BIOL 47/147 – Genomics: From Data to Analysis

Fall 2020
Meeting on ZOOM and in SLACK
Class Meeting times: Tuesdays and Thursdays 2:50PM-4:40PM
X-hour: Fridays 5:10-6:00PM

Instructor: Olga Zhaxybayeva (aka Professor Z)
Email: olgazh@dartmouth.edu
Office Phone: 603-646-8616

Office Hours: X-hours; Additional times TBD; By appointment, if needed
X-hour: Office hours

Course Goals

With advances in sequencing technology, analyses of DNA, RNA and protein data have become central in many biological and medical research projects. Through lectures, discussion sessions, hands-on *in silico* exercises and in-class problem-solving, we will explore how genomic data analysis advances biological knowledge.

By the end of this course, you will be:

- Familiar with algorithms of nucleotide and amino acid sequence data analysis and able to apply them using specific software and online tools.
- Able to explain reasoning behind the widely-used bioinformatic algorithms and methods.
- Able to analyze and evaluate authentic genomic datasets.

Pre-Requisites

Undergraduate Students: BIOL 13 *OR* BIOL 15 *OR* Permission of Instructor

Overview of the Course Modules

Detailed schedule is on *Canvas*

What is in a Genome?

Learning Objectives:

- Be able to identify components and features of a genome
- Be able to discuss genomic architecture and signatures of prokaryotes and eukaryotes
- Be able to use computational tools to predict genes in prokaryotes and eukaryotes
- Be able to examine genomic features of completed genomes using browser-based software and genomic databases

Essentials of DNA and Amino Acid Sequence Comparisons

Learning Objectives:

- Understand homology and its relationship to sequence similarity
- Understand local alignment algorithms and assessment of significance of sequence similarity
- Master running programs that find and compare similar sequences and learn to interpret the results

Elements of Protein Structure and its Evolution
Learning Objectives:

- Understand the concept of protein motifs and domains
- Learn how to interpret sequence logos
- Learn how to represent motifs as regular expressions and how to run a PHI-BLAST search
- Understand the concept of a position specific scoring matrix and a profile
- Master running PSI-BLAST and RPS-BLAST (CDD) searches

Accounting for Insertion and Deletion of Genetic Material Over Time: Multiple Sequence Alignments. Measuring Genetic Change over Time
Learning Objectives:

- Describe algorithms for aligning more than two sequences
- Master using multiple sequence alignment programs
- Describe alignment program benchmarking techniques
- Be able to assess alignment quality
- Be able to calculate genetic distances from nucleotide and amino acid sequences
- Know the difference between observed and expected actual number of substitutions
- Discuss why it is important to select appropriate substitution models
- Learn how to model DNA (protein) substitutions

Inference and Visualization of Evolutionary Histories of Genes. Inference of Natural Selection at the Molecular Level
Learning Objectives:

- Learn major classes of phylogenetic reconstruction methods
- Learn bootstrapping technique for statistical assessment of tree reconstruction
- Master visualization and Interpretation of produced phylogenetic trees
- Understand how to measure selection in protein-coding genes
- Learn models and statistical tests for detecting selection in molecular data
- Master interpreting results of programs for estimation of selection

How do we determine DNA sequence of an organism? What genes are actually used by an organism? DNA Sequencing and Assembly. Transcriptomics
Learning Objectives:

- Learn basic principles of currently used "next-gen" DNA sequencing technologies

- Be able to calculate how much DNA needs to be sequenced for a genomic project
- Learn basic principles of post-sequencing genome assembly
- Be able to determine how genome "completeness" by examination of how many scaffolds and contigs it contains
- Learn basic principles of the RNA-Seq analysis
- Be able to interpret results of an RNA-Seq analysis

Genetic Make-up of Complex Microbial Communities: Metagenomics.
Learning Objectives:

- Understand the rationale behind going "metagenomic" in microbial community analysis
- Understand methods of taxonomic classification of microbial communities
- Master using RDP tools to classify unknown 16S rRNA sequences
- Be able to calculate metrics that quantify population diversity
- Explore recent advances in our understanding of the human microbiome

Human Genomics. Linking Genes to Traits: Genome-Wide Association Studies.
Learning Objectives:

- Describe genetic variation in a population
- Understand concept of genetic association
- Understand concept of multiple test correction in genome wide association studies
- Know how to perform allelic trend test
- Learn to navigate the human population genomics data via Ensembl portal
- Learn to link phenotypic and genomic data via OMIM database
- Understand the concept of linkage disequilibrium and how it is used to detect genomic regions under selection

Reading and Video Materials

Most of the reading for this course will come from the "Concepts in Bioinformatics and Genomics" by Nomand and McCurdy, Oxford University Press, 2016. This is not a perfect textbook, as genomics is a quickly changing field of study. Since we cannot have physical copies on reserve in the libraries, the selected chapters/pages will be available as scanned PDF files on *Canvas*. These will be supplemented by chapters from other texts, review articles and online tutorials, also available via *Canvas*. This syllabus, power point presentations, links to pre-lecture videos, links to library reserves and additional readings, and assignments will be posted to Dartmouth's *Canvas* site (http://canvas.dartmouth.edu). To access the site, use your NetID, password and Duo authentication. You will need to use Dartmouth's VPN client to access some of the posted resources.

