

University of Guelph
College of Biological Science
Department of Molecular and Cellular Biology
COURSE OUTLINE

Genomic Methods for Bioinformatics, BINF*6110,
Winter, 2017

Teaching team

Terry Van Raay
SSC 3460
tvanraay@uoguelph.ca
Office Hours: By appointment

Steffen Graether
SSC 2255
graether@uoguelph.ca
Office Hours: By appointment

RNA-Seq with Laurence Tessier: tessierl@uoguelph.ca.
Please, prior to March 9, apply for a SHARCNET account. More instructions will be provided in class before then.
Please send your SHARCNET username to Laurence via email.
If you are using a computer with a Windows operating system, you will need to install PuTTY (<http://www.chiark.greenend.org.uk/~sgtatham/putty/>) and WinSCP (<https://winscp.net/eng/index.php>) prior to the first RNA-seq class (you will need these to access Sharcnet and transfer files between Sharcnet and your computer).
Please bring a laptop to all sessions. Please also install R (<https://www.r-project.org/>) on your computer.

Course schedule

Lectures: **Thursdays, 1:30 – 4:20 PM SSC 2315**

Learning goals and rationale

Overall, it is our objective for students to be able to generate information about their gene of interest using genomics related databases, use the tools to analyze large datasets such as RNA-seq, and use the tools to analyze protein structure datasets.

By the end of my course, students should be able to:

- 1) Explain the different types of information that can be obtained from a genome project (eg., aneuploidy or genome evolution).
- 2) Understand the tools in use to generate genomic data.

- 3) Integrate different databases, such as a genome browser and its associated databases (eg. Human Genome Browser) with other databases (eg., Genbank), and synthesize the various elements displayed in these databases.
- 4) Apply information gathered from various databases to a gene of interest.
- 5) Understand methods to manipulate genomes.
- 6) Use the UNIX command-line to run basic commands.
- 7) Manage large biological datasets such as RNA-Seq.
- 8) Use tools to convert raw RNA-Seq data to test a biological hypothesis.
- 9) Understand protein structure and the representation of protein structures.
- 10) Use online and command-line tools to analyze protein structures.

Course Resources

The following textbooks, which will be on reserve in the library:

Discovering Genomics, Proteomics and Bioinformatics, 2nd Edition, by A. Malcolm Campbell and Laurie J. Heyer

Genomes 3, T.A. Brown

The readings are meant to help those students with a weaker biochemistry background.

Protein structures: Any introductory biochemistry textbook that discusses amino acids, protein sequences, secondary and tertiary structure.

X-ray crystallography:

http://chemwiki.ucdavis.edu/Analytical_Chemistry/Instrumental_Analysis/Diffracton/X-ray_Crystallography

Protein NMR review:

Barrett, P. J. et al. The Quiet Renaissance of Protein Nuclear Magnetic Resonance. *Biochemistry* 52, 1303–1320 (2013).

Course Content

This course will be run using CourseLink.

Major course components:

- 1) Lecture
- 2) Assignments

Tentative Schedule

Week	Topics Covered in Lecture (Subject to Change)
1 January 12 th	What constitutes a genome? Structure of the Human Genome Genome Maps Annotation of a genome Databases
2 January 19 th	Genomic Web Servers Databases continued Microarrays Types of information generated from a microarray
3 January 26 th	Comparative Genomics Comprehensive Genomic analysis Genetic Manipulation
4 February 2 nd	A Brief Introduction to Protein Structure and Methods Amino acids, secondary and tertiary structure X-ray crystallography, Protein NMR
5 February 9 th	Protein Sequence Bioinformatics Pairwise alignment, multiple-sequence alignment, motif discovery
6 February 16 th	Protein Structure Bioinformatics 1 Visualization of protein structures Aligning protein structures Fold identification
February 20-24 th	Reading Week – no classes
7 March 2 nd	Protein Structure Bioinformatics 2 Docking proteins to ligands Modelling protein structures
8 March 9 th	Introduction to the UNIX command-line/BIO-tour
9 March 16 th	Introduction to mRNA Biology and RNA-Seq Quality Control and Assessment of RNA-seq
10 March 23 rd	Aligning sequences and mapping reads to a genomic sequence Finding and Using Reference Genomes
11 March 30 th	Review of genes, isoforms and splice variants Quantification of mRNA using RNA-Seq data
12 April 6 th	Understanding Biology using Gene Ontology Identification of genes with shared biological roles

