Computational Biology Curriculum

Last updated: Fall 2018

Foundational courses

The following are suggested courses for students with little prior experience who want to become computationally competent.

Introductory Programming

GS 559 - Introduction to Statistical and Computational Genomics
Introductory Python for absolute beginners. Emphasis on basic probability and statistics, introduction to computer programming, and relevant web databases.

MCB 517A - Tools for Computational Biology
Introduction to practical data analysis for computational biology using R, Python, and the command line.

BIOL 519 - Data Science for Biologists
Focuses on the basics of data wrangling, data analytics, statistics and visualization. In-class exercises, labs, and homeworks are in Python.

CSE 583 - Software Development For Data Scientists
Provides students outside of CSE with a practical knowledge of software development that is sufficient to do graduate work in their discipline. Modules include Python basics, software version control, software design, and using Python for machine learning and visualization.

Advanced Programming

GS 540 - Introduction to Computational Molecular Biology: Genome and Protein Sequence Analysis
This class will vastly improve your programming ability, and is worth the effort if you truly want to be a computational(ly competent) biologist. It does entail a heavy workload; plan on 15-20 hours/week.

GS 541 - Introduction to Computational Molecular Biology: Molecular Evolution
More of a survey course. Nice exposure to a lot of methods in computational bio overall. Far less work than GS 540, but still requires decent programming skills to complete. Note this is the foundational required course required for the highly recommended CMB certificate program (see below).

Statistics

There is not an ideal series. Pick the option right for you from these or from another department's catalog. *Note that BIOST 517/518 is discouraged. The summer stats course common among MCB students is probably insufficient for the kinds of analyses common in computational biology.*

GS 560 - Introduction to Statistical Genomics
Provides the key statistical concepts and methods necessary for extracting biological insights from data. In-class exercises and homework assignments are written in R.

STAT 509 - Introduction to Mathematical Statistics
Examines methods, tools, and theory of mathematical statistics

STAT 512-513 - Statistical Inference
Covers random variables, hypothesis testing, and decision theory.

Electives

PABIO 536 (co-listed as BIME 536/PHG 536) - Bioinformatics and Gene Sequence Analysis.
Non-programming intro to common databases, etc.

EPI 554 - Intro to Epidemic Modeling
Great introductory class to programming deterministic mathematical models for infectious diseases.

CSE 544 - Data Management

CSE 546/**STAT 535** - Statistical Learning
Covers statistical learning over discrete multivariate domains, exemplified by graphical probability models. Emphasizes the algorithmic and computational aspects of these models. *Note: content in past quarters has emphasized theory over practical application; check with the scheduled professor to confirm the style currently offered.*

CSE 512 - Data Visualization

CSE 527 - The computer science version of GS 540 for students coming from a CS background. Introduces computational methods for understanding biological systems at the molecular level. Problem areas include mapping and sequencing, sequence analysis, structure prediction, phylogenetic inference, and regulatory analysis.

CSE 391 - System and Software Tools
Another undergraduate course that does not count for MCB credit, but covers highly

useful utilities for working in comp bio (e.g. command-line interface, file and string manipulation, regular expressions, and using version-control systems.)

CSE 586 – Intro to synthetic biology
For students with a preexisting background in computer science or computational biology.

BIOST 578 - Bayesian Statistics (special topic)
Not offered regularly; check for availability.

GS 570 - Phylogenetic Inference
Not currently offered.

Additional Resources

Programming

fredhutch.io
Short courses on Python, R, the command line, and other utilities in computational biology. Courses are offered at the Hutch and fill quickly. Sign up for their email list.

Code Academy
Online introduction to Python

UW Computer Science Course Guide
Guide from the CS department on which programming courses to take. Note that these are undergraduate level courses (do not count for MCB credit), but may help bridge the gap between GS559 and GS540 for students who are serious about learning to program.

Certificate programs

Certificate programs pull students from many departments, and can be a great way to tap into a learning community of peers with shared interests. It can also help structure your curriculum and adds a line to your CV. MCB supports students' participation in the following programs.

- Computational Molecular Biology (CMB)
- eScience Advanced Data Science Option
- Statistical Genetics