

BIOINF 524/525 Winter 2017
Foundations of Bioinformatics and Systems Biology
<http://tinyurl.com/bioinf525-w17>

Course Director: Dr. Barry Grant (DCM&B, bjgrant@med.umich.edu)

Description: BIOINF524 is a three credit three module course covering (1) *Foundations of Bioinformatics*, (2) *Statistics in Bioinformatics*, and (3) *Systems Biology*. BIOINF525 allows for each one credit module to be registered for separately. BIOINF 524/525 is intended to provide non-bioinformaticians in any field of biomedical sciences with new quantitative tools in bioinformatics, statistics, and network modeling that will be directly applicable to their ongoing research. Throughout the course, students will also build proficiency with the R programming language, which will be used as a foundation in most lab sessions.

Module 1:	January 10 – February 2	(four lectures and four labs).
Module 2:	February 7 – March 16	(five lectures and five labs).
Module 3:	March 21 – April 13	(four lectures and four labs).

Schedule:

Lectures:	Tuesdays 2:30 - 4:00 PM, Rm. 2062 Palmer Commons Bldg.
Labs:	Thursdays 2:30 - 4:00 PM, Rm. 2062 Palmer Commons Bldg.

Prerequisites: A familiarity with basic biomedical concepts and basic knowledge of computer usage. No programming skills are needed. You **MUST** bring your own wifi enabled laptop to lectures and labs to fully participate in this class.

Grading: Satisfactory/unsatisfactory grading will be based on a combination of lecture and lab involvement together with weekly homework and quiz assignment performance.

Module 1 (1 credit): **Foundations of Bioinformatics**
Module 1 Director: Dr. Barry Grant (DCM&B, bjgrant@med.umich.edu)

Bioinformatics is driving the collection, analysis and interpretation of big data for biomedicine and the biosciences. This module provides an introduction to practical issues of computer-based handling and interpretation of biomolecular and genomic datasets. We specifically target bioinformatics software and data resources freely available on the Internet.

Lecture (1-1): *Introduction to bioinformatics*

Instructor: Dr. Barry Grant

Time: Jan 10 (Tuesday), 2:30 - 4:00 PM

Topics:

What are bioinformatics and computational biology? Major NCBI and EBI bioinformatics resources. How do we actually do bioinformatics? Major areas of research and application.

Lab (1-1): *Bioinformatics databases and major online resources*

Instructor: Dr. Barry Grant

Time: Jan 12 (Thursday), 2:30 – 4:00 PM

Topics:

The purpose of this lab session is to introduce a range of core bioinformatics databases and associated online services whilst investigating the molecular basis of several common human disease.

BIOINF 524/525 Winter 2017
Foundations of Bioinformatics and Systems Biology
<http://tinyurl.com/bioinf525-w17>

Lecture (1-2): Sequence alignment and database searching

Instructor: Dr. Barry Grant

Time: Jan 17 (Tuesday), 2:30 - 4:00 PM

Topics:

Homology and sequence alignments, Dynamic programming, Sequence similarity and database searching, Interpreting search results – E value, Needleman and Wunsch, Smith and Waterman, The BLAST algorithm.

Lab (1-2): Sequence alignment algorithms and applications

Instructor: Dr. Barry Grant

Time: Jan 19 (Thursday), 2:30 – 4:00 PM,

Topics:

Aligning novel sequences with previously characterized genes or proteins provides important insights into their common attributes and evolutionary origins. In this lab session we will explore the principles underlying the computational tools that can be used to compute and evaluate sequence alignments.

Lecture (1-3): Structural bioinformatics

Instructor: Dr. Barry Grant

Time: Jan 24 (Tuesday), 2:30 - 4:00 PM

Topics:

Major goals, current research challenges, and application areas of structural bioinformatics. Key concepts covered include: Sequence-structure-function relationships; Energy landscapes; Physics and knowledge based modeling approaches for describing the structure, energetics and dynamics of biomolecules computationally.

Lab (1-3): Protein structure visualization, analysis and small molecule docking

Instructor: Dr. Barry Grant

Time: Jan 26 (Thursday), 2:30 - 4:00 PM

Topics:

The RCSB PDB resource. Interactive biomolecular structure visualization with VMD. Structural analysis of protein families and prediction of protein flexibility with Bio3D. *In silico* docking and virtual screening strategies for drug design.

Lecture (1-4): Genome informatics

Instructor: Dr. Ryan Mills

Time: Jan 31 (Tuesday), 2:30 - 4:00 PM

Topics:

High throughput sequencing technologies, biological applications and bioinformatics analysis methods. Web interfaces to genome databases at Ensembl and UCSC.

Lab (1-4): Web resources for analyzing genomic data

Instructor: Drs. Ryan Mills and Barry Grant

Time: Feb 2 (Thursday), 2:30 - 4:00 PM

Topics:

The Galaxy web-based platform for genomic data analysis. Manipulation and analysis of next generation sequencing data sets.

BIOINF 524/525 Winter 2017
Foundations of Bioinformatics and Systems Biology
<http://tinyurl.com/bioinf525-w17>

Module 2 (1 credit): **Introduction to Statistics in Bioinformatics**
Module 2 Director: Dr. Niko Kaciroti (Biostatistics, nicola@umich.edu)

Basic statistics as used in bioinformatics, especially standard statistical tests of significance and when they apply. Applications to genetics, experimental and observational medical data, as well as exploration of multiple testing issues that arise in bioinformatics and other experimental settings. NOTE: Dr. Kaciroti is the instructor for all Module 2 lectures and labs.

