Fall 2020

Human Molecular Biologic Markers Data Analysis Laboratory
QBS132-2

Angeline S. Andrew, PhD

Summary:
This course covers both the computational analysis of data derived from human tissue samples and the biological basis of the features being assessed. The course covers a variety of modern, high-throughput technologies used in translational research. Molecular biologic markers will be studied in the context of both inherited variations, and lifestyle or environmental exposures. Students will practice applying various bio-statistical analytic approaches to molecular data in the context of human health studies.

Learning objectives of the course:

- Understand the biological motivation for high-throughput experiments, the particular types of challenges, artifacts, and biases encountered in working with specific types of molecular markers and tissues, and learn approaches for overcoming them.
- Gain practical experience working with a diversity of molecular marker dataset types.
- Discuss the similarities and differences among various analytic pipelines and identify areas in need of novel approaches / method development.
- Gain experience working with the programming language R, and with additional specialized analytic tools that have been developed for analysis of molecular marker data.

Class Meetings: Offered Fall term, one hour per week. Additional time will be required for completing the analysis exercises. (0.5 credits): Thursdays 8:30-9:30 am in Rubin xxx.

Assignments: Students will be expected to complete weekly analyses of the example molecular biologic marker datasets using a variety of programmatic approaches and software tools, as detailed in the Schedule below. Results / output showing that you ran the assigned example analyses will be submitted via email *within one week* of the assigned date.

Assigned	Introductory material	Activity	Dataset
9/17	Dose-response relationships. Assessment of inherited genetic variations.	Disease risk factor analysis: Exposures, genetic variation	Lung W1
9/24	Epistasis and the biological basis of statistical interactions.	GxG, GxE interaction identification tools	Lung W2
10/1	Arrays as relative measures of mRNA abundance. Context of biological pathways.	Analyzing gene expression array data, map to pathways	Bladder W3
10/8	Sequencing analysis workflow. Transcript abundance, alternative splicing, structural variation.	RNA-seq: analysis of mapped sequence reads	Oral W4
10/15	Expression regulation through Protein-DNA binding, chromatin modifications. Peak detection.	Protein-DNA interactions: analysis of ChIP-seq data	Breast W5
	Attend a talk at Life Sciences Symposium (runs 9am-4pm in Life Sciences Center, Hanover)		
10/22	Epigenetic regulation via DNA methylation, long range interactions, histone modification.	DNA methylation dysregulation: analysis of methylation array data, dataset normalization	Colon W6
due 10/29	**Student initiated project analysis plan:**~250 words	**Rough DRAFT of analysis plan**	Student choice
10/29	Microbiome analysis.	Analysis of 16s rRNA sequence reads	Gut W7
11/5	Single cell genomics	Analysis of single cell RNA seq data	PBMC W9
11/12	Discrimination analysis	Classify diabetes status	Diabetes
11/19	**Student initiated project presentations**	**Student initiated project materials due by email**	Student choice

Fall 2020

Student Initiated Project: Students will choose ONE of the following project options (A or B)

Choice A. New activity module.

- Students will identify an analytic software tool to assess a molecular marker data type not already covered in the course. Following the same structure as the existing course activities, they will construct an activity module and apply it to a dataset of their choice. Activity plan: ~250 words, *due week 7*.
- Students will adapt code to demonstrate the utility of the chosen software tool.
- Final product will be an activity package sent by email: Activity package will include:
    - 1) Brief Background – Introductory Powerpoint slides explaining context of the data type, including how raw data is generated, and the analysis need being fulfilled
    - 2) Raw dataset used to run demonstration.
    - 3) File with analytic code (e.g.an R-file), fully commented with explanatory information.
    - 4) Raw results of the analysis, (e.g. tables, figures) with interpretive legends or titles

OR

Choice B. New analysis of marker vs. exposure/disease hypothesis.

- Students will generate a hypothesis that is testable using existing data from one of the molecular marker datasets provided in this course. They will construct an analysis plan to test this hypothesis. Analysis plan: ~250 words, *due week 7*.
- Students will conduct an analysis to test their hypothesized molecular marker vs. exposure/disease relationship.
- Final product will be in written report format (double spaced), sent by email: Report will include:
    - 1) Brief Background – context of problem, gap in knowledge (~1 page),
    - 2) Hypothesis,
    - 3) Statistical Analysis Methods description (~1/2 page),
    - 4) Raw results of the analysis (~3 tables, or figures), with interpretive legends or titles,
    - 5) Bulleted interpretation / conclusions (~3 bullets),
    - 6) File with analytic code (e.g.an R-file), fully commented

Grading policies:

- 10% Effort and attendance
- 50% Completion of assigned analyses
- 40% Student Initiated Project

Fall 2020

Institutional Policies

Some students may wish to take part in religious observances that occur during this academic term. If you have a religious observance that conflicts with your participation in the course, please meet with me before the end of the second week of the term to discuss appropriate accommodations.

Students with disabilities who may need disability-related academic adjustments and services for this course are encouraged to see me privately as early in the term as possible.

At Dartmouth, we value integrity, responsibility, and respect for the rights and interests of others, all central to our Principles of Community. We are dedicated to establishing and maintaining a safe and inclusive campus where all have equal access to the educational and employment opportunities Dartmouth offers. We strive to promote an environment of sexual respect, safety, and well-being. In its policies and standards, Dartmouth demonstrates unequivocally that sexual assault, gender-based harassment, domestic violence, dating violence, and stalking are not tolerated in our community. The Sexual Respect Website (https://sexual-respect.dartmouth.edu) at Dartmouth provides a wealth of information on your rights with regard to sexual respect and resources that are available to all in our community. Please note that, as a faculty member, I am obligated to share disclosures regarding conduct under Title IX with Dartmouth's Title IX Coordinator. Confidential resources are also available, and include licensed medical or counseling professionals (e.g., a licensed psychologist), staff members of organizations recognized as rape crisis centers under state law (such as WISE), and ordained clergy (see https://dartgo.org/titleix_resources). Should you have any questions, please feel free to contact Dartmouth's Title IX Coordinator or the Deputy Title IX Coordinator for the Guarini School. Their contact information can be found on the sexual respect website at: https://sexual-respect.dartmouth.edu.