Methods of Assessment

Assessment				
Form of Assessment	Weight of Assessment	Due Date of Assessment	Course Content /Activity	Learning Outcome Addressed
Assignment #1	5%	January 19 th	Lecture and Tutorial	2,3
Assignment #2	10%	January 26 th	Lecture and Tutorial	1,2,3,4
Assignment #3	10%	February 2 nd	Lecture and Tutorial	1,2,3,4,5
Courselink Quizzes	10%	TBD	Lecture and Tutorial	6, 7, 8
RNA-Seq Project	40%	TBD	Lecture and Tutorial	6, 7, 8
Structure Representation	10%	March 2 nd	Lecture and Tutorial	9, 10
Structure Alignment	15%	March 16 th	Lecture and Tutorial	9, 10

Assignment #1

Here you will use some basic Graphical User Interface (GUI, pronounced Gooy) tools to analyze some DNA that you will be provided. This will introduce you to the numerous databases that exist that contain information about the sequence.

Assignment #2

Students will be assigned a unique human gene and use the various online genomic databases to explore their gene of interest (GOI)

Assignment #3

A continuation of exploring unique elements about your GOI, including methods to manipulate your gene.

RNA-Seq Project: Analysis of RNA-Seq data for differential expression. Each student will be provided with the same raw read counts, and each student will be assigned a unique research hypothesis. After determining an appropriate model, each student must analyze the data appropriate software to identify differentially expressed genes. You will then use a variety of tools to identify the genes (or groups of genes) that pertain to your research hypothesis. Be sure to place your findings in the context of established biology, and describe how your findings relate to other relevant work in the field. Your paper should be no more than 15

pages, double-spaced not including figures and references, and should follow standard journal format (Abstract, Introduction, Materials & Methods, Results, Discussion, References) with figures attached at the back of the document.

Software used:

FastQC:

Andrews S: FASTQC. A quality control tool for high throughput sequence data. [<http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/>].

STAR:

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29:15-21.

Bowtie and Bowtie2:

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10, R25.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359.

Tophat, Tophat2, and Cuffdiff:

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc.* 7, 562–578.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14, R36.

HTSeq:

Simon Anders. HTSeq: Analysing high-throughput sequencing data with Python. <http://www-huber.embl.de/users/anders/HTSeq/>, 2011

EdgeR:

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140.

DEXSeq:

Anders S, Reyes A, Huber W. Detecting differential usage of exons from RNA-seq data. *Genome Res.* 2012;22:2008-17.

Cytoscape:

Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, Christmas R, Avila-Campilo I, Creech M, Gross B, Hanspers K, Isserlin R, Kelley R, Killcoyne S, Lotia S, Maere S, Morris J, Ono K, Pavlovic V, Pico AR, et al. Integration of biological networks and gene expression data using Cytoscape. Nat Protoc. 2007;2:2366-82.

PANTHER:

Mi H, Muruganujan A, Thomas PD. PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. Nucleic Acids Res. 2013;41:D377-86.

Structure Representation Assignment:

Students will learn how to obtain protein structure files from the PDB database (<http://rcsb.org>) and to show key structural features of proteins.

Software used:

Pymol: http://www.pymolwiki.org/index.php/Main_Page

Structure Alignment Assignment:

A central dogma of biochemistry is that the sequence of a protein determines its structure, which in turn determines its function. Often the structural conservation occurs even after sequence conservation is no longer obvious. In this assignment you will align a family of functionally related proteins and create figures to highlight their similarity and differences.

Software used:

Dalilite: http://ekhidna.biocenter.helsinki.fi/dali_lite/start

Course and University Policies

When You Cannot Meet a Course Requirement

When you find yourself unable to meet an in-course requirement because of illness or compassionate reasons, please advise the course coordinator in writing, with your name, id#, and e-mail contact, and be prepared to provide supporting documentation. See the undergraduate calendar for information on regulations and procedures for Academic Consideration: <http://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-ac.shtml>

Late assignments will be penalized 10 percentage points for every 24 hour period starting at 00:01 AM on the day after the assignment is due. Assignments will not be accepted after 5 days. If there is a valid reason why this cannot be achieved see below for the university guidelines.

