MCB 416A/516A

**Bioinformatics and Functional Genomics Analysis**

Spring 2023

**Instructor:**

Guang Yao, Ph.D.

Associate Professor, Molecular and Cellular Biology

guangyao@arizona.edu

Office Hours: Wednesdays 11:00 am-1:00 pm

**Teaching Assistant:**

Nick Bielski

bielski@email.arizona.edu  
Office Hours: Tuesdays 2:00-3:00 pm; Fridays 9:30-10:30 am

**Course description:**

The course introduces bioinformatics and statistical methods for the analysis of large-scale experimental data in functional genomics. It uses the analysis of next-generation RNA-sequencing data as a leading example and prepares the students to perform independent analyses of genomic data in an interdisciplinary environment such as a research lab or pharmaceutical company.

**Learning outcomes:**

After completing the course, the students (in 416A and 516A) will be able to

* understand and describe basic biological concepts and techniques in functional genomics;
* understand and apply basic bioinformatics and statistical methods in analyzing functional genomics data;
* perform large-scale data analysis using the high-performance computing (HPC) platform and R/Bioconductor software suite.

Additionally, graduate students (in 516A) will be able to

* apply concepts and methods learned in this course to appreciate current research topics and big data analysis in functional genomics or related –omics fields;
* develop and present a mini-lecture to demonstrate their learned knowledge in a chosen research topic.

**Primary audience:**

Graduate students and senior undergraduate students in life sciences and related fields (e.g., computer sciences, Math/Stat, and engineering) who have basic quantitative training and are interested in learning cutting-edge methods, and understanding the underlying statistical principles, in analyzing large-scale functional genomics data in biomedical research.

**Prerequisite:**

* Ability of basic programming, with CSC 250 (Essential Computing for the Sciences) or equivalent strongly recommended
* Understanding of basic statistics, with Math/Data 363 (Introduction to Statistical Methods) or equivalent strongly recommended
* Understanding of molecular genetics, with MCB 301 (Molecular Basis of Life) or equivalent strongly recommended

**Textbook and course materials:**

There is no required textbook for the course.

Lecture notes, handouts, weekly calendar, and other course information and materials are provided on the course website in D2L.

**Topics:**

*Module 1: Introduction to Next-generation sequencing (NGS) and RNA-seq techniques*

* NGS and RNA-seq procedure and platforms
* RNA-seq applications

*Module 2: Introduction to* *high performance computing (HPC) and R/Bioconductor*

* HPC overview and access
* UNIX Command line script and HPC scheduling/running script
* Overview of R and Bioconductor packages
* Frequently used R commands
* Data structures and visualization in R

*Module 3: Transcriptome (RNA-seq) data analysis*

* Read quality control and preprocessing
* Mapping and alignment
* Transcriptome assembly
* Quantification of gene expression
* Differential expression analysis

*Module 4:* *Bioinformatics analysis*

* Functional genomics databases
* Statistical testing and machine learning in data analysis
* Gene ontology and gene set analysis
* Pathway and network analysis

*Module 5: Student group topics and project presentations*

**Equipment requirements:**

For this class you will need access to a computer with internet connection.

**Course requirements/assignments:**

* **Homework (45 pts):** There will be three homework assignments (15 pts each), as indicated on the D2L course calendar.

There will be a 10% penalty each day for late homework submission.

* **Quizzes and Exams (75 pts)**:
  + Three *scheduled* in-class quizzes (15 pts each), as indicated on the D2L course calendar. Quizzes are closed-book (no notes or other resources may be used) unless otherwise noted.
  + Eighteen *unscheduled* short quizzes will be given in class, with questions based on most recent lecture materials. Three lowest short quiz scores will be dropped for each student.
  + No midterm and final exams.
  + Total points for scheduled quizzes = 45 pts.
  + Total points for unscheduled short quizzes = 30 pts.
* **Final project (45 pts):** There will be one final project in which students select a research article of interest, download its raw sequencing data, and reproduce its RNA-seq data analysis. The students will also perform at least one functional analysis (e.g., pathway or gene set analysis) using a different method(s) from the authors’ and discuss the obtained result in the context of the paper.

The project will be done in groups of up to 3 students each. Each groupwill write a final report (including source code) and give an oral presentation.

