BIOL / CMSI / HNRS 398-01
BIOLOGICAL DATABASES
http://myweb.lmu.edu/dondi/fall2008/cmsi398

Fall 2008　　　　　　　　Kam D. Dahlquist, PhD　　　　John David N. Dionisio, PhD
Doolan 219　　　　　　　　*kdahlquist@lmu.edu*　　　*dondi@lmu.edu,* AIM: *dondi2LMU*
TR 9:25–10:40am　　Office Hours: MW 1–3pm or by appointment　　TR 3–6pm or by appointment
3 semester hours　　　　　　Seaver 218; (310) 338-7697　　　　Doolan 106; (310) 338-5782

Course Objectives

This course is built upon L. Dee Fink's *taxonomy of significant learning,* as applied to biological databases. Long after the course concludes, our hope is that:

- You understand how biological information is encoded in the genome and can apply this knowledge to a variety of biological tasks and problems

- You understand the core concepts, structure, and functions of a database, ranging from individual files to a full relational database management system, and can perform useful tasks with such data

- You show discipline and proficiency in day-to-day science and engineering *best practices*, such as maintaining journals and notebooks, managing your files and code, and critically evaluating scientific and technical information

- You recognize and care about how the biological and technological issues presented in this course relate to and affect society, our daily lives, and ourselves

- You have some skills and tools for "leaving your comfort zone," flourishing outside of it, and learning more about biology and computer science on your own

- You learn how to effectively communicate and work with colleagues from different disciplines

Classroom and Laboratory Environment

We are all responsible for maintaining a classroom and laboratory environment that is safe and conducive to learning. As such, we will observe the following:

1. You are responsible for your own learning and for being a good class citizen.

2. We will act with honesty and integrity at all times.

3. We will always treat individuals with respect.

4. Class will start promptly on time.

5. You are expected to come to class having done the assigned reading and preparatory work.

6. You are expected to bring the required materials to each class session.

7. Cell phones, pagers, and other communication or music devices will be turned off.

Materials and Texts

- David P. Clark and Lonnie Dee Russell, *Molecular Biology Made Simple and Fun.* Cache River Press, 1997. (a.k.a. "the monkey book")

- Zoé Lacroix and Terence Critchlow, *Bioinformatics: Managing Scientific Data*, Morgan Kaufmann, 2003.

- Recommended but not required: Jeffrey Ullman and Jennifer Widom, *A First Course in Database Systems*, Third Edition. Prentice Hall, 2008.

- Assorted handouts, articles, and sample code to be distributed throughout the semester.

- Keck Lab account

Additional information is also available on the Web; do not hesitate to look for further sources of information regarding the concepts, techniques, tools, and paradigms that we will discuss.

Course Work and Grading

Graded coursework consists of assorted homework and exercises (25%), 1 midterm (25%), an ongoing course notebook/journal (25%), and 1 biological database project (25%). Letter grades are determined as follows: ≥ 90% gets an A– or better; ≥ 80% gets a B– or better; ≥ 70% gets a C– or better.   The instructors may curve grades upward based on qualitative considerations such as degree of difficulty, effort, class participation, time constraints, and overall attitude throughout the course. Grades are never curved downward.

Homework and In-Class Exercises

Homework and in-class exercises will be assigned throughout the semester, generally once per week. Each assignment will vary, but there will be a total of 100 points to earn in this category throughout the semester. Grading is based on effort and completeness; if you do the work and submit it on time, you will get full credit, regardless of correctness. The effort you put into your homework will assist you in doing well on the midterm exam and final project. The homework submission deadline is always the beginning of class on the designated due date. Submissions after the deadline receive half credit, and will not be accepted after one week has passed.

Midterm Examination

The midterm examination, worth 100 points, is currently scheduled for September 30. An unexcused absence from the midterm will result in a grade of zero.

