1 Course Details

1.1 Calendar Description

This course will familiarize students with tools for the computational acquisition and analysis of molecular biological data. Key software for gene expression analyses, biological sequence analysis, and data acquisition and management will be presented. Laboratory exercises will guide students through application of relevant tools.

Restrictions: Restricted to students in Bioinformatics programs. Students in other programs may consult with course instructor.

1.2 Course Description

Overview: The main goal of this course is to guide graduate students through an introduction to the analysis of biological data using computational and statistical tools, with emphasis upon the analysis of molecular biology data. The course will largely focus upon developing programming skills in the R language for quality checking, analyzing, and visualizing data. The course also includes an introduction to several key web-based tools, SQL, and command-line software tools for phylogenetics. We also will cover how to acquire and analyze data from selected biological databases important for bioinformatics, including sequence databases, biodiversity databases, and functional gene annotation resources. It is important to recognize the origins and limitations of these data in addition to their utility. We will discuss core bioinformatics algorithms (e.g. for alignment, clustering, phylogenetics) and population genetics metrics and principles that are important for making analytical decisions and interpreting results. We also will promote good practices for organizing your data and analyses, prepare reproducible analyses and well-commented code, and introduce software tools that facilitate version control and collaborative coding. As bioinformatics is a fast-moving discipline, we will also spend time practising strategies for how to learn to use to use new tools and to conduct new analyses.
**Curriculum Note:** This course is complementary to others in the bioinformatics graduate program. In the fall semester of 2019, programming in the Unix environment and in the Python language are covered in Bioinformatics Programming (BINF*6410). Students in the Master of Bioinformatics program must also take Topics in Bioinformatics (BINF*6890), which covers diverse concepts in bioinformatics and emphasizes critical thinking and communication skills. The winter semester core bioinformatics courses are Genomic Methods (BINF*6110), in which large-scale genomic analysis is covered, and Statistical Bioinformatics (BINF*6970). Students from other graduate programs may wish to discuss their background and the suitability of these courses with the instructors prior to enrolling.

**Pre-Requisites:** Students accepted into the Master of Bioinformatics and MSc in Bioinformatics programs should have the necessary background for this course. No prior programming experience is assumed. Students are expected to have taken at least one course at the undergraduate level in each of genetics and statistics.

1.3 **Timetable**

Where: SSC 1306

When: Tuesdays and Thursdays 11:30 AM - 12:50 PM. September 5 - November 28.

(Note: There is no class on Tuesday October 15 for the Fall Study Break.)

1.4 **Final Exam**

There is no final exam for this course.

---

2 **Instructional Support**

2.1 **Instructional Support Team**

**Instructor:** Sarah Adamowicz Associate Professor, Department of Integrative Biology

**Email:** sadamowi@uoguelph.ca

**Telephone:** +1-519-824-4120 x53055

**Office:** SSC 2447

**Office Hours:** Tuesdays 1:30-2:30 PM from September 10 - December 3, 2019.

Location: SSC (Science Complex) 2447.
2.2 Teaching Assistants

Teaching Assistant: Jacqueline May PhD in Bioinformatics Candidate  
Email: mayj@uoguelph.ca  
Office: SSC2454  
Office Hours: Jacqueline May is available for consultations with students. Please email Jacqueline for an appointment.

3 Learning Resources

3.1 Required Resources

R and RStudio (Software)
Classes will generally start with a short background lecture, interspersed with discussions and learning activities, followed by a hands-on computer tutorial. Please bring your own laptop to every class, and install needed software in advance of class.

Prior to the first class, please install R on your computer:
https://www.r-project.org/

Prior to the second class, please install RStudio:
https://www.rstudio.com/

Announcements will be made throughout the semester regarding R packages or additional software to install prior to the next class.

Papers and Textbook Chapters (Readings)
Relevant published articles related to the course content for each day will be posted through CourseLink. The first-listed article for each class is required reading for that class. The other posted readings are recommended or supplemental for those interested in more depth on that topic.

We will also be consulting a wide range of online resources, such as software manuals and vignettes for Bioconductor packages. Links to relevant resources will be posted in the class slides and in the comments sections of the example code.

Additionally, selected chapters from the following manuals and books will be recommended to accompany specific modules. All are available as open-access PDFs directly online or are available as a PDF book for download through the University of Guelph library site (https://www.lib.uoguelph.ca/).

1. Paradis E, 2005. R For Beginners. (Freely available through the following link: https://cran.r-project.org/doc/contrib/Paradis-rdebuts_en.pdf)
4 Learning Outcomes

4.1 Course Learning Outcomes

By the end of this course, you should be able to:

1. obtain data from key databases relevant for bioinformatics and to understand the sources and limitations of these data.

2. filter, manipulate, analyze, and visualize bioinformatics data (with non-exclusive emphasis on the R programming language and software resources available through Bioconductor).

3. conduct reproducible analyses and use software tools for version control and collaboration.

4. understand and apply selected algorithms commonly used in bioinformatics, including for sequence alignment and clustering.

5. adapt the above skills to learn new tools and conduct new analyses not explicitly covered in class.

