



DEPARTMENT OF BIOCHEMISTRY  
AND  
PURDUE UNIVERSITY LIBRARIES

**BCHM 695 Syllabus**  
**Introduction to R and Bioconductor**  
**Summer 2017**

**Dr. Pete E. Pascuzzi, Instructor**

Office: LILY L-31 (in the Life Sciences Library)  
WALC later in summer.  
Phone: 494-3620  
e-mail: [ppascuzz@purdue.edu](mailto:ppascuzz@purdue.edu)  
Office hours: between lecture and lab, after lab or by appointment

**Siwen Wang, Teaching Assistant**

Office: BCHM A343  
Phone: 494-7672  
email: [wang1266@purdue.edu](mailto:wang1266@purdue.edu)  
Office hours: TBD

**Candy Mao, Teaching Assistant**

Office: BCHM 211  
Phone: 494-0657  
email: [mao47@purdue.edu](mailto:mao47@purdue.edu)  
Office hours: TBD

**COURSE DESCRIPTION**

This course will provide an introductory, hands-on experience for life science researchers in bioinformatics using R and Bioconductor. Emphasis will be placed on accessing, formatting, and visualizing genomics data. Most analyses will deal with “little” data (no mapping or assembly of short reads), but some techniques to work with “big” data (e.g. BAM files) will be covered. Lecture and lab will both be held in a computer lab, so lecture will be “hands-on”. Working in small groups is encouraged.

**COURSE OBJECTIVES**

- ✓ Students will learn the fundamentals of bioinformatics analyses of genomics data using R and Bioconductor.
- ✓ Students will gain a greater appreciation for bioinformatics and the parallels with “wet bench” experiments.
- ✓ Students will be introduced to the concept of “literate programming” and how it can be applied to document their work and write legible reports.
- ✓ Students will be prepared for more advanced courses in R or bioinformatics, or for continued self-learning.

## LEARNING OUTCOMES

- ✓ Students will be able to write R scripts that utilize many of functions and packages that R and Bioconductor provides for bioinformatics.
- ✓ Students will be able to access genome-scale data sets at public repositories from within the R environment.
- ✓ Students will be able to visualize genome-scale data sets with R.
- ✓ Students will learn common techniques to analyze genome-scale data with R.
- ✓ Students will learn techniques to work with “big” data in R.

## TEXTBOOK

None. However, there are many eBooks on R programming available through the library. Here are a few that are recommended:

**The Art of R Programming**, Norman Matloff  
**R in a Nutshell**, Joseph Adler  
**Introductory Statistics with R**, Peter Daalgard

Another recommended resource on statistics for biologists is on reserve at the library:

**Statistics at the Bench**, M. Bremer and R.W. Doerge

## LECTURE TIME AND PLACE

Lecture	TR	8:40 – 9:30	LILY G428
Lab	TR	9:50 – 11:40	LILY G428

## ASSESSMENT

This course is being offered for a letter grade. Grades will be determined through successful completion of homework, exercises, projects, and class participation. Final grades will be based on the following:

Four homeworks	10%
Four in-class exercises	10%
Four programming projects	30%
Final exam	50%
Class participation and attendance	0, 5 or 10 bonus points

Homeworks and exercises will be graded on a scale 0 - 10 with point values clearly indicated. Every effort will be made to give partial credit, but, remember, computers take you very literally! Programming projects are not graded on style or efficiency, only on successful completion of the assigned task(s). The final which will be graded as a traditional exam. Perfect attendance is required for 10 bonus points, and one absence is allowed for 5 bonus points. Pre-arranged, excused absences may be allowed at the discretion of the instructor. In either case, active participation in class discussions is expected.

## GRADING SCALE

98.00 and above	A+	78.00 to 79.99	C+
92.00 to 97.99	A	72.00 to 77.99	C
90.00 to 91.99	A-	70.00 to 71.99	C-
88.00 to 89.99	B+	68.00 to 69.99	D+
82.00 to 87.99	B	62.00 to 67.99	D
80.00 to 81.99	B-	60.00 to 61.99	D-
		59.99 and Below	F

## **EXTRA CREDIT**

There will be no opportunity for extra credit other than class attendance.

## **COURSE MANAGEMENT SYSTEM**

We will not use Blackboard for this course. Instead, we will use the Purdue University Research Repository (PURR). Each student must create an account with PURR, and create a private PURR project with Dr. Pascuzzi and the teaching assistant as collaborators. This project site will be used to turn in all completed assignments and to track your progress in the class. There will be a common PURR project for the course where students will retrieve the lectures, data files and other material as required.

## **OBTAINING EXTRA HELP**

Dr. Pascuzzi will be available to answer your questions immediately after class, during office hours, or by appointment (arranged in class or by e-mail). Alternatively, you can submit questions by e-mail that can be answered in class or by return e-mail. The teaching assistants will be available during office hours, by email or by appointment. If you are struggling, get help!

## **ACADEMIC MISCONDUCT**

Academic misconduct of any kind will not be tolerated in any course offered by the Department of Biochemistry. Information on Purdue's policies with regard to academic misconduct can be found at

[http://www.purdue.edu/studentregulations/student\\_conduct/regulations.html](http://www.purdue.edu/studentregulations/student_conduct/regulations.html)

Any incidence of academic misconduct will be reported to the Office of the Dean of Students. Academic misconduct may result in disciplinary sanctions including expulsion, suspension, probated suspension, disciplinary probation, and/or educational sanctions. In addition, such misconduct will result in punitive grading such as:

- receiving a lower or failing grade on the assignment, or
- assessing a lower or failing grade for the course

Punitive grading decisions will be made after consultation with the Office of the Dean of Students. Please note reported incidences of academic misconduct go on record for reference by other instructors. Further, a record of academic misconduct is likely to influence how current/future situations are handled.

