

Applied Bioinformatics 2018

Draft Schedule

MDI Biological Laboratory
Salisbury Cove, ME 04672

July 7 – July 12, 2018

A hands-on course for advanced graduate students, post-doctoral trainees, and researchers at all levels interested in incorporating bioinformatics into their research.

All sessions and seminars will be held on the second floor of the Maine Center for Biomedical Innovation (MCBI) on the campus of the MDI Biological Laboratory.

Saturday, July 7, 2018. Course Introduction and Overview

5:00 pm – 6:00 pm – Registration and housing check in (Dining Hall)

6:00 pm – 7:00 pm Dinner (Dining Hall-)

7:00 pm – 9:00 pm – Course Introduction and Overview (Ben King, University of Maine, MCBI)

- Boundaries with biology, statistics, computer science
 - Contemporary biological examples
 - Cell Biology
 - Evolution
 - Biomedical
 - Statistical Challenges and Solutions
 - Raw Computational Challenges and Solutions
 - Problems of Data Representation and Solutions
-

Sunday, July 8, 2017. Sequence Analysis

7:00 am – 9:00 am Continental Breakfast (Dining Hall)

9:00 am – 10:30 am – Introduction to High-Throughput Sequencing (Kelley Thomas, University of New Hampshire, MCBI)

- Technologies and Applications
- History
- Chemistry
- Instruments

- Costs
- High-level analysis workflow

10:30 am – 10:45 am – Break

10:45 am – 12:00 pm – Overview of Sequence Analysis Workflow (Ben King, University of Maine, MCBI)

- Workflow outline
 - Read diagnostics
 - Trimming
 - Read alignment
 - Visualization of alignments
- Common file formats
- Unaligned reads (FASTQ, etc.)
- Aligned reads (SAM/BAM)

12:00 pm – 1:00 pm – Lunch (Dining Hall)

1:00 pm – 2:30 pm – Read Alignment and Analysis Workshop (Part 1) (Ben King, University of Maine, MCBI)

- Importing reads
- Trimming and QC
- Read mapping to reference sequence(s)

2:30 pm – 3:00 pm – Break

3:00 pm – 6:00 pm – Read Alignment and Analysis Workshop (Part 2) (Ben King, University of Maine, MCBI)

- Quantifying gene expression
- Variant detection
- Visualization of aligned reads

6:00 pm – 7:00 pm – Dinner (Dining Hall)

7:00 pm – 8:00 pm – Evening Lecture (Steve Munger, The Jackson Laboratory, MCBI)

Monday, July 9, 2018. Gene, Protein and Sequence Tools

7:00 am – 9:00 am Continental Breakfast (Dining Hall)

9:00 am – 10:30 am – Gene, Protein and Sequence Resources (Ben King, University of Maine, MCBI)

- NCBI Entrez system
- UniProt
- Gene Ontology
- miRNA data bases
- RNA-Seq data repositories
 - NCBI Gene Expression Omnibus and EBI Array Express

- NCBI Short Read Archive EBI European Nucleotide Archive

10:30 am – 10:45 am – Break

10:45 am – 12:00 pm – Genome Browsers & Data Retrieval (Ben King, University of Maine, MCBI)

- UCSC Genome Browser
- UCSC Table Browser
- Ensembl
- Biomart

12:00 pm – 1:00 pm – Lunch (Dining Hall)

1:00 pm - 2:00 pm – Analysis of High Throughput Data (Tom Hampton, The Geisel School of Medicine at Dartmouth, MCBI)

- Exploratory Analysis
- Normalization
- Inference

2:00 pm – 2:45 pm – R Power Tools: Way Beyond Word & Excel (Tom Hampton, The Geisel School of Medicine at Dartmouth, MCBI)

- Why R
- Packages: CRAN, Bioconductor
- Reproducible, "literate" statistics

2:45 pm - 3:15 pm – Break

3:15 pm - 4:00 pm – Introduction to R Studio (Britton Goodale, The Geisel School of Medicine at Dartmouth, MCBI)

4:00 pm - 5:00 pm – R Statistical Computing Environment I (Tom Hampton, The Geisel School of Medicine at Dartmouth, MCBI)

- Basic math, stats and plots

6:00 pm – 7:00 pm – Dinner (Dining Hall)

7:00 pm – 8:00 pm – Research Seminar (Kelley Thomas, University of New Hampshire, MCBI)

Tuesday, July 10, 2018. R Statistical Computing Environment

7:00 am – 9:00 am Continental Breakfast (Dining Hall)

9:00 am – 10:00 am – R Statistical Computing Environment II (Tom Hampton, The Geisel School of Medicine at Dartmouth, MCBI)

- Variables and Functions
- Simulation

10:00 am – 10:45 am – Advanced R and Exploratory Data Analysis I (Tom Hampton, The Geisel School of Medicine at Dartmouth, MCBI)

- Introduction to the dataset

10:45 am – 11:00 am – Break

11:00 am – 12:00 pm – Advanced R and Exploratory Data Analysis II (Tom Hampton, The Geisel School of Medicine at Dartmouth, MCBI)

- PCA
- Clustering
- CART models

12:00 pm – 1:00 pm – Lunch (Dining Hall)

1:00 pm – 2:45 pm – EdgeR and Differential Expression (Katja Koeppen, The Geisel School of Medicine at Dartmouth, MCBI)

- Specify Design
- Normalization
- Estimating Common Dispersion
- Identify Differentially Expressed Genes

