

Advanced Bioinformatics Laboratory
BIMM 185, Spring 2017
MW 5:00-7:50
York 1310

Instructor:

Dmitry Rodionov

Research Assistant Professor,
Department of Bioinformatics
Sanford Burnham Prebys Medical Discovery Institute
Email: drodionov@ucsd.edu

Office Hours (optional): TBD

Office Address: Humanities and Social Sciences Bldg. 1145LB

Office Phone: 858-246-0965

Instructor Assistant (IA)

Simon Leyn

Research Scientist
Department of Bioinformatics
Sanford Burnham Prebys Medical Discovery Institute
Email: sleyn@ucsd.edu

Office Hours (optional): TBD

Private appointments to discuss accommodations or special circumstances can be arranged by email.

This course emphasizes the hands-on application of bioinformatics methods to biological problems. Students will gain experience in the application of existing publicly available software, as well as in combining approaches to answer specific biological questions. The emphasis of this course will be given to the advanced bioinformatics tools for genome analysis, functional genomics, and biological networks in application to microbial genomes. This course is open to bioinformatics majors only.

Learning Objectives

By the end of the course, students will:

- Be able to explain why bioinformatics approaches are important to understand and interpret results in many different biological areas of study.
- Be able to utilize existing bioinformatics tools to analyze various kinds of biological data and solve complex scientific questions.
- Develop critical analysis and research skills that can be applied to understand and use new bioinformatics tools that are developed in the future.
- Be able to access publicly available biological data sets and bioinformatics tools.

Grading Criteria

Grading will be based on weekly lab reports (60%), in-class quizzes and discussion participation (10%) and independent project report (30%). For the first 7 weeks, students will work on a weekly lab project in class on Monday and Wednesday. Reports for weekly problem assignments are due the following Monday at the start of class, and will be returned within one week. For the last 3 weeks of the quarter, students will work on an independent project.

Attendance is mandatory, but in the case of valid excuses (must be approved by professor before scheduled class), quizzes and participation points can be made up.

Team work vs. individual work

In the classroom, while you are working on coding and analysis for weekly project, you are encouraged to chat with your neighbors, form teams, look things up online, and ask the instructor and IA for help. However, when writing your lab reports, work independently, and make sure that the lab report is in your own words and reflects your own understanding.

You may use scientific literature, but must cite it in the text. DO NOT directly copy/paste code to/from other students in the class or websites. DO NOT post questions about assignments on sites like seqanswers or stack overflow (you can research existing threads on these websites, just don't start new ones for the coursework).

Academic Honesty

There is a zero tolerance policy regarding academic dishonesty. See the Academic Integrity Agreement (which you must sign to receive a passing grade in the class) for more details on the topic. Violators of these policies may be subject to UCSD rules for academic integrity.

Course schedule (planned)

Week 1 (Apr 3/5): “Genes and proteins in biological databases”

Introduction into genes and proteins, translation/genetic code, molecular sequence databases (Genbank, Uniprot), protein families and functional classifications (Pfam, CDD, TCDB, P2TF, CAZy), 3D structures (PDB), homology, pairwise and multiple sequence alignment, BLAST and protein motif detection tools. Gene finding and annotation by ORFfinder. UGene toolkit.

Week 2 (Apr 10/12): “Microbial genomes and genomic databases”

Overview of genome organization, features in intergenic regions, operons. Genomic projects and sequence repositories (GOLD, Genbank). Unfinished/complete genomes. Genome file formats. Genome browsers and integrated analysis platforms (IMG/MER, PubSEED, PATRIC, MicrobesOnline).

Week 3 (Apr 17/19): “Comparative genomics and genome evolution”

Comparative genomics approaches for gene annotation, using genome context, gene fusions, gene co-occurrence. Concepts of orthologs and paralogs, horizontal gene transfer, gene duplications, genome evolution. Proteome comparisons, massive computing of orthologs and pangenomes (services in IMG/MER and KBase).

Week 4 (Apr 24/26): “Metabolic pathways and models”

Conception of biochemical pathways, enzyme classification, reactions and metabolites. Metabolic pathways databases (KEGG, BioCyc). Metabolic reconstructions and missing genes. Subsystems approach for metabolic reconstruction (PubSEED). Whole-genome metabolic models, flux balance analysis (ModelSEED, KBase).

Week 5 (May 1/3): “Transcriptional regulatory networks and regulons”

Mechanisms for regulation of transcription in bacteria. Transcription factors, riboswitches. Global and local regulons, modulons and stimulons. DNA motif detection (MEME). Comparative genomics approaches (phylogenetic footprinting, consistency check). Tools for regulons reconstruction (RegPredict, Riboswitch scanners). Databases of transcriptional regulons (RegulonDB, RegPrecise). Transcriptional networks.

Week 6 (May 8/10): “Microbial genome assembly and annotation”

Approaches for whole genome sequencing and assembly. Genome assembly and annotation services (PATRIC, RAST, KBase). QC Analysis and correction of short reads. Next Generation Sequencing (NGS) data analysis. Detecting and annotating genetic variations.

Week 7 (May 15/17): “Molecular phylogenetics and trees”

Algorithms for phylogenetic tree construction (ML, NJ). Rooted/unrooted trees, bootstrapping. Construction of species trees (16S or ribosomal proteins). Phylogenetic analysis of large protein families with multiple duplications. visualization of properties on trees (iTOL, Dendroscope etc.). Evolution of transcriptional networks and regulons.

Weeks 8, 9 and 10 (May 22 - June 7): Independent projects.

May 31: A mini-course on metagenomics by invited lector (Alex Tyakht)