Genomics BIOD25H3 — Syllabus

Instructor: Dr. Guillaume J. Filion
Prerequisite: BIOC15H3 — Genetics
Lectures (online): Wednesday 14.00–16.00
Office hours (online): Thursday 12.00-13.00

What is genomics?
On April 14 2003, the Human Genome Project was officially completed. This announced the dawn of the *genomic era*, an era where scientists look for answers about the past, the present and the future of organisms in their genomes.

Genomics is about sequencing genomes and analyzing them. By extension, it also means acquiring “big data” from cells (*e.g.*, transcriptomics or proteomics). This course is technology-centric and puts as much emphasis on *what* we have learned about genomes as on *how* we have learned it. The first part of the course covers the methods to sequence, assemble, annotate and query genomes. The second part covers the applications of full genome sequencing in forensics, health, history, agriculture and biotechnology.

Learning outcomes
In this course, you will revisit a concept that you encountered in BIOB11H and BIOC15H: the gene. We are going to build on what you learned about transcription, splicing, translation, gene regulation, mobile DNA, sequencing methods and mutations.

More specifically, after this course, you will be able to:

1) describe modern sequencing technologies,
2) illustrate the use of those technologies for different purposes,
3) define the features of file formats for sequencing data,
4) use software to map reads to a genome,
5) examine mapping results and comment on the quality,
6) describe methods to compare genomes and to retrace their evolution,
7) describe methods to assign functions to genes,
8) name important discoveries made from omics approaches,
9) explain the principles of genome organization in the different kingdoms,
10) apply the principles of sequencing to forensic and medical problems,
11) describe the theoretical foundations of GWAS,
12) analyze the results of GWAS data,
13) discuss the ethics of personalized genomics data,
14) summarize academic papers in the field of genomics,
15) write large-audience academic content about genomics.

Sure, but what's going to be at the exam?
One of the main objectives of this course is to make you familiar with the primary research literature. Instead of a final exam, you will have to write a blog post about a significant finding in genomics. That's also a way to build a visible curriculum for yourself and reach out to the community through your work. The rubric will be detailed in class so that you get a clear idea of what you have to write and how.

A second objective of this course is to give you the basic knowledge to analyze high-throughput sequencing data. The mid-term exam will require you to use the high-performance cluster of the university and peek under the hood to analyze some real sequencing data. It's OK if you have no experience with programming or with Linux; we'll cover what you need during the classes.

A third objective of this course is to give you a working knowledge of genomics in practical applications. What better way to train yourself than solving some genetic whodunnit? You will get the chance to test your technical and logical skills on some forensic cases.

There will be 10 low-stake quizzes so that you can make sure that you acquire and understand the content of the course. Quizzes have more than 10 questions, but I will count only your top 10 answers. In other words, it's OK to make a few mistakes: you will get the maximum grade for a quiz if you get 10 answers right.

Below are the weights of the different activities:

Weekly quizzes: 15%
Whodunnit: 10%
Mid-term exam: 25%
Blog post milestone 1: choose an article: 5%
Blog post milestone 2: title / TL;DR of the post: 5%
Blog post milestone 3: figure(s) of the post: 10%
Blog post milestone 4: final version of post: 30%

Course structure
I am the instructor (Guillaume Filion) and the TA for this course is Michelle Harwood. I will be delivering most of the lectures and answering your questions. Michelle will give lecture #10 on genome-wide association studies (GWAS) and she will help you with the content if you have some difficulties. We will explain in class how to contact us and how to attend office hours.

The course consists of 2 hours of online lectures per week. The typical session will include some discussions and Q&A on the material to read, followed by the lecture proper, complementing the material to read and introducing new concepts.

The fifth session is a hand-on tutorial using the "Teach" computer cluster of the university, to which you will connect from your computer.

Penalties and bonus points
To reflect your future experience in the professional world, there will be no timed evaluation. All the evaluations will be "take home" with a defined deadline.

Work submitted after the deadline is accepted, but there will be a 20% penalty per fraction of 24 hours after the deadline. For instance, if the deadline is January 19, 2020 at 21:00, a 20% penalty will be applied for handing in your work on January 19, 2020 at 21:01. The penalty rises to 40% from January 20, 2020 at 21:01 *etc*. It goes without saying that I invite you to submit your work well before the deadline to avoid penalties. If you have a registered disability, the deadline may be extended according to your needs; the penalty system remains the same but kicks in from a different deadline.

You have **2 jokers**, each removing a 20% penalty for late submission. They can be used for the same assignment in case the penalty is 40% or higher, or for different assignments. You do not need to say that you use a joker: if required, they will be applied automatically at the end of the semester to maximize your final grade.

It may be that you are unable to study for a prolonged period of time in ways that will affect your capacity to submit your work on time. In this case, you should notify me as soon as possible, and definitely before the deadline. Remember that the rules are meant to reward punctuality and to promote a climate of productivity; they are not meant to lower your motivation or make you feel under the weather. We will find fair ways to deal with every case requiring special attention.