Course Notebook/Journal

One of the most important skills you can develop as a scientist is keeping an excellent laboratory notebook. The computer science equivalent of the biology paper-based lab notebook is documentation, both in the form of comments in the code, commit logs, and user or developer manuals. For this course you will practice the documentation skills needed by both users and developers of bioinformatics tools by keeping an electronic course notebook or journal. The technology we will use is a MediaWiki Web site, open only to the students and instructors of this course, that we will create and edit during the semester. You will create an individual user page and make weekly entries that the instructors will read and grade. You will use the Wiki to complete some of the homework assignments as well. The following guidelines apply:

- Your weekly Wiki entry is due every midnight on Saturday.
- You will earn 5 points per weekly submission; failure to submit on time will result in a grade of zero for that week.
- The instructors will read and comment on how to improve your journal entries.
- You will earn an additional 25 points for your overall course notebook on the Wiki at the end

of the course. You will be allowed to go back and revise previous journal entries for the end-of-course evaluation. You will be able to earn back half-credit for filling in missed submissions.

Generally, your journal entries will consist of:

- Workflow and other documentation for hands-on exercises and projects
- Reflection on your learning. Each week you will answer questions such as:
  - What have I learned?
  - What isn't clear to me (yet)?

Biological Database Project

In the biological database project, you will apply what you have learned in the first two units of the course to perform an original bioinformatics research project. Students will work in interdisciplinary teams to create a Gene Database for the GenMAPP program for a new species not currently supported by the program. You will then use GenMAPP to analyze publicly available microarray data for that species. This project is worth 100 points towards your final grade for the course. You will be graded on three aspects of the project:

1. Biology
   - Results of your data analysis
   - Discussion of the results
2. Computer science
   - The Gene Database itself
   - README/documentation
3. Reflection
   - Self-assessment
   - Peer-assessment

Extra Credit: Seminars

Students may accumulate up to 2.5% of their final grade in extra credit by attending Biology Department seminars and completing the seminar sheets. Each seminar attended is worth 0.5% with up to 5 seminars (2.5%) total. Students arriving late to the seminar will not be granted credit — so don't be late! Certain non-Biology Department seminars may be approved in advance for extra credit at the instructors' discretion. To receive credit for these non-Biology Department seminars, you must turn in a one-page summary of the seminar within one week of the date of the seminar or they will not count as extra credit.

Attendance

Attendance at all class sessions is not absolutely required. However, there will be a sufficient number of group activities and in-class assignments that will be difficult to make up. Thus, if you must miss one or more class sessions, it is your responsibility to keep up with the course work. Note that the add/drop/withdraw-with-100%-refund deadline is August 30. The withdrawal or credit/no-credit status deadline is October 31.

University Policy on Academic Honesty

Loyola Marymount University expects high standards of honesty and integrity from all members of its community. All students are expected to follow the LMU honor code. As stated in the *LMU Undergraduate Bulletin 2008-2010*, pp. 58–59, "Examples of academic dishonesty include, but are not limited to, the following: all acts of cheating on assignments or examinations, or facilitating other students' cheating; plagiarism; fabrication of data, including the use of false citations; improper use of non-print media; unauthorized access to computer accounts or files or other privileged information and improper use of Internet sites and resources." You are required to sign the Honor Code Agreement for this course.

Students with Disabilities

Academic accommodations are available for students with disabilities who are registered with the Disability Support Service (DSS) Office within the Learning Resource Center (Daum Hall, Room 224). Please schedule an appointment with us early in the semester to discuss any accommodations for this course for which you have been approved.

Course Schedule

This schedule may change based on the actual ebb and flow of the class; deadlines, exams, university, and national dates (italicized) are less likely to change than course topics.

August/ September	Central dogma, the genetic code, the Human Genome Project, sequence databases; text files: formats, conversion, searching, regular expressions
August 30	*Last day to add or drop a class without a grade of W; last day to withdraw from the semester for a 100% refund*
September 30	Midterm
October	Cancer, biological pathways; tabular data, visualization, XML
October 20	*California voter registration deadline: http://www.sos.ca.gov/elections/elections_vr.html*
October 31	*Withdraw/credit/no-credit deadline*
November	Project workshop; relational databases
November 4	*Election day — go out and vote!*
November 26–28	*Thanksgiving; no class*
December	Project workshop
December 11	*Project due*

You can view the class calendar on the Web at *http://ical.mac.com/dondi/LMU*. If you have an iCalendar-savvy client (i.e., Mozilla Calendar, Ximian Evolution, KOrganizer, Apple iCal, etc.), you can subscribe to the class calendar at *webcal://ical.mac.com/dondi/LMU.ics*. On-the-fly updates and adjustments to the class schedule will be reflected in this calendar.