Accessibility

The University of Guelph is committed to creating a barrier-free environment.

Providing services for students is a shared responsibility among students, faculty and administrators. This relationship is based on respect of individual rights, the dignity of the individual and the University community's shared commitment to an open and supportive learning environment. Students requiring service or accommodation, whether due to an identified, ongoing disability or a short-term disability should contact Student Accessibility Services (and Centre for Students with Disabilities) as soon as possible.

For more information, contact Student Accessibility Services at 519-824-4120 ext. 56208 or email <mailto:csd@uoguelph.ca> or see the website: <http://www.csd.uoguelph.ca/csd/>

Academic Misconduct

The University of Guelph is committed to upholding the highest standards of academic integrity and it is the responsibility of all members of the University community – faculty, staff, and students – to be aware of what constitutes academic misconduct and to do as much as possible to prevent academic offences from occurring. University of Guelph students have the responsibility of abiding by the University's policy on academic misconduct regardless of their location of study; faculty, staff and students have the responsibility of supporting an environment that discourages misconduct. Students need to remain aware that instructors have access to and the right to use electronic and other means of detection.

Please note: Whether or not a student intended to commit academic misconduct is not relevant for a finding of guilt. Hurried or careless submission of assignments does not excuse students from responsibility for verifying the academic integrity of their work before submitting it. Students who are in any doubt as to whether an action on their part could be construed as an academic offence should consult with a faculty member or faculty advisor.

The Academic Misconduct Policy is detailed in the Undergraduate Calendar: <http://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-amisconduct.shtml>

E-mail Communication

As per university regulations, all students are required to check their <mail.uoguelph.ca> e-mail account regularly: e-mail is the official route of communication between the University and its students.

Drop Date

The last date to drop one-semester courses, without academic penalty, is the 40th class day. To confirm the actual date please see the schedule of dates in the Undergraduate Calendar. For regulations and procedures for Dropping Courses, see the Undergraduate Calendar:

<http://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-drop.shtml>

Copies of out-of-class assignments

Keep paper and/or other reliable back-up copies of all out-of-class assignments: you may be asked to resubmit work at any time.

Recording of Materials

Presentations which are made in relation to course work—including lectures—cannot be recorded or copied without the permission of the presenter, whether the instructor, a classmate or guest lecturer. Material recorded with permission is restricted to use for that course unless further permission is granted.

Campus Resources

The Academic Calendar is the source of information about the University of Guelph's procedures, policies and regulations which apply to undergraduate, graduate and diploma programs:

<http://www.uoguelph.ca/registrar/calendars/index.cfm?index>

If you are concerned about any aspect of your academic program:

- make an appointment with a program counsellor in your degree program. <http://www.bsc.uoguelph.ca/index.shtml> or <https://www.uoguelph.ca/uaic/programcounsellors>

If you are struggling to succeed academically:

- There are numerous academic resources offered by the Learning Commons including, Supported Learning Groups for a variety of courses, workshops related to time management, taking multiple choice exams, and general study skills. You can also set up individualized appointments with a learning specialist. <http://www.learningcommons.uoguelph.ca/>

If you are struggling with personal or health issues:

- Counselling services offers individualized appointments to help students work through personal struggles that may be impacting their academic performance. <https://www.uoguelph.ca/counselling/>
- Student Health Services is located on campus and is available to provide medical attention. <https://www.uoguelph.ca/studenthealthservices/clinic>
- For support related to stress and anxiety, besides Health Services and Counselling Services, Kathy Somers runs training workshops and one-on-one sessions related to stress management and high performance situations. <http://www.uoguelph.ca/~ksomers/>

If you have a documented disability or think you may have a disability:

Student Accessibility Services (formerly Centre for Students with Disabilities) can provide services and support for students with a documented learning or physical

disability. They can also provide information about how to be tested for a learning disability. For more information, including how to register with the centre please see: