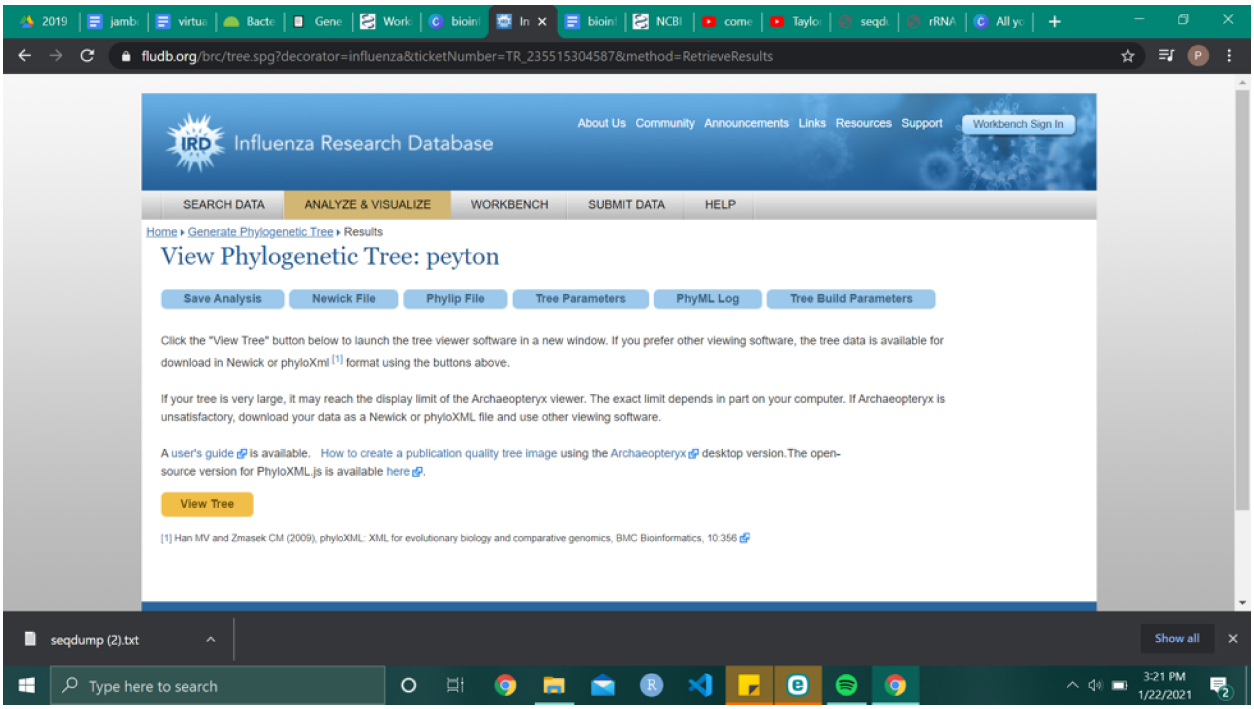
There are certain methylobacterium samples that align less than others, and these sequences, along with the sequences pertaining to species that are not methylobacterium, are most distantly related from my mystery sequence.

4. Phylogenetic Tree

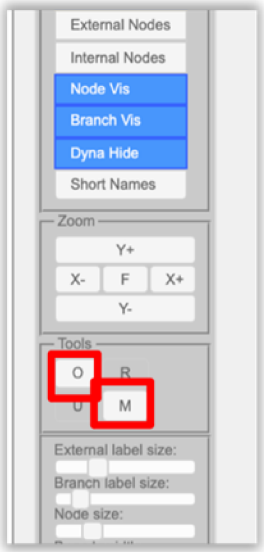
What's happening in this step: Now, we're going to create a visualization of relatedness among the top hits and the reference sequences.
To do this, we'll be using the Influenza Research Database's tool found at this link: https://www.fludb.org/brc/tree.spg?method=ShowCleanInputPage&decorator=influenza



same as before, we'll be uploading our combined FASTA. make sure that you check unaligned FASTA format, and click submit!



click view tree, and check out your phylogeny!



