



BST434: Genomic Data Analysis Spring 2020

Room: Zoom

Time: MW 9:10-10:50am

Phone: (585) 273-3177

Email: matthew_mccall@urmc.rochester.edu

Office Hours: by appointment

Course website: Slack

Prerequisites: {BST430 + BST411} or {DSC 462 + DSC 465} or {permission of instructor}

Basic Information

Classes will be held weekly by Dr. McCall with occasional guest lectures on specific topics.

Office hours can be arranged via email, phone, or slack. After the introductory period, the course will generally consist of a lecture on Monday and discussion of an article and working through the vignette of the corresponding R package on Wednesday.

Course Description

This course provides an introduction to techniques used in modern genomic experimentation and the corresponding statistical methods and software available to visualize, analyze, and interpret these data. Specific topics include mRNA/microRNA expression, copy number variants, single nucleotide variants, DNA methylation, and microbial abundance.

Course Aims and Objectives

Students will gain an in-depth knowledge of modern genomic experimentation and the corresponding statistical methods and software available to visualize, analyze, and interpret these data. Additionally, students will learn how to conduct reproducible data analyses and generate reproducible reports.

Course Policies and Expectations

Students will be responsible for completing weekly assignments, respectfully participating in classroom discussion, performing an in-depth exploration of a particular topic of interest, and presenting their topic near the end of the course. Attendance in the course is mandatory; students missing more than 2 lectures without pre-approval or documented illness will receive a reduced score proportional to the number of absences.

Materials and Access

There are no required texts; however, Quickstart Molecular Biology: An Introduction for Mathematicians, Physicists, and Computational Scientists may be a helpful reference for biological concepts.

Assignments and Grading Procedures

Attendance: 10%

Participation / weekly assignments: 60%

Final project & presentation: 30%



Academic Integrity

Academic integrity is a core value of the University of Rochester. Students who violate the University of Rochester University Policy on Academic Honesty are subject to disciplinary penalties, including the possibility of failure in the course and/or dismissal from the University. Since academic dishonesty harms the individual, other students, and the integrity of the University, policies on academic dishonesty are strictly enforced. For further information on the University of Rochester Policy on Academic Honesty, please see the Jurisdiction and Responsibility for Academic and Nonacademic Misconduct section in the Regulations and University Polices Concerning Graduate Studies <http://www.rochester.edu/GradBulletin/PDFbulletin/Regulations.pdf>

Accommodations for Students with Disabilities

Students needing academic adjustments or accommodations because of a documented disability must contact the Access Services Coordinator. For information regarding access services and support at SMD, please refer to our webpage: <https://www.urmc.rochester.edu/education/graduate/current-students/disability-supports-services.aspx>

Final Projects

The final project will consist of a written report (2-4 pages) and an oral presentation (~20 minutes). Both the written report and oral presentation should describe the biological significance, methodological challenge(s), and the current state of the field. The written report should be prepared similarly to the Introduction section of a manuscript. Students will select a topic by the end of the 2nd week of the course. No more than two students will be permitted to select the same topic. By the end of the 4th week, students should have identified 3-5 relevant papers that will form the foundation of their project. Students are encouraged to discuss their projects with other students and the instructor via Slack or during office hours. Presentations will occur during the final weeks of class. Written reports will be due on the last day of class.

Potential Project Topics

Challenges of quantifying microRNA expression
Methods to model cell type mixtures
Alternatives to analyzing counts
Compositional analysis of sequencing data
Estimation of cellular networks
Multi-omics data integration
Chromatin Accessibility (ATAC-seq)
Cell morphology
Expression quantitative trait loci (eQTL)
Co-expression and covariation of genomic features



Course Schedule

01/13	Review of R/Rstudio and Reproducible Research
01/20	Introduction to Molecular Biology
01/25	Brief history of Genomics
01/27	High-throughput sequencing
02/01 & 02/03	Introduction to Bioconductor Paper: Orchestrating high-throughput genomic analysis with Bioconductor Packages: SummarizedExperiment, GenomicRanges
02/08 & 02/10	RNA Expression Paper: Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2 Package: DESeq2
02/15 & 02/17	Distance and Normalization Paper: Visualizing Data using t-SNE Packages: Rtsne, DESeq2, iSEE
02/22 & 02/24	Batch Effects Paper: Preserving biological heterogeneity with a permuted surrogate variable analysis for genomics batch correction Package: sva
03/01	Project proposal lightning talks
03/03	No Class (SMD study break)
03/08 & 03/10	Hierarchical Models in Genomics Paper: DEXUS: identifying differential expression in RNA-Seq studies with unknown conditions Package: dexus
03/15 & 03/17	DNA Methylation Paper: Detection and accurate false discovery rate control of differentially methylated regions from whole genome bisulfite sequencing Package: dmrseq
03/22 & 03/24	Copy Number Variants Paper: Modeling Read Counts for CNV Detection in Exome Sequencing Data Package: exomeCopy
03/29 & 03/31	Protein-DNA Binding Paper: csaw: a Bioconductor package for differential binding analysis of ChIP-seq data using sliding windows Package: csaw
04/05 & 04/07	Microbiome Analysis Paper: DADA2: High-resolution sample inference from Illumina amplicon data Package: dada2
04/12 & 04/14	Single Cell Sequencing (guest lecture Andrew McDavid)
04/19 & 04/21	Genome Editing (guest lecture Aslihan Petenkaya)
04/26 & 04/28	Final Presentations