Fall 2011 BIOL B216: Introduction to Genomics and Bioinformatics

Joshua Shapiro
office: 112 Park Science
phone: 610-526-5099
email: jshapiro01@brynmawr.edu

Course Objectives

The proliferation of genomic data has fundamentally transformed biology, making previously laborious and expensive experiments easier and cheaper, and enabling entirely new areas of research and fundamentally altering all fields of biology and medicine. This course will introduce you to the questions that can be asked and answered with genomic data, and to the computational tools available to analyze that data. The primary goals are:

- to become familiar with the tools and databases available for bioinformatic analysis, with an appreciation of the quantitative concepts that underly those tools.
- to learn how genomic data are being used to provide new insights throughout biology and medicine.
- to develop the ability to formulate and investigate genomic research questions, and to effectively communicate your questions, methods, and results.

Course Description

Lectures in the first half of the class will introduce some of the common techniques and algorithms used in genomic analysis, including sequence alignment, BLAST, gene expression profiling, and prediction of protein structure and gene function. As the course progresses, we will move from techniques to applications and results. We will explore how genomic data has been used to explore a variety of topics in biology, such as evolutionary history, identifying genetic causes of disease, the genomics of cancer, and the emerging field of metagenomics. Since genomics is a new and rapidly-changing field, much of this latter portion will consist of topics chosen from the primary literature, which we will discuss in class.

The computational lab is an essential component of this course. That said, there is no assumption of previous experience beyond knowing how to move files around and use the web, word processors and spreadsheets. In the first set of labs, you will become comfortable with tools for sequence alignment, gene expression analysis, and handling of genome-scale data sets. The second half of the lab will be devoted to developing and completing a final project; you and a partner will work together to formulate a question that can be answered using publicly available data, then use the skills you have developed, and whatever new tools you find useful, to work toward an answer to your question.

Lecture

MWF 10 – 11 am
Park 10

Lab

Wednesday 1 – 4 pm
Park 10

Office Hours and Contact info

MF 2–3:30PM, and by appointment.

Feel free to email me any time, but see notes below. I will answer emails (and Piazza queries) fairly quickly during normal business hours, but generally less quickly in the evenings and weekends (possibly not until the next day).

We will be using Piazza.com for Q&A during the course, so any questions that might be of general interest should be asked there, rather than by email. This will give you the opportunity to see the questions that other students have already asked, and to answer each other's questions. Active participation on Piazza (questions and/or answers) will result in extra credit. To sign up, go to: www.piazza.com/brynmawr/fall2011/biolb216 (*if you are not a Bryn Mawr student, I will have to add you manually. Email me.)

Textbook and Readings

Bioinformatics and Functional Genomics by Jonathan Pevsner

Additional readings from the primary literature and reviews will be posted on Moodle. These will mostly be for recent topics that the textbook does not yet cover, especially as we progress later into the semester.

Evaluation and Grading

Moodle quizzes (10) 10%
Take-home exams (2) 30%
Final exam.....20%

Lab writeups (4) 20%
Final Project 20%

Late assignments will incur a 10% penalty per day unless arrangements are made for an extension at least 24 hours prior to the due date.

Moodle Quizzes

Quizzes will be posted on Moodle after the Friday class of most weeks, due the following Monday before class (no quiz the first week or weeks with exams). Each will consist of a few short answer questions, with 50% of the grade based on completion of the quiz on time.

Exams

Mid-semester exams will be take-home format, and 1-2 hours in length. The final exam will be cumulative, but will emphasize topics covered in the later part of the course. Any scheduling conflicts must be addressed well in advance of the exam. Make-up exams will not be given, except in cases of illness or unforeseen circumstances, or with prior arrangement.

Labs

Lab activities will be entirely computational. If you have your own laptop, please bring it to every lab; loaner laptops are available for those who do not have their own.

Lab writeups will be short (1-3 page) reports of results obtained in lab, including responses to questions presented in the lab instructions. These should be completed individually.

Final Project

The final lab project will be conducted in teams of 2, with the two lab members collaborating on all aspects of the project. There will be three phases: A proposal, preliminary results (on which you will receive feedback and suggestions for further work), and a final paper and oral presentations. The final paper should include clear statements of which aspects of the analysis and writing were performed by each team member.

Accommodations

Students who think they may need accommodations in this course due to the impact of a learning, physical, or psychological disability are encouraged to meet with me privately early in the semester to discuss their concerns. Students should also contact Stephanie Bell, Coordinator of Access Services (610-526-7351 or sbell@brynmawr.edu), as soon as possible, to verify their eligibility for reasonable academic accommodations. Early contact will help to avoid unnecessary inconvenience and delays.

Schedule

(very much subject to change through the semester, depending on progress and interests)

Week 1	*Lecture*	Why study genomes? Principles and background
Aug 29 – Sept 2	*Reading*	Pevsner Chapter 1
	Lab	No lab: shopping period
Week 2	*Lecture*	Sequence alignment of DNA and proteins
Sept 5 – 9	*Reading*	Pevsner Chapters 2-3
	Lab	Introduction to online tools: websites and databases
Week 3	*Lecture*	BLAST and its relatives
Sept 12 – 16	*Reading*	Pevsner Chapters 4-5
	Lab	Project Ia: Identifying gene families with BLAST
Week 4	*Lecture*	Multiple Sequence Alignment
Sept 19 – 23	*Reading*	Pevsner Chapter 6
	Lab	Project Ib: Multiple sequence alignment
Week 5	*Lecture*	Proteomics: structure and function
Sept 26 – 30	*Reading*	Pevsner Chapters 10-11
	Lab	Project Ic: Building gene phylogenies
		Exam I
Week 6	*Lecture*	Whole genome sequencing
Oct 3 – 7	*Reading*	Pevsner Chapter 13
	Lab	Introduction to R; reading data files and calculating statistics

Fall Break

Week 7	*Lecture*	Analysis of global gene expression
Oct 17 – 21	*Reading*	Pevsner Chapters 8-9
	Lab	Project IIa: Differential gene expression
Week 8	*Lecture*	Quantitative trait locus mapping
Oct 24 – 28	*Reading*	Primary literature: TBA
	Lab	Project IIb: Expression clustering and gene set analysis
Week 9	*Lecture*	Genome-wide association studies
Oct 31 – Nov 4	*Reading*	Primary literature: TBA
	Lab	Individual Projects: brainstorming and planning
Week 10	*Lecture*	High-throughput sequencing
Nov 7 – 11	*Reading*	Primary literature: TBA
	Lab	Individual Projects: data collection and initial analysis
		Project Proposal due Nov 7
		Exam II

Week 11	*Lecture*	Functional genomics
Nov 14 – 18	*Reading*	Pevsner Chapter 12
	Lab	Individual Projects: analysis
Week 12	*Lecture*	Epigenomics
Nov 21 –23	*Reading*	Pevsner Chapter 16
Thanksgiving	*Lab*	Individual project meetings, no lab **Preliminary results due Nov 21**
Week 13	*Lecture*	The human genome
Nov 28 – Dec 2	*Reading*	Pevsner Chapters 19-20
	Lab	Individual Projects: final analysis, revision
Week 14	*Lecture*	Metagenomics
Dec 5 – Dec 7	*Reading*	Primary literature: TBA
	Lab	Project Presentations **Final lab papers due Dec 7**