Microbial 'Omics (BIOS 25420)

for
University of Chicago Biological Sciences Division
Upper Level Elective for Biology Majors

Description

Every ecological niche our planet has to offer, including the human body itself, is home to an astonishing number of microbial cells that form complex communities. The last several years witnessed tremendous advances in molecular and computational approaches which now offer unprecedented access to these communities through new 'omics strategies. Developing an overall understanding of these strategies -including the ability to identify their appropriate applications and shortcomings- has quietly become a de facto necessity in the journey of an independent life scientist. The primary aim of this course is to empower its participants and enable them to study the ecology, evolution, and functioning of naturally occurring microbial populations. Through equal proportions of theory and practice, the participants will use state-of-the-art computational methods to work with real-world microbial data and recognize the current conceptual framework that helps us wrap our collective mind around the most diverse form of life on our planet.

Assessment

- *Mini quizzes* (25%). Participants will be referred to articles, reviews, videos, or news articles that are relevant to the lecture topic at the end of every week, and occasionally asked to share their take from these material through mini quizzes at the last 10 minutes of the following Thursday slot.
- *Class citizenship* (25%). At the end of each lecture participants will send a short e-mail with their own summary of the main concepts discussed during the class and their understanding in their own words. They will also include short questions about concepts that were discussed during the lecture yet remained unclear.
- *Mini data analysis challenges* (25%). Participants will be given four to five data analysis challenges to put their new skills into use. They will be expected to demonstrate their ability to (1) interpret the significance and relevance of these questions, (2) utilize tools and strategies they learned during the course, (3) investigate best practices beyond the course material, and (4) evaluate their findings in the context of the existing literature. Participants will report their computational analyses and findings in a reproducible manner.
- *Take home final exam* (25%). This final written exam will cover most key concepts discussed throughout the class.

Previous feedback

We have learned from the previous year's feedback that the most challenging aspects of this course include (1) the need to work with the command line environment (for a crash course I would suggest this resource) and (2) ability to secure enough storage space on personal computers to complete data analysis tasks (our largest analysis will require about 75Gb free space on your personal computer or external disk). While the course was taken by individuals with diverse backgrounds (including mathematics and computer science), the extent of prior exposure to biological concepts did not seem to have a dramatic influence on learning (so, if you find microbial 'omics interesting, you are very likely going to do well in this course even if you have a very modest training in biology).

Syllabus Plan[1]

Week #01

In Theory (Tue)

Course Logistics & A brief introduction to microbial life on Earth

In Practice (Thu)

Installing anvi'o and a brief introduction to the platform

Learning Objectives

- Describe the extent of microbial diversity on Earth, their involvement in key biogeochemical processes, as well as human health and disease.
- Remember seminal studies that contributed to our understanding of the diversity, functioning, and metabolic potential of naturally occurring microbial communities and approaches to study them.
- Explain the old and new intellectual and technical challenges that prevent us from defining fundamental units of microbial life, and the *art* of moving forward without any answers.

Week #02

In Theory (Tue)

An overview of data-driven strategies to survey environmental microbiomes

In Practice (Thu)

A tutorial on metagenomic read recruitment and profiling

Learning Objectives

- Recognize currently available 'omics data types (such as *metagenomics*, *metatranscriptomics*, *metaproteomics*, *metametabolomics*), approaches (such as *pangenomics*, *phylogenomics*), and questions they *can* and *can't* answer.
- Identify the state-of-the-art practices to analyze each data type.
- Recognize the available computational solutions to gain insights into fundamental questions in microbiology.
- Explain the power of metagenomic read recruitment strategies and interpret ecological and evolutionary insights we can infer through this strategy.

[1] Reading through this syllabus plan let's keep in mind what Winston Churchill had once said about plans: "*plans are of little importance, but planning is essential*". While the syllabus is taking care of what is essentials by clarifying the intention of the course, the course director will happily consider deviations from the planned lectures and activities if there is consensus among the course participants on switching to an unlisted topic of broad interest or activity around microbial 'omics.

Week #03

In Theory (Tue)

Genome-resolved metagenomics: opportunities, pitfalls, and frustrations

In Practice (Thu)

Hands-on exercise to recover genomes from a real-world metagenome

Learning Objectives

- Recognize the difference between microbial isolates, enrichments, single-cell amplified genomes, and metagenome-assembled genomes.
- Explain the importance of the ability to acquire genomic information from microbes we have not yet cultivated.
- Tell the basics of algorithms and strategies to reconstruct microbial genomes directly from metagenomes.
- Appreciate the limitations and opportunities associated with genome-resolved workflows.
- Employ genome-resolved strategies to recover genomes from not-too-complex metagenomes and scrutinize their quality.

Week #04

In Theory (Tue)

Pangenomics: comparative genomics in the era of genomic explosion

In Practice (Thu)

Application of pangenomics to a marine microbial genus

Learning Objectives

- Explain the concepts of core and accessory genome, as well as open and closed pangenomes.
- Define gene clusters in pangenomes through sequence homology.
- Interpret ecological and evolutionary insights pangenomes offer.
- Apply functional enrichment analyses in pangenomes to investigate determinants of fitness.
- Demonstrate the application of pangenomics to a human gut commensal.

Week #05

In Theory (Tue)

Metapangenomics: A nexus between pangenomes and metagenomes

In Practice (Thu)

Turning the pangenome from the previous week into a metapangenome

Learning Objectives

- Explain the emerging opportunities to investigate the functioning and the ecology of microbial populations by linking pangenomes and metagenomes.
- Comprehend the power of characterizing a single genome across metagenomes.
- Use metapangenomics to put a pangenome into the context of publicly available metagenomes.

Week #06

In Theory (Tue)

Phylogenomics: inferring evolutionary relationships between microorganisms

In Practice (Thu)

Placing a metagenome-assembled genome of unknown origin into the Tree of Life

Learning Objectives

- Identify commonly used genes, statistics, and heuristics to infer phylogenomic relationships across distantly related organisms.
- Recognize historical events that led to the emergence of the current Tree of Life, and why scientists cant even.
- Appreciate technical and theoretical limitations of inferring deep branching patterns confidently.
- Apply reasonable phylogenomic analyses to discuss potential origins and taxonomic affiliations of novel genomes.

Week #07

In Theory (Tue)

Microbial population genetics: tools, terminology, and open questions

In Practice (Thu)

Inferring the ecology of environmental subpopulations through single-nucleotide variants

Learning Objectives

- Learn ecological and evolutionary implications of clonality and heterogeneity within environmental populations.
- Identify approaches to study single-nucleotide variants, and methods to reconstruct haplotypes.
- Comprehend differences and overlaps between population genetics approaches in animal populations and microbial populations
- Characterize variation within a metagenomic sample and make use of it for exploratory analyses or hypothesis testing.

Week #08

In Theory (Tue)

Metaepitranscriptomics: translational regulation and its ecological implications

In Practice (Thu)

High-throughput sequencing of Transfer RNA transcripts

Learning Objectives

- Identify limitations of DNA and RNA molecules to study microbial life.
- Recognize another aspect of life to regulate its fitness in shortest scales of time.

Week #09 (with Aly A. Khan)

In Theory (Tue)

Applications of machine learning to microbiome data

In Practice (Thu)

Paper discussions

Learning Objectives

- Recognize different machine learning classification algorithms.
- Explain how to formulate a discriminative machine-learning task with training, validation, and test data.
- Interpret various machine learning models.
- Describe some applications of machine learning methods in microbiome data, such as predicting tissue of origin or disease status.

Week #10

No class due to the late start of the semester.