5 Teaching and Learning Activities

5.1 Seminar

Thu, Sep 5 - Thu, Nov 28

Topics: Classes are organized to contain a mixture of
activities. Typically, classes will start with a background lecture of 15-40 minutes, often interspersed with discussion or conceptual activities, followed by a guided hands-on computer tutorial.

6 Assessments

In this course, assignments are used not only for assessment but also as learning tools. You should work on your assignments regularly. Do not leave these assignments to the night before!

Detailed instructions for each assignment will be posted to CourseLink at least two weeks before the due date and at least three weeks in the case of the Final Project. An overview is provided below.

6.1 Assessment Details

Assignment #1 (20%)

Date: Fri, Oct 4

Assignments #1 and #2 follow a similar format. Each consists of three components. In Part A (background), you will be asked to define key terms and provide short answers to demonstrate your understanding of important concepts, algorithms, and functions. In Part B (coding skills), you will need to write code that performs specific data analysis steps. The Part B tasks will involve writing code similar to examples we cover in class. For Part C (application of knowledge to solve new problems), you will design and complete a micro-project that builds upon skills and concepts covered until that point in the course.

You should focus on being precise in your explanations of algorithms and functions. Code needs to be correct and do what it is meant to do (always check). In part C, you will additionally be assessed on creativity and novelty of your micro-project in terms of going beyond the class materials. Your code should be correct, well-commented, and reproducible. You will also need to provide a short written summary interpreting your findings.

Assignment #1 is due by 5:00 PM on Friday Oct. 4th.

Assignment #2 (20%)

Date: Fri, Oct 25

See above for the description of Assignment #1.

Additionally, as the course progresses, you should aim to write code that is elegant (i.e. streamlined and not excessively convoluted for a given task) as well as computationally efficient. You should also pay careful attention to the preparation of your visualizations, considering whether the main message is conveyed, informative labeling, colour/symbol
choice, etc.

**Assignment #2** is due by 5:00 PM on Friday October 25th.

**Assignment #3 (Group Project) (15%)**
For Assignment #3 (Group Project), you will swap code (your Part C from either Assignment #1 or Assignment #2) with a peer in your group. The assignment involves making improvements to your peer's code and using GitHub to manage the collaboration and code edits. You should discuss the project together and may work on the code together. Each person will individually prepare a short-write up about the code improvements and collaboration process, which is individually graded.

(Why GitHub? GitHub is an important code repository as well as a tool for version control and collaboration. By the end of your program, we would highly recommended that you post examples of your work to GitHub and provide a link to your GitHub page on your CV when applying for bioinformatics-related jobs.)

**Assignment #3 is due by 5:00 PM on Friday November 15th.**

**Assignment #4 (Seminars) (5%)**
For students in the Software Tools class, attendance at the Bioinformatics Seminar Series is mandatory. The series will help you to expand your knowledge of the field of bioinformatics.

Students should attend all three seminars of the F19 semester and choose two for this short writing assignment about the seminars.

Seminars are held monthly from 2:30-3:30 PM on Tuesdays. In the F19 semester, the dates are: Sept. 24, Oct. 22, and Nov. 19. This time is scheduled around the BINF courses. In the case of any students in Software Tools from other graduate programs who have a course conflict, you must notify the instructor of the conflict within the first two weeks of class, and an adjusted small writing assignment will be assigned to you of the same length, scope, and value (5%).

**Assignment #4 is due by 5:00 PM on Friday November 29th.**

**Assignment #5 (Final Project) (40%)**
**Date:** Tue, Dec 10
Assignment #5 involves completing a final course project consisting of written paragraphs interspersed with commented code blocks and visualizations. Your project should include: introduction, description of dataset, data exploration and quality control (commented code block and visualizations), description of main software tool used, main analyses (commented code block and visualizations), interpretation of results and discussion.

You will choose your own topic, but you should clear your topic selection with the instructor at least two weeks before the due date. Your project must incorporate at least one
software tool beyond those covered in class. Being able to read software documentation and do new analyses of interest to you is important in bioinformatics.

Detailed instructions for each assignment will be posted to CourseLink.

**Assignment #5 is due by 5:00 PM on Tuesday December 10th.**

---

### 7 Course Statements

#### 7.1 Class Attendance

Class attendance is mandatory for Software Tools. This is not a Distance Education course, and the class structure is organized around participation during class. It is hard to replace the learning activities exclusively through study on your own. So, please do plan to attend every class to maximize your success.

However, if you do need to miss a class (e.g. due to illness or compassionate reasons), it is important to check the slides and the example code. Slides, code examples, and coding activities/challenges will be posted to CourseLink for each class.

Throughout the semester, you should regularly consult CourseLink for announcements and posted files.

#### 7.2 Group Activities

Throughout the course, we will engage in regular discussions and coding activities in pairs or small groups during class time. Pedagogical research indicates that you will learn better if you regularly work in groups and engage in active learning activities. So, please come to class prepared to engage with your peers.

We will change up the groups from time to time so that you can meet new people and work with individuals with varying personalities and academic backgrounds. Collaboration is common in bioinformatics in the workplace as well, and so this is good practice for your career beyond graduate studies.