Class exercises will be available via *Google Docs* on *Dartmouth's Google Drive*. To access the Dartmouth's Google Suite, use your NetID, password and Duo authentication.

Meeting Places

The regular class period will meet on Zoom Please use Dartmouth's Zoom account, as this meeting place is limited to Dartmouth users only. These class periods will be recorded and the videos will be made available via *Canvas* (see **Consent to Recording** section below)

The office hours will take place in a different Zoom room. The office hours will *not* be recorded.

We will also use the dedicated Slack channel for various communications during the class periods as well as outside of the class periods:

- announcements/reminders,
- discussions of muddy points and problem sets,
- communications within teams,
- communications between teams and professor during breakout sessions.

Teaching Approach

Class periods will be a combination of mini-lectures, discussions of the assigned reading and videos, software demonstrations, and computer-based exercises completed in groups. The exercises will give you ample opportunities to master analytic and problem-solving skills using real data.

To make time for interactive in-class exercises and to avoid zoom fatigue, most lecture materials will be provided in the form of pre-recorded videos, which you will need to watch *before* coming to class. You will also be assigned quizzes to reflect on the pre-lecture video and reading materials. These quizzes will not be graded for content, but they will count toward your participation grade. I will read your muddy points reports before coming to class. I will try to address most muddy points during the class periods, but due to time limitations some will be addressed in after-class Slack postings.

On Independent Project

Our research theme this term is "Coronaviruses"! During the course I will introduce you to the genomes of coronaviruses, including the genome of SARS-CoV-2. For your project, you will be assigned one gene from a coronavirus genome. You will analyze it using tools and methods that we will be learning during the term, following the provided instructions. You are expected to work on the project throughout the term, and to keep "analyses notebook" and to report interim results in three "progress quizzes". At the end of the term, you will make a short video presentation of your findings, which will be reviewed and graded by me and your classmates. You will also summarize the project in a research paper. Details will be discussed in class and specific instructions provided via *Canvas* throughout the term.

Additional Assignment for Graduate Students only

You will select a gene related to your own research and will create and perform an analysis of the gene using tools and methods that we will be learning during the term. You will report your findings using a Canvas quiz. Details will be discussed in class.

Grading

Undergraduate Students:

Exam #1 (30%) – October 6-8 (on Gradescope; 2hrs)

Exam #2 (30%) – November 3-5 (on Gradescope; 2hrs)

Independent Project (35%):

- progress quizzes (5%) – due throughout the term;
- presentation video (10%) - due November 17, 11:59PM;
- analyses notebook (5%) - due December 1, 11:59PM;
- final paper (15%) - due December 1, 11:59PM;

In-class participation (5%)

- active engagement during the class periods;
- pre-lecture quizzes (will be posted via *Canvas*)

Graduate students:

Exam #1 (30%) – October 6-8 (on Gradescope; 2hrs)

Exam #2 (30%) – November 3-5 (on Gradescope; 2hrs)

Independent Project (35%):

- progress quizzes (4%) – due throughout the term;
- presentation video (9%) - due November 17, 11:59PM;
- analyses notebook (4%) - due December 1, 11:59PM;
- final paper (13%) - due December 1, 11:59PM;
- “your own gene” analysis (5%) – due October 30, 11:59PM;

In-class participation (5%)

- active engagement during the class periods;
- pre-lecture quizzes (will be posted via *Canvas*)

Expectations

Here is what I expect from you:

(1) to carefully read and watch the assigned material before class,

(2) to enthusiastically participate in class discussions and group work,

(3) to help your group mates with in-class hands-on exercises.

Meeting live online poses some unique participation challenges. There are a few participation guidelines I would like you to follow to help things go as smoothly as possible and minimize disruptions to our class sessions.

- **Please keep your microphone muted unless you are talking.** If you hold down your spacebar to unmute while you are talking, the microphone will automatically mute itself again when you let go. This will limit intrusions of local environmental noise and audio feedback.
- **Please use headphones during class whenever possible.** Otherwise, voices and noises will be picked up by your microphone and echo audibly in our call whenever your microphone is unmuted. In some cases, we may even experience loud feedback. A quiet meeting location is preferred, if possible.
- **If you have the necessary equipment and internet access to do so, please turn on your video during class.** This will make it easier to track who is talking and help us maintain a sense of community at a distance.
- It will likely be more challenging than usual to maintain attention during class. To some extent this may be unavoidable given the circumstances. Whenever possible, **please try to limit multi-tasking**, such as scrolling on your phone, navigating out of the Zoom call and into other programs or windows, etc. unless its directly relevant to what we're doing in class.
- I encourage you to **post your questions**, or to type "I have a question", to the live chat within Zoom during class as they arise. I will make sure to make breaks to go over the questions.