If you want to improve your grades, you can get some bonus points in the following ways:

1) If you finish the course having submitted all your evaluations on time *without* using any joker, you will get a +5% bonus to the final grade.
2) You can proof-read the blog posts of your classmates and give them feedback about what is good and what needs improving. Each blog post you review will give you a +1% bonus to the final grade, with an upper limit of +5%.
3) If you participate in some extracurricular events related to genomics (*e.g.*, conference, seminar, online forum...), you show me some proof of participation in order to get a +1% bonus to your final grade per event, with an upper limit of +5%.

What else do I need to know?
We will be using the textbook *Genomes 4* by T.A. Brown. You will have to read some sections before the classes (see calendar below). During the classes, we will discuss the content together, cover some complementary knowledge and do some exercises. It's important that you come prepared, otherwise the classes will be too hard to follow.

January 13, 2020 — The human genome
Introduction, genomics, the Human Genome project.

Deadline: January 19, 2020, 21:00 pm
Submit quiz on lecture 1 — The human genome (1.5%)

Read textbook sections 1.1, 2.2, 4.1, 4.2 and 4.4.
January 20, 2020 — Sequencing technologies
Polymerase chain reaction (PCR), Sanger sequencing, high-throughput sequencing technologies, Illumina, Ion Torrent, PacBio, Nanopore.

Deadline: January 26, 2020, 21:00 pm
Submit quiz on lecture 2 — Sequencing technologies (1.5%)

Read textbook sections 1.2 (optional), 5.2, 5.3, 12.1, 12.2 and 12.4.
January 27, 2020 — Transcriptomics
Gene regulatory networks, splicing, noncoding RNAs, reverse transcription, quantitative RT-PCR, single-cell transcriptomics, spatial transcriptomics.

Deadline: January 31, 2020, 21:00 pm
Choose a research article to write a blog post about (5%)

Deadline: February 2, 2020, 21:00 pm
Submit quiz on lecture 3 — Transcriptomics (1.5%)

Read textbook sections 1.3 (optional), 5.4, and 6.1
February 3, 2020 — Databases and algorithms
Homology search (BLAST), read alignments (BWA), visualization (UCSC Genome Browser), FASTA, FASTQ and SAM formats, introduction to Linux and high-performance computing.

Office hour dedicated to troubleshooting connection to the Teach cluster

Deadline: February 9, 2020, 21:00 pm
Submit quiz on lecture 4 — Databases and algorithms (1.5%)

February 10, 2020 — Hands-on computer session
Map reads from high-throughput sequencing experiments and analyze the results.

Reading week

Deadline: February 22, 2020 21:00 pm
Midterm exam (take-home work on the Teach cluster, 25%)

Read textbook sections 4.3, 7.4, 9.2, 17.3

February 24, 2020 — Repeats and genome assembly

Cot curves, transposons and retrotransposons, integrated viruses, repeated elements, the assembly problem, physical maps, shotgun assembly, mixed reads.

Deadline: March 2, 2020, 21:00 pm

Submit quiz on lecture 6 — Repeats and genome assembly (1.5%)

Read textbook sections 3.2.

March 3, 2020 — DNA fingerprints and forensics

Genetic diversity, SNPs, RFLP, microsatellites, mitochondrial genome, ancient DNA, DNA spray, criminal cases.

Deadline: March 9, 2020, 21:00 pm

Submit quiz on lecture 7 — DNA fingerprints and forensics (1.5%)

Read textbook sections 8.1, 8.2, 8.3 and 9.1.

March 10, 2020 — Metagenomics

Ribosomal RNA, gut microbiome, sea microbiome, metagenome assembly, microbes and health.

Deadline: March 16, 2020, 21:00 pm

Submit quiz on lecture 8 — Metagenomics (1.5%)

Read textbook sections 16.1, 7.1, 7.2, 7.3, 18.1 and 18.2

March 17, 2020 — Evolutionary genomics

Genomes of prokaryotes and eukaryotes, multiple alignments, evolutionary trees, ancestral reconstruction, structured coalescent, origins of past epidemics.

Deadline: March 20, 2021, 21:00 pm

Submit title and TL;DR of your blog post (5%)

Deadline: March 23, 2020, 21:00 pm

Submit quiz on lecture 9 — Evolutionary genomics (1.5%)

Read textbook sections 3.3 and 3.4.

March 24, 2020 — Genome-wide association studies

Linkage disequilibrium, *haplotypes, applications of GWAS in agriculture and in health, personalized medicine, the missing heritability problem.*

Deadline: March 27, 2021, 21:00 pm

Submit figure(s) of your blog post (10%)

Deadline: March 30, 2020, 21:00 pm

Submit quiz on lecture 10 — Genome-wide association studies (1.5%)

Read textbook sections 16.2 and 17.1 and 17.2.

March 31, 2020 — Genome editing

DNA repair, knock-out, knock-in, CRISPR, gene therapy.

Deadline: April 4, 2020, 21:00 pm

Submit “whodunnit” (10%)

Deadline: April 6, 2020, 21:00 pm

Submit quiz on lecture 11 — Genome editing (1.5%)

Deadline: April 7, 2021, 21:00 pm

Submit final version of your blog post (35%)

Deadline: April 12, 2021, 21:00 pm

Submit feedback on other blog posts (bonus points)