There are two quick manipulations you should carry out to make the tree easier to compare with others in your class.
First click the "M" button for "mid-point re-root" - this makes a guess at where to place the ancestor.
Then click the "O" button for "order all nodes" - this prettifies things by ordering branches by length (but doesn't change the interpretation of the tree).

Enter the name of your top blast sequence in the search bar to highlight it, and download an image of your tree with your highlighted sequence.
then you're ready to upload the image to the class slideshow!

SUBMIT YOUR TREE: ***Follow this link to the class slideshow*****. Find the slide with your name on it and import the image of your phylogenetic tree. Add the information about your isolate's host plant.**

D. STOP AND ANSWER: According to the tree, what are your mystery organism's closest relatives? Is this what you expected? Explain.

My mystery organism's closest relatives are other sequences of the ribosomal RNA gene of methylobacterium, according to BLAST results. According to other reference species on my tree, the other closest relatives to my isolate are other varying species of methylobacterium including *brachiatum, pseudosassicola*, and *phyllosphaerae*. *Yes, seen as my isolate is a methylobacterium, it was expected that its closest relatives would also be.*

E. STOP AND ANSWER: Once student results begin to fill the powerpoint, answer the following questions:

a) Find another student who has an isolate that appears genetically identical to yours.
What is the name of that isolate? 2-ZX
Compare sequences-- just give them an initial glance. Do they look like they might be the same? Explain.

No, these sequences aren't exactly the same, but they are very similar in terms of alignment. Many nucleotide bases align with each other, with a few exceptions of substitutions and gaps. The differences increase towards the end of the sequence.

If you really wanted to know for sure whether they were exactly the same all the way through, how would you do that?

In order to do this, I would simply repeat this process with the sequence of my isolate and this similar isolate. I would combine the sequences into a file and run it through the website that highlights corresponding pairs and similarities. The more highlighted these two sequences are, the more similar, and if they are completely highlighted, that means they are exactly the same all the way through.

b) Find another student who has an isolate that is not identical, but looks from your trees to be closely related.
What is the name of that isolate? 6-CL
Describe how you know that your sequences are closely related but not identical. You can base this claim purely on the phylogenetic tree, or also on the DNA sequence or BLAST results.

I know this sequence is closely identical to mine solely based on the phylogenetic tree. This isolate's closest relatives are also partial ribosomal RNA gene sequences, which are the closest relatives to my isolate as well according to BLAST results. Furthermore, at a node only one further back on 6-CL's phylogenetic tree, it branches off to other species of methylobacterium from the reference list, which are also close relatives to my isolate. This includes methylobacterium *brachiatum, pseudosassicola*, and *phyllosphaerae*.

c) Find a student who has an isolate that is not *Methylobacterium*
If your own isolate was not *Methylobacterium*, find a different student!
What is the name of that isolate? 4-MC
What kind of organism do you think that isolate is? Sphingomonas bacterium

d) Find a student who has an isolate from the same plant species as yours. [Note: not all plants had two isolates, so if you don't find a match, skip this question.]
What is the name of that isolate? 1B-CC
Are your isolates genetically similar? Explain how you know.

Yes, our isolates are genetically similar. Both our isolates are from the same plant species, which means the cultured organism is likely to be the same. This is confirmed by our closest relatives, which according to BLAST are the same. Not only this, but the references are organized in the same locations, with other species of methylobacterium similarly related to our isolates.

e) BONUS: Did anyone in your class find something that could potentially be a novel species of *Methylobacterium*? (Assume that the reference sequences we provided include all the known species of *Methylobacterium*, which is quite an assumption.) Explain how you would recognize that.

The isolate 6A-RM is its own outgroup when compared to methylobacterium, despite still being placed closely on the phylogenetic tree. This could possibly signify a novel species of Methylobacterium having been discovered.