2:45 pm – 3:15 pm – Break

3:15 pm – 5:00 pm – Gene Set Enrichment (Britton Goodale, The Geisel School of Medicine at Dartmouth, MCBI)

- Concepts: Hypergeometric distribution
- Gene Ontology and KEGG Pathway annotation

6:00 pm – 7:00 pm – Dinner and Discussion on Scientific Careers with Diana Lamppu (Homology Medicines, Inc.) (Dining Hall)

7:00 pm – 8:00 pm – Large Scale Computing (Chris Dagdigian, BioTeam, MCBI)

Wednesday, July 11, 2018. Hands-on Bioinformatics & Ingenuity Pathway Analysis

7:00 am – 9:00 am Continental Breakfast (Dining Hall)

9:00 am – 10:45 am – Small Group Exercise Exploratory Analysis in R (Katja Koeppen, Tom Hampton, Britton Goodale, The Geisel School of Medicine at Dartmouth, MCBI)

- Practice exploratory data analysis in R using an example dataset
- Goal: create at least one figure provide a biological interpretation of the data

10:45 am – 11:00 am – Break

11:00 am – 12:00 pm – Small Group Exercise Exploratory Analysis in R (Katja Koeppen, Tom Hampton, Britton Goodale, The Geisel School of Medicine at Dartmouth, MCBI)

- Group presentations of results

12:00 pm – 1:00 pm – Lunch (Dining Hall)

1:00 pm – 3:00 pm Accessing NLM with R and UNIX (Zhongyou Li, The Geisel School of Medicine at Dartmouth, MCBI)

- Introduction to the National Library of Medicine's EDirect Data Portal
- Introduction to UNIX
- Exercise: Downloading Gene Expression Omnibus with EDirect
- Visualization of Results in R: Word Clouds

3:00 pm – 3:30 pm – Break

3:30 pm – 6:00 pm – Applied Bioinformatics Consultation Clinic (Faculty, MCBI)

- Bring us your data, we'll help you out

6:00 pm – 8:00 pm – Lobster Bake (Dining Hall)

Thursday, July 12, 2018. Machine Learning

7:00 am – 8:00 am Continental Breakfast (Dining Hall)

8:00 am – 9:00 am – Beyond What is Known – Machine Learning (Tom Hampton, The Geisel School of Medicine at Dartmouth, MCBI)

- What makes machine learning special: Prediction
- Concepts: Clustering, Classification, Regression
- Supervised and Unsupervised Approaches
- Limits to machine learning: The Gradient Descent Example
- How to generate, quantify and test models
- R packages for Machine Learning

9:00 am – 9:15 am – Break

9:15 am – 10:15 pm – Hands-on Machine Learning in R (Katja Koeppen, The Geisel School of Medicine at Dartmouth, MCBI)

- Clustering
- Classification
- Regression
- Random Forest

10:15 am – 11:15 pm – Advanced Machine Learning Topics (Jaclyn Torino, University of Pennsylvania, MCBI)

11:15 am – 12:00 pm – Course Summary & Evaluations (MCBI)

12:00 pm – 1:00 pm – Lunch & Departure (Dining Hall)

Faculty

Course Directors: Ben King, Ph.D. (University of Maine) and Bruce A. Stanton, Ph.D. (The Geisel School of Medicine at Dartmouth)

Chris Dagdigian, Ph.D. Co-Founder and Director of Infrastructure, BioTeam.

<https://bioteam.net/bio/chris-dagdigian/>

Britton Goodale, Ph.D. Postdoctoral Fellow, Department of Microbiology and Immunology, The Geisel School of Medicine at Dartmouth.

<http://www.dartmouth.edu/~toxmetal/program-resources/training/brittongoodalebio.html>

Thomas Hampton, Ph.D. Senior Bioinformatics Analyst, Department of Microbiology and Immunology, The Geisel School of Medicine at Dartmouth.

<http://www.dartmouth.edu/~lbcobre/faculty/thomashampton.html>

Ben King, Ph.D. Assistant Professor of Bioinformatics, University of Maine.

<https://umaine.edu/biomed/home/faculty/benjamin-king/>

Katja Koeppen, Ph.D. Research Scientist, Department of Microbiology and Immunology, The Geisel School of Medicine at Dartmouth.

<http://www.dartmouth.edu/~kkoeppen/>

Diana Lampu, Ph.D. Vice President of Global Program Management and Strategic Planning, Homology Medicines, Inc. <https://www.linkedin.com/in/diana-lamppu-1b11616/>

Zhongyou Li, Graduate Student, Department of Microbiology and Immunology, The Geisel School of Medicine at Dartmouth.

Steve Munger, Ph.D., Assistant Professor, The Jackson Laboratory.

<https://www.jax.org/research-and-faculty/faculty/steven-munger>

Bruce A. Stanton, Ph.D. Andrew C. Vail Professor of Microbiology and Immunology, The Geisel School of Medicine at Dartmouth.

<https://geiselmed.dartmouth.edu/faculty/facultydb/view.php?uid=193>

W. Kelley Thomas, Ph.D., W. Kelley Thomas, Ph.D., Hubbard Professor in Genomics and Director, Department of Molecular, Cellular and Biomedical Sciences, University of New Hampshire. <https://colsa.unh.edu/faculty/thomas>

Jaclyn Torino, Ph.D., Postdoctoral Fellow, Integrated Genomics Laboratory, University of Pennsylvania. <http://www.jaclyn-taroni.com/>