Grading for the final project will be based on the quality of a) the written report (50%) and b) the presentation (50%) including answering questions.

* **Mini-lecture presentation (516A only, 35 pts):** Graduate students will form groups (2-3 students each) to work with the instructor and identify functional genomics-related topics. Each group will study the chosen topic, develop and present a lecture (discussing techniques, applications, analysis principles, case studies etc.), including class exercises.

Grading for group topic presentation will be based on the organized and presented materials (50%, in the format of slides and slide descriptions), as well as the quality of the presentation, discussion, and exercise (50%).

**Make-up Assignments:** Make-ups will not be given for missed assignments. In cases of *emergencies*, contact the instructor within *48 hours* of the assignment due time to discuss special arrangements if needed.

**Final grades**:

* The total points for undergraduate students (in 416A) will be 165 points.
* The total points for graduate students (in 516A) will be 200 points.
* The final grades will be given following the scale (percentage of the total points) below:

90-100 % = A

80-89% = B

70-79% = C

60-69% = D

0-59% = E

**Absence and Class Participation Policy:** The UA’s policy concerning Class Attendance, Participation, and Administrative Drops is available at <http://catalog.arizona.edu/policy/class-attendance-participation-and-administrative-drop>

The UA policy regarding absences for any sincerely held religious belief, observance or practice will be accommodated where reasonable. See <http://policy.arizona.edu/human-resources/religious-accommodation-policy>.

Absences pre-approved by the UA Dean of Students (or Dean Designee) will be honored. See <https://deanofstudents.arizona.edu/absences>

**Code of Academic Integrity:** Students are encouraged to share intellectual views and freely discuss the principles and applications of course materials. Students may work together in discussion and consultation on course assignments, but *all materials submitted to the course instructor must be the exclusive written work of the student indicated.* *Students may neither collaborate or consult with other students* *nor use other resources during quizzes.*

*Any form of cheating or plagiarism will result in a failing grade for the course, and in other appropriate disciplinary actions,*as described by the UA Code of Academic Integrity at <http://deanofstudents.arizona.edu/academic-integrity/students/academic-integrity>

The University Libraries have some excellent tips for avoiding plagiarism, available at <https://new.library.arizona.edu/research/citing/plagiarism>

**Class Recordings and Notes:** For lecture recordings, which are used at the discretion of the instructor, students must access content in D2L only. Students may not modify content or re-use content for any purpose other than personal educational reasons. All recordings are subject to government and university regulations. Therefore, students accessing unauthorized recordings or using them in a manner inconsistent with UArizona values and educational policies are subject to suspension or civil action.

*Selling class notes and/or other course materials to other students or to a third party for resale is NOT permitted without the instructor’s written consent*. Violations to this and other course rules are subject to the Code of Academic Integrity and may result in course sanctions. Additionally, students who use D2L or UA e-mail to sell or buy these copyrighted materials are subject to Code of Conduct Violations for misuse of student e-mail addresses. This conduct may also constitute copyright infringement.

**Special Needs and Accommodations:** Students who need special accommodations or services should contact the Strategic Alternatives Learning Techniques (SALT) Center (<http://www.salt.arizona.edu>) and/or the Disability Resources Center (<http://drc.arizona.edu>, 520-621-3268). The appropriate office must document the need for accommodations.

**Classroom Behavior Policy:** Disruptive behaviors that interfere with or obstruct class teaching/learning activities are prohibited. These behaviors include disruptive conversations, non-class related readings, etc. Also included is the non-class related use of laptops, tablets, mobile devices, and other technology. In the event of disruptive behaviors in the classroom, official policies and procedures will be followed as described at <https://policy.arizona.edu/education-and-student-affairs/disruptive-behavior-instructional-setting>

**Policy Against Threatening Behavior:** <https://policy.arizona.edu/education-and-student-affairs/threatening-behavior-students>

**Additional Resources for Students:**

UA Academic policies and procedures are available at <http://catalog.arizona.edu/policies>

Student Assistance and Advocacy information is available at <http://deanofstudents.arizona.edu/student-assistance/students/student-assistance>

***The information contained in the course syllabus, other than the grade and absence policies, may be subject to change with reasonable advance notice, as deemed appropriate by the instructor.***