Lecture (2-1): *Framework for statistical analysis of biomedical data*

Time: Feb 7 (Tuesday), 2:30 - 4:00 PM

Topics:

Probability distributions, quantifying central values and variability, quantifying association, graphical displays of data

Lab (2-1): *Using R for descriptive statistics and summarizing data*

Time: Feb 9 (Thursday), 2:30 - 4:00 PM

Topics:

Introduction to R and RStudio. Using R we will explore probability distributions, quantifying central values, variability, association, and generating informative graphical displays of data

Lecture (2-2): *Approaches to statistical estimation and testing*

Time: Feb 14 (Tuesday), 2:30 - 4:00 PM

Topics:

Estimation and standard errors, standard errors for means, correlations, and log odds ratios, formal hypothesis testing, tests involving means, correlations, and log odds ratios, power.

Lab (2-2): *Statistical estimation and hypothesis testing*

Time: Feb 16 (Thursday), 2:30 - 4:00 PM

Topics:

Estimation and standard errors, standard errors for means, correlations, and log odds ratios, formal hypothesis testing, one and two sample tests involving means, power.

Lecture (2-3): *Analyses involving associations*

Time: Feb 21 (Tuesday), 2:30 - 4:00 PM

Topics:

Pearson correlation, t-test, odds ratios, discussion of a research article

Lab (2-3): *Pearson correlation, t-test, and log odds ratios*

Time: Feb 23 (Thursday), 2:30 - 4:00 PM

Topics:

Tests based on Pearson correlation t-test, and log odds ratios

Lecture (2-4): *Linear regression*

Time: Mar 7 (Tuesday), 2:30 - 4:00 PM

Topics:

Single and multiple variable linear regression, Bonferroni correction, power for regression analysis

BIOINF 524/525 Winter 2017
Foundations of Bioinformatics and Systems Biology
<http://tinyurl.com/bioinf525-w17>

Lab (2-4): Regression models

Time: Mar 9 (Thursday), 2:30 - 4:00 PM

Topics:

Fitting regression models for prediction and effect estimation, inference for regression effects, R^2 , diagnostics, comparing models

Lecture (2-5): Introduction to graphical methods for multivariate data analysis

Time: Mar 14 (Tuesday), 2:30 - 4:00 PM

Topics:

Clustering methods, Multidimensional scaling and Principal component analysis

Lab (2-5): Clustering and principal component analysis

Time: Mar 16 (Thursday), 2:30 - 4:00 PM

Topics:

Multivariate data, Heat maps and dendrograms, clustering methods, principal component analysis.

.....
Module 3 (1 credit): Biological networks and systems-level modeling

Module 3 Director: Dr. Peter Freddolino (Biological Chemistry, petefred@umich.edu)

Computational analysis of biological networks, OMICs data (genomics, transcriptomics, metabolomics, proteomics). Application of advanced analysis and modeling approaches to study pathways and networks. Emphasis on using existing high throughput data sets alongside newly generated data to analyze and interpret research findings.

Lecture (3-1): Introduction to systems biology

Instructor: Dr. Peter Freddolino

Time: Mar 21 (Tuesday), 2:30 - 4:00 PM

Lab (3-1): Network analysis for systems biology

Instructor: Dr. Peter Freddolino

Time: Mar 23 (Thursday), 2:30 - 4:00 PM

Topics:

Representation of data as graphs. Pathway and network exploration and visualization with Cytoscape (including the MetScape app and the RCytoscape package) and analysis with igraph.

Lecture (3-2): High throughput sequencing methods in systems biology

Instructor: Dr. Peter Freddolino

Time: Mar 28 (Tuesday), 2:30 - 4:00 PM

Topics:

Overview of high throughput sequencing-based methods used to investigate biological networks, along with an introduction to databases and analysis considerations for each.

BIOINF 524/525 Winter 2017
Foundations of Bioinformatics and Systems Biology
<http://tinyurl.com/bioinf525-w17>

Lab (3-2): *Mapping genetic regulatory networks using high-throughput sequencing*
Instructors: Dr. Peter Freddolino
Time: Mar 30 (Thursday), 2:30 - 4:00 PM

Topics:

Finding and interpreting RNA-seq and ChIP-seq data sets to study regulatory networks, including identification of differentially expressed genes, location of transcription factor binding sites, and inference of regulatory motifs.

Lecture (3-3): *Network modeling for hypothesis testing and generation*

Instructor: Dr. Peter Freddolino
Time: Apr 4 (Tuesday), 2:30 - 4:00 PM

Topics:

Conceptual understanding of kinetic models and constraint-based models of metabolic networks; approaches for translating biochemical schematics into quantitative frameworks.

Lab (3-3): *Modeling and inference of metabolic networks*

Instructors: Dr. Peter Freddolino
Time: Apr 6 (Thursday), 2:30 - 4:00 PM

Topics:

Students will learn how to represent simple metabolic pathways in SBML, and use the *sybil* R package to simulate the effects of mutations in a model network.

Lecture (3-4): *Machine learning approaches in systems biology*

Instructor: Dr. Peter Freddolino
Time: Apr 11 (Tuesday), 2:30 - 4:00 PM

Topics:

Introduction to machine learning and overview of key algorithms (linear classifier, SVM, decision trees and random forests). Existing and potential areas for application of machine learning in the analysis of biological networks.

Lab (3-4): *Application of machine learning to biological network analysis*

Instructor: Dr. Peter Freddolino
Time: Apr 13 (Thursday), 2:30 - 4:00 PM

Topics:

Overview of R modules for machine learning; comparison of a variety of approaches for predicting regulatory modules, protein function, and protein-protein interactions.
