

MICR4203/5203 – Principals of Bioinformatics

Instructor: Robert L. Burnap, Department of Microbiology & Molecular Genetics

Email: rob.burnap@okstate.edu (allow 1 business day for response); Please use office hours for routine discussion of the course material and mechanics.

Office: 130P Henry Bellmon Research Center,

Office Hours: in person, by appointment.

Course Website: Online Classroom (<https://canvas.okstate.edu/courses/43714/>)

Lectures: Tuesdays –Thursday 9:00-10:20 am 113 Life Sciences East

Required Textbook: *Concepts in Bioinformatics and Genomics* 2016 First Edition Jamil Momand, Alison McCurdy, Contributions by Silvia Heubach, and Nancy Warter-Perez ISBN: 9780199936991

Course Description: The course provides life scientists and pre-professionals a foundation in the concepts of biological sequence information and the inferential techniques used to assign structure, function, and evolutionary relationships among genes and proteins. It focuses on the theory and application of available software used for biological sequence, structural, and phylogenetic analyses and introduces the principals of comparative genomics. We also introduce current approaches to assigning functions to genes and gene networks and how characteristics such as metabolic capacities are inferred from genomic analysis of predicted enzymes. It is inherently multidisciplinary and involves, most notably, molecular biology and computer science.

The *first part* of the course emphasizes the mechanics of molecular evolution and how bioinformatic tools, notably scoring matrices, are designed to simulate these processes in a way that allows predictive powers for things such as sequence alignments.

The *second part* of the course extends these concepts to the characterization of gene families and applied lessons on the rudiments of inferring phylogeny.

The *third part* of the course introduces genome-wide analyses including comparative genome structure, use of clustering algorithms for analyzing global gene expression and inferring metabolic capacity through consideration of gene content.

Although computers and biology have been united for decades now, the impetus for bioinformatics derives from the need to make biological sense out of the explosion of genetic sequence data that floods in from modern high-throughput sequence projects. This course is oriented to the intelligent understanding and application of computers to the problems of modern molecular biology. Although the course does not require extensive computer programming, it covers basic algorithms used, for example, for analyzing sequence similarity focusing on the development of scoring matrices and introduces file manipulation using exercises in the Python language. This course is the foundation course for the OSU Bioinformatics Certificate Program.

Prerequisites: Cell & Molecular Biology MICR3033 or Biochemistry BIOC3653 or equivalent. An excellent understanding of the fundamentals of molecular genetics (i.e. DNA->RNA->protein) and the basics of protein structure is assumed. A good familiarity with the operation of a personal computer is also assumed.

Graded Material:

Homework assignments= Assigned Project Tasks: A major component of the class is to perform a set of homework tasks relating to the course. Reading assignments from your textbook and primary literature will be given in class. Since the text does not cover all of the topics covered in the course, additional readings in the form of handouts or library materials will be required. *Homework will often be given as an email message that you will receive before 5 pm the day before class.* **Pay attention to your Okstate email.**

Quizzes: Quizzes may be unannounced and will typically cover material covered in reading assignments or to assess homework. For the latter, homework will be maintained in your digital notebook, but the actual assessment on whether comprehension was obtained will be made using a quiz covering recent homework.

Exams: Three exams (2 midterm 1 comprehensive final exam) will be given. Each exam will consist of a combination of short answer, computation, and short essay type questions. Questions will be based upon material from lectures, discussion, and reading assignments. Questions based upon material in reading assignments, but not lectured on, may appear on exams. No make-up exams will be given - if you have an excused absence you must make arrangements prior to the scheduled exam time. Excused exam absences are granted only for truly emergency conditions. Final Exam is scheduled for Thursday, December 12 8:00-9:50am. **Note: arrangements for final examinations at earlier dates will not be made.**

Graduate Credit:

Graduate students are highly encouraged to take the class, but will be required to complete a project that employs the basic principals of the course. Students must consult with the instructor to define the project, which, minimally, will be comprised of a well-researched statement of a bioinformatics problem and a documentation of a set of completed tasks aimed at its solution. Typically, this will involve a project on the student's thesis topic. The final project will be given in digital notebook form according to the initial specifications, which typically follow the tasks for the assigned gene sequence. The graduate project will be worth an additional 100 points (600 points total for the course).

Graded Points:

Exams (3) at 90, 110 and 150pts	350 pts
Quizzes/Discussion (both online and in-class)	50 pts
Assigned Project Tasks (homework)	100 pts
Total Points for Course	500 pts

Grades for tests and submitted items will be posted at the D2L site within one week. Every effort will be made by the instructor to ensure the latest grade information is available prior to key course dates (exams, 'W' withdrawal dates).

Letter Grade Assignments:

450-500 pts	= A
400-449 pts	= B
350-399 pts	= C
300-349 pts	= D
Below 300 pts	= F

Late Policy for Assignments: Late assignments will have 10% possible points on the first day, and an additional 10% deducted for each of the subsequent days the assignment is late.

Drop Policy

Information about university drop policy and dates is at this website:

<http://registrar.okstate.edu/>

Click on "class schedules," and "short, internet, and outreach courses"

To drop this course, contact the Registrar's office, (405) 744-6876, or drop through SIS (Student Information Services).

Academic Honesty: Cheating on exams or any form of academic dishonesty will not be tolerated. Flagrant violations may result in an automatic "F" for the entire course. It is our hope and your responsibility that this situation does not arise. Participating in a behavior that violates academic integrity (e.g., unauthorized collaboration, plagiarism, multiple submissions, cheating on examinations, fabricating information, helping another person cheat, unauthorized advance access to examinations, altering or destroying the work of others, and fraudulently altering academic records) will result in your being sanctioned. Violations may subject you to disciplinary action including the following: receiving a failing grade on an assignment, examination or course, receiving a notation of a violation of academic integrity on your transcript (F!), and being suspended from the University. You have the right to appeal the charge. Contact the Office of Academic **Affairs, 101 Whitehurst, 405-744-5627, <http://academicintegrity.okstate.edu/>.**

Withdrawals and Incompletes: Please consult university schedules and policy. Incompletes ("I") will be given only in truly exceptional circumstances and students must be passing the course.

Internet Netiquette Guidelines (adapted from OSU Preparing Online Instruction course)

A melding of the words "network" and "etiquette", **netiquette** refers to the manner in which communication is conveyed in an electronic environment.

Outline of the Course:

The course consists of ten conceptual modules that introduce the principals of bioinformatics:

1. Molecular genetic background relevant to bioinformatics
2. Algorithmic thinking and expression for molecular genetics
3. Databases and Molecular File Structure
4. Biological Sequence Alignments
5. Pairwise Alignments and Database Searches
6. Patterns, Profiles, and Multiple Alignments
7. Inferring Evolutionary History: Phylogenetic Methods
8. Analysis of Genomes
9. Protein Structure/Function Prediction
10. Proteome and Gene Expression Analysis
11. Systems Biology & Metabolomic Reconstruction