To provide you with an unambiguous definition of academic misconduct, the following text has been excerpted from "Academic Integrity: A Guide for Students", written by Stephen Akers, Ph.D., Executive Associate Dean of Students (1995, Revised 1999, 2003), and published by the Office of the Dean of Students in cooperation with Purdue Student Government, Schleman Hall of Student Services, Room 207, 475 Stadium Mall Drive West Lafayette, IN 47907-2050.

"Purdue prohibits "dishonesty in connection with any University activity. Cheating, plagiarism, or knowingly furnishing false information to the University are examples of dishonesty." [Part 5, Section III-B-2-a, [Student Regulations](#)] Furthermore, the University Senate has stipulated that "the commitment of acts of cheating, lying, and deceit in any of their diverse forms (such as the use of substitutes for taking examinations, the use of illegal cribs, plagiarism, and copying during examinations) is dishonest and must not be tolerated. Moreover, knowingly to aid and abet, directly or indirectly, other parties in committing dishonest acts is in itself dishonest." [University Senate Document 72-18, December 15, 1972]

More specifically, the following are a few examples of academic dishonesty which have been discovered at Purdue University.

- substituting on an exam for another student
- substituting in a course for another student
- paying someone else to write a paper and submitting it as one's own work
- giving or receiving answers by use of signals during an exam
- copying with or without the other person's knowledge during an exam
- doing class assignments for someone else
- plagiarizing published material, class assignments, or lab reports
- turning in a paper that has been purchased from a commercial research firm or obtained from the internet
- padding items of a bibliography
- obtaining an unauthorized copy of a test in advance of its scheduled administration
- using unauthorized notes during an exam
- collaborating with other students on assignments when it is not allowed
- obtaining a test from the exam site, completing and submitting it later
- altering answers on a scored test and submitting it for a regrade
- accessing and altering grade records
- stealing class assignments from other students and submitting them as one's own
- fabricating data
- destroying or stealing the work of other students

Plagiarism is a special kind of academic dishonesty in which one person steals another person's ideas or words and falsely presents them as the plagiarist's own product. This is most likely to occur in the following ways:

- using the exact language of someone else without the use of quotation marks and without giving proper credit to the author
- presenting the sequence of ideas or arranging the material of someone else even though such is expressed in one's own words, without giving appropriate acknowledgment
- submitting a document written by someone else but representing it as one's own"

## **CLASS ATTENDANCE**

In accordance with University policy, you are expected to attend every scheduled class. If you have a valid reason for missing class such as a University-sponsored activity, religious observances, illness, or family emergency, the instructor or TA will assist you in obtaining information and materials you may have missed. Students who skip class without a valid excuse should not expect the instructor or TA to supply class notes or provide special help. For the official university policy, see:

[www.purdue.edu/odos/services/classabsence.php](http://www.purdue.edu/odos/services/classabsence.php) and  
[http://www.purdue.edu/studentregulations/regulations\\_procedures/classes.html](http://www.purdue.edu/studentregulations/regulations_procedures/classes.html)

## **EMERGENCY PREPAREDNESS**

In the event of a major campus emergency, course requirements, deadlines and grading percentages are subject to changes that may be necessitated by a revised semester calendar or other circumstances. To get information about changes in this course consult the class Blackboard site or e-mail or phone the instructor.

## **ON-LINE COURSE EVALUATIONS**

During the last two weeks of the semester, you will be provided an opportunity to evaluate this course and your instructor(s). To this end, Purdue has transitioned to online course evaluations. On Monday of the fifteenth week of classes, you will receive an official email from evaluation administrators with a link to the online evaluation site. You will have two weeks to complete this evaluation. Your participation in this evaluation is an integral part of this course. Your feedback is vital to improving education at Purdue University. I strongly urge you to participate in the evaluation system.

## **NON-DISCRIMINATION POLICY STATEMENT**

Purdue University's non-discrimination policy will be upheld in this classroom. Purdue University is committed to maintaining a community which recognizes and values the inherent worth and dignity of every person; fosters tolerance, sensitivity, understanding, and mutual respect among its members; and encourages each individual to strive to reach his or her own potential. In pursuit of its goal of academic excellence, the University seeks to develop and nurture diversity. The University believes that diversity among its many members strengthens the institution, stimulates creativity, promotes the exchange of ideas, and enriches campus life.

Purdue University views, evaluates, and treats all persons in any University related activity or circumstance in which they may be involved, solely as individuals on the basis of their own personal abilities, qualifications, and other relevant characteristics.

For more information, see [http://www.purdue.edu/purdue/ea\\_eou\\_statement.html](http://www.purdue.edu/purdue/ea_eou_statement.html).

## **LECTURE / LAB SCHEDULE (subject to minor changes)**

<b>Week 1</b>	<b>May 16 and 18</b>	<b>R - Setup and Basics</b>
<b>Week 2</b>	<b>May 23 and 25</b>	<b>R - Data Import, Manipulation and Export</b>
<b>Week 3</b>	<b>May 30 and June 1</b>	<b>Bioconductor - Ranged Data, Data Visualization</b>
<b>Week 4</b>	<b>June 6 and 8</b>	<b>Bioconductor - Biological Sequences</b>
<b>Week 5</b>	<b>June 13 and 15</b>	<b>Bioconductor - Microarrays and Gene Expression</b>
<b>Week 6</b>	<b>June 20 and 22</b>	<b>Bioconductor - Ontology and Pathway Analysis</b>
<b>Week 7</b>	<b>June 27 and 29</b>	<b>Bioconductor - RNAseq Analysis</b>
<b>Week 8</b>	<b>July 6</b>	<b>Final</b>