SIOB 278 | Spring 2017
Ocean 'Omics: Metagenome Data Processing & Analysis

Class tutorials: Wednesdays 10:00 am – 12:00 pm, 4135 Hubbs Hall
Independent work with SIO-BUG folks: Wednesdays 2:00 pm – 4:00 pm, 2230 Hubbs Hall (optional)

DATE	TOPICS / ACTIVITIES
Apr 05	**Introduction to metagenomics, research questions, and experimental design** *AWS access; Command line operations; Local software needs; Selection of independent data sets*
Apr 12	**Data types, expectations, and quality control** *Sequencing technologies; File formats; Data quality control processing (FastQC, Trimmomatic)*
Apr 19	**Assembly of metagenome sequence data** *Assembly processing (IDBA-UD, MetaSPAdes); Assembly statistics & validation (Quast)*
Apr 26	**Check point 1: Research questions, discussion, and papers**
May 03	**Gene prediction and functional annotation of assembled contigs/scaffolds** *Annotation pipeline (Prokka); Web-based analysis portal (JGI-IMG); Annotation viewing (Artemis)*
May 10	**Binning of assembled sequences I: DNA compositional metrics** *Nucleotide frequency analysis (%GC); Coverage (mapping); Analysis & visualization (Anvi'o)*
May 17	**Binning of assembled sequences II: taxonomy assignments** *Homology searches (BLAST, DIAMOND, DarkHorse); Visualization of binned data (Anvi'o)*
May 24	**Check point 2: Research questions, discussion, & papers**
May 31	**Ad-hoc analyses: population genome assembly, gene & pathway discovery**
Jun 07	**Presentation of independent research projects**

Instructors: Eric Allen   Jessica Blanton
Email: eallen@ucsd.edu   jmblanton@ucsd.edu
Office: 4170 Hubbs Hall   4130 Hubbs Hall

Course Web Site: http://tritoned.ucsd.edu (tutorials, lecture notes, readings, etc.)
Course Requirements: Laptop with SSH client (Mac Unix shell or Windows MobaXterm); Familiarity with Unix command line interface – a nice introductory tutorial can be found here; AWS account ("Free Tier")

Course Prospectus: The field of metagenomics is a rapidly evolving discipline driven by advances in DNA sequencing technologies and computational biology. Metagenomic approaches have been applied to diverse marine microbial habitats ranging from seawater and sediment samples to the microbiomes associated with marine animals. Ultimately, these methods are applied to reveal the taxonomic and metabolic diversity of environmental microorganisms. This seminar course will provide a hands-on workshop environment to introduce methods and concepts involved in the *de novo* processing of metagenomic sequence data. At the end of this course, you should be able to:

- Demonstrate expertise in basic Unix command line operations & AWS cloud computing resources
- Access publically available biological data sets and bioinformatics tools
- Evaluate software tools and methods for metagenome data processing
- Execute (and troubleshoot) a variety of bioinformatic packages and analysis pipelines
- Design experiments for metagenome discovery and analysis