You can expect me to:

(1) Bring expertise into the classroom.
(2) Stimulate interest in the course material.
(3) Provide consultations during the hands-on activities and be available to answer questions on lectures and hands-on exercises.
(4) Return graded assignments promptly.

I recognize that due to ongoing circumstances you may find yourself in a situation that makes it challenging to keep up with the course material or participate in any fashion. If such a situation arises, please reach out to me as soon as possible so we can find a solution. If you are already aware of circumstances that will affect your ability to participate in the course regularly or occasionally, please arrange a meeting with me at the start of the term so we can plan ahead.

Late Assignments

Please speak with me at the beginning of the term if you anticipate circumstances that might affect your ability to get your work in on time and reach out if such a situation arises along the way. Reasonable allowances will be made for students facing time zone differences, work conflicts, and other extenuating circumstances, as long as we connect about these matters prior to the assignment deadlines.

Remote Learning Tips
https://students.dartmouth.edu/undergraduate-deans/students/academic-advising/remote-learning-tips

Academic Honor

The Dartmouth Honor Principle applies to all work you submit for a grade in this course. During the in-class activities, however, you are *encouraged* to collaborate with others while designing analyses, running analyses, and drawing conclusions. Since you all will be using the same software for your exercises and projects, I also encourage you to consult with your classmates if you run into technical difficulties with running programs or accessing databases.

(The detailed description of the Dartmouth Honor Principle is available at https://students.dartmouth.edu/judicial-affairs/policy/academic-honor-principle)

Your Needs and Wellness

If you have questions about whether you need to receive academic adjustments and services, please contact the Student Accessibility Services (SAS) office (Carson Hall, Suite 125, 646-9900; all inquiries and discussions are confidential) as early in the term as possible. Once SAS has authorized services, please contact me privately.

If you have a religious observance that conflicts with your participation in the course, please meet with me before the end of the second week of the term to discuss appropriate accommodations.

The academic environment at Dartmouth is challenging, our terms are intensive, and classes are not the only demanding part of your life. There are a number of resources available to you on campus to support your wellness, including your undergraduate dean (http://www.dartmouth.edu/~upperde/), Counseling and Human Development (http://www.dartmouth.edu/~chd/), and the Student Wellness Center (http://www.dartmouth.edu/~healthed/). I want you to be aware of these resources and encourage you to use them as needed

Title IX

At Dartmouth, we value integrity, responsibility, and respect for the rights and interests of others, all central to our Principles of Community. We are dedicated to establishing and maintaining a safe and inclusive campus where all have equal access to the educational and employment opportunities Dartmouth offers. We strive to promote an environment of sexual respect, safety, and well-being. In its policies and standards, Dartmouth demonstrates unequivocally that sexual assault, gender-based harassment, domestic violence, dating violence, and stalking are not tolerated in our community.

The Sexual Respect Website (https://sexual-respect.dartmouth.edu) at Dartmouth provides a wealth of information on your rights with regard to sexual respect and resources that are available to all in our community. Please note that, as faculty members, we are obligated to share disclosures regarding conduct under Title IX with Dartmouth's Title IX Coordinator. Confidential resources are also available, and include licensed medical or counseling professionals (e.g., a licensed psychologist), staff members of organizations recognized as rape crisis centers under state law (such as WISE), and ordained clergy

(see https://sexual-respect.dartmouth.edu/reporting-support/all-resources/confidential-resources). Should you have any questions, please feel free to contact Dartmouth's Title IX Coordinator (Kristi.Clemens@Dartmouth.edu) (and deputies if appropriate).

Consent to Recording

(1) Consent to recording of course

a) By enrolling in this course, you affirm your understanding that this course and associated **group** meetings involving students and the instructor, may be recorded within any digital platform used to offer remote instruction for this course;
b) You further affirm that the instructor owns the copyright to their instructional materials, of which these recordings constitute a part, and distribution of any of these recordings in whole or in part without prior written consent of the instructor may be subject to discipline by Dartmouth up to and including expulsion;
c) You authorize Dartmouth and anyone acting on behalf of Dartmouth to record your participation and appearance in any medium, and to use your name, likeness, and voice in connection with such recording; and
d) You authorize Dartmouth and anyone acting on behalf of Dartmouth to use, reproduce, or distribute such recording without restrictions or limitation for any educational purpose deemed appropriate by Dartmouth and anyone acting on behalf of Dartmouth.

(2) Requirement of consent to one-on-one recordings

By enrolling in this course, you affirm that you will not under any circumstance make a recording in any medium of any one-on-one meeting with the instructor without obtaining the prior written consent of all those participating, and you understand that if you violate this prohibition, you will be subject to discipline by Dartmouth up to and including expulsion, as well as any other civil or criminal penalties under applicable law.