I also encourage students to form peer study groups to review course materials outside class time and to engage in other activities beyond the course materials, such as analyzing additional datasets to develop your skills further. You may also work together to solve the “coding challenges” posed to you at the end of some classes. Taking the time to work through problems of increasing difficulty will help you.

For graded assignments, it is important to complete your work on your own. The exception is the single group assignment. For that assignment, you should discuss your assignment and can work together on the code. Each person submits an individual short write-up, which is graded individually. You need to prepare your own written remarks.
7.3 Course-Specific Statement on Academic Integrity

You are encouraged to work in peer groups to practise your coding skills, to discuss concepts, and to seek advice about useful software and information resources. However, you must complete your individual assignments independently. Electronic resources (such as TurnItIn) may be used to assess the originality of your assignments. Use quotations sparingly, such as for profound statements or definitions. Otherwise, you should paraphrase from any sources you cite for the written portions of your assignments. You are free to consult online resources to learn about various ways of coding and approaching bioinformatics problems. If you draw heavily from a specific source (such as a particular entry on Stack Overflow) when completing an assignment, then you should cite that source and indicate how you adapted the code for your purposes. You always need to check that your code works as intended.

You will work together in a small group for one group assignment (assignment #3). You should discuss your assignment and may work on the code together. You should complete the short write-up for that assignment on your own. The assignment is graded individually.

Please see below for the university-level statement on academic integrity for further information.

7.4 Due Dates

Assignments are due by the due DATE AND TIME listed for each assignment.

Late work that is submitted within one hour (exactly 60 minutes) of the deadline will receive a penalty of 5%.

Otherwise, late work will receive a deduction of 10% per 24-hour period.

See below for the university guidelines regarding cases of missed work due to illness or compassionate reasons.
Assignments should be submitted to the course Dropbox available through CourseLink.

8 College of Biological Science Statements

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

- Make an appointment with a program counsellor in your degree program. B.Sc. Academic Advising or Program Counsellors

8.2 Academic Support
If you are struggling to succeed academically:

- Learning Commons: 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.
- Science Commons: Located in the library, the Science Commons provides support for physics, mathematic/statistics, and chemistry. Details on their hours of operations can be found at: Chemistry & Physics Help and Math & Stats Help

8.3 Wellness
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.
- Student Health Services is located on campus and is available to provide medical attention.
- 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.selfregulationskills.ca/

9 University Statements
9.1 Email Communication

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

9.2 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 instructor (or designated person, such as a teaching assistant) in writing, with your name, id#, and e-mail contact. The grounds for Academic Consideration are detailed in the Undergraduate and Graduate Calendars.

Undergraduate Calendar - Academic Consideration and Appeals
https://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-ac.shtml

Graduate Calendar - Grounds for Academic Consideration
https://www.uoguelph.ca/registrar/calendars/graduate/current/genreg/index.shtml

Associate Diploma Calendar - Academic Consideration, Appeals and Petitions
https://www.uoguelph.ca/registrar/calendars/diploma/current/index.shtml

9.3 Drop Date

Students will have until the last day of classes to drop courses without academic penalty. The deadline to drop two-semester courses will be the last day of classes in the second semester. This applies to all students (undergraduate, graduate and diploma) except for Doctor of Veterinary Medicine and Associate Diploma in Veterinary Technology (conventional and alternative delivery) students. The regulations and procedures for course registration are available in their respective Academic Calendars.

Undergraduate Calendar - Dropping Courses
https://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-drop.shtml

Graduate Calendar - Registration Changes
https://www.uoguelph.ca/registrar/calendars/graduate/current/genreg/genreg-regregchg.shtml

Associate Diploma Calendar - Dropping Courses
https://www.uoguelph.ca/registrar/calendars/diploma/current/c08/c08-drop.shtml

9.4 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.

9.5 Accessibility

The University promotes the full participation of students who experience disabilities in their academic programs. To that end, the provision of academic accommodation is a shared
responsibility between the University and the student.

When accommodations are needed, the student is required to first register with Student Accessibility Services (SAS). Documentation to substantiate the existence of a disability is required; however, interim accommodations may be possible while that process is underway.

Accommodations are available for both permanent and temporary disabilities. It should be noted that common illnesses such as a cold or the flu do not constitute a disability.

Use of the SAS Exam Centre requires students to book their exams at least 7 days in advance and not later than the 40th Class Day.

For Guelph students, information can be found on the SAS website https://www.uoguelph.ca/sas

For Ridgetown students, information can be found on the Ridgetown SAS website https://www.ridgetownc.com/services/accessibilityservices.cfm

9.6 Academic Integrity

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 encourages academic integrity. 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.

Undergraduate Calendar - Academic Misconduct
https://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-amisconduct.shtml

Graduate Calendar - Academic Misconduct
https://www.uoguelph.ca/registrar/calendars/graduate/current/genreg/index.shtml

9.7 Recording of Materials

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

The Academic Calendars are the source of information about the University of Guelph’s procedures, policies, and regulations that apply to undergraduate, graduate, and diploma programs.

Academic Calendars
https://www.uoguelph.ca/academics/calendars