MCB3637: Practical Methods in Microbial Genomics, Fall 2021

Dr. Jonathan Klassen - jonathan.klassen@uconn.edu

Tues & Thurs, 12:30-3:30 BCH 202

Course Philosophy:

Welcome to MCB3637, Practical Methods in Microbial Genomics! The goal of this course is for you to understand how to generate and use microbial genomes, starting with raw data and moving through genome assembly, annotation, and analysis. You will also learn basic computational biology skills, particularly the phenomenal advantages of UNIX-style commands for "Big Data" biology and basic scripting using Python. Given the revolution that has occurred to make DNA sequencing incredibly cheap and widely available, skills to deal with these data are essential for working in the biological sciences. It is my goal to introduce you to these vital skills during this class.

My teaching philosophy follows the classic saying: "give someone a fish and you feed them for a day; teach someone to fish and you feed them for a lifetime." Accordingly, although I will always present background concepts for everything that you do in this class, it will be up to you to apply these concepts in new ways during the exercises and assignments. As a scientist, you will never be given all of the steps to solve a particular problem, but rather you need to be able to figure out for yourself how to obtain the solution and prove that it is correct. Accordingly, most of our time together will be spent solving computational problems, after I provide some brief background giving you all of the tools that you will need for each problem. You are expected to use whatever resources you can find to achieve this goal, including my lecture notes, consultation with your peers and myself, online and other published resources, or anything else you can find. All of the exercises are designed to be achievable during class time, except for some more computationally intensive analyses that have a more "fire-and-forget" character and that may need to be advanced outside of class time. Please consult with me if you find you are taking longer than the allocated class time so that I can help you find strategies to overcome the obstacles that you are hitting. **ASKING FOR MY HELP ON AN EXERCISE OR ASSIGNMENT INCURS NO PENALTY.**

Evaluation: 20% midterm exam, 20% final exam, 10% exercises, 20% projects, 30% in-class participation (e.g., attendance, communication, focus, following instructions). Exams will include practical lab exercises.

Late penalties: Exercises *absolutely will not* be accepted after the start of the following class. These are designed to be completed during the class and submitted before you leave. Each exercise exemplifies the basic concepts discussed in each class and allows you to obtain assistance from your classmates and me. You are welcome to complete these exercises outside of class time, but this will be your responsibility; I will be unable to provide the same level of assistance outside of our scheduled class periods as during class. Late assignments will lose 5% per day as a late penalty and will not be accepted after 1 week following the due date (a "0" mark will be assigned). Assignments are more substantial than exercises and will synthesize concepts from multiple classes. That said, I am very willing to accommodate medical and personal emergencies. Please discuss these (broadly) with me in person so that we can devise an appropriate solution warranted by the situation.

Attendance and Conduct: Because this course is designed so that you use your peers and me as resources to complete the course assignments, I require that you attend as many of the classes as possible. If you have to miss a class, please let me know in advance and recognize that you will still be responsible for completing that day's assignment on your own. I am quite willing to accommodate non-statutory religious observances and other personal class conflicts; please discuss these with me in advance.

All of your conduct in this course is governed by the established UConn Community Standards – see www.community.uconn.edu and/or ask me for further details.

Course Outline:

Week 1: (Aug 31, Sept 2): Introduction to DNA sequencing; NCBI part 1 (Ex #1a)

Week 2: (Sept 7, 9): NCBI part 2 (Ex #1b); The UNIX terminal (Ex #2)

Week 3 (Sept 14, 16): Python introduction and strings (Ex #3); Pattern matching (Ex #4)

Week 4 (Sept 21, 23): Lists & loops (Ex #5, start Assign #1); Tables & dictionaries (Ex #6)

Week 5 (Sept 28, Sept 30): Work day; Work day

Week 6 (Oct 5, 7): Sequencing libraries & assessing data quality (Ex #7); Read trimming (Ex #8)

Week 7 (Oct 12, 14): Midterm review (Assign #1 due); **Midterm exam**

Week 8 (Oct 19, 21): Illumina *de novo* genome assembly (start Assign #2); File parsing

Week 9 (Oct 26, 28): Assembly quality (start Assign #3; Assign #2 due); Oxford Nanopore *de novo* genome assembly (Ex #9)

Week 10 (Nov 2, 4): Work day; Gene prediction (Ex #10; Assign #3 due; start Assign #4);

Week 11 (Nov 9, 11): Genome alignment (Ex #11); Work day;

Week 12 (Nov 16, 18): Genome annotation using BLAST (Ex #12); Annotation using HMMs (Ex #13)

Week 13 (Nov 30, Dec 2): Work day; Read alignment (Ex #14; Assign #4 due)

Week 14 (Dec 7, 9): Genome assembly awards; Review

The final exam will be scheduled by UConn during Dec 13-19 and will take place in BCH 202.

NOTE: I list these topics as a rough outline, and reserve the right to switch topics based on class interest and comfort with the presented material.

Lab website: All lectures, exercises, and assignments will be posted to the course website: http://mcb3637.clas.uconn.edu/.

Bioinformatics server: Lab work can either be done using the computers in BCH 202 to log into the Xanadu server maintained by the UConn Computational Biology Core facility run by the Institute for Systems Genomics (https://bioinformatics.uconn.edu/). The Xanadu server is preloaded with all of the software that we will use during this course. Each student has already been assigned an account for this course that I will email to you after the first class. Any material that you wish to keep must be transferred from this server at the end of the course or it will be lost when your account is closed by the Xanadu sysadmins. I will also show you how to access the Xanadu server from a personal computer so that you can continue your work outside of scheduled class hours. Please let me know if access to a computer is an issue for you so that we can identify an accommodation.

Email and office hours: I will not hold formal office hours. However, I strongly encourage you to email me specific questions. *Please include the text "MCB3637" in the email subject line so that it does not get lost in the morass that is my inbox.* Feel free to also cc your colleagues to both access their wisdom (try as I might, I can't get to every email immediately) and because inevitably others are likely struggling with the same question that you are. Including me in such email discussions will enable me to credit active participants via the participation marks. (You can also forward me a discussion chain after the fact.) *There are absolutely no stupid questions.* Not asking questions is the most common reason for students to struggle with this course. Remember that I expect you to have no pre-existing computational experience and that my job is to guide you through the early steps of your computational career. I am also happy to meet with you individually to hash out more complicated questions – please email me an outline of your question (including "MCB3637" in your subject line) so that we can arrange a time to meet that is compatible with both of our schedules.

COVID-19 safety protocols: Until further notice, to ensure a safe learning environment for everyone, masks/face coverings must be worn at all times when inside buildings, including in the classroom, regardless of vaccination status. If a student is not wearing a mask/face covering, I will ask them to put one on immediately or leave the classroom. Repeatedly failing to follow this expectation will result in a referral to Community Standards. Activities that involve temporarily removing the mask, such as eating or drinking are not allowed; please leave the classroom for such activities. If you see me not wearing a mask/face covering, you should feel comfortable asking me to put one on immediately. Likewise, you should draw my attention to unmasked individuals in our learning spaces so that I can ensure that they follow proper safety protocols. More information about the proper use of masks is available from UConn Environmental Health and Safety at https://ehs.uconn.edu/wp-content/uploads/sites/1131/2020/06/Face-Covering-Handout.pdf.

Although social distancing will not be required inside classrooms for vaccinated individuals, please respect the wishes of others who prefer to maintain social distancing. For their protection, unvaccinated individuals are requested to maintain 6 feet social distance from others.

Remote learning policy: UConn's current policies state that in-person classes should be held when scheduled, and so no online version of MCB3637 will be offered in parallel with the in-person class sessions. Should you need to miss class due to illness or quarantine you will still be responsible for the work done in your absence but should contact me to arrange for appropriate instructions and explanations that will allow you to complete the assigned work on your own, along with modified due dates if appropriate. Absences for COVID will be treated as would those for any other illness. In the currently unlikely event that UConn requires it, we have the capacity to switch to an online format, especially due to our reliance on off-site servers for all work. A personal computer will be required in this event, and accommodations for this should be discussed with me at that time if needed. I have previously offered a related course (MCB5672 – Applied Bioinformatics; part of the UConn MCB Professional MSc program) online that will serve as a prototype for how MCB3637 might accommodate such a sudden shift in teaching modality in the unlikely event that it is needed.