

# BIOL 498/CMSI 698

## SPECIAL STUDIES: BIOINFORMATICS

<http://myweb.lmu.edu/dondi/spring2006/cmsi698>

### Spring 2006

Doolan 222

R 6:30 – 9:30pm

3 semester hours

Kam D. Dahlquist, PhD

**e-mail:** [kdahlquist@lmu.edu](mailto:kdahlquist@lmu.edu)

**Ofc Hrs:** MW 10:30am–12:00pm, R 5–6pm

Seaver 218; (310) 338-7697

John David N. Dionisio, PhD

[dondi@lmu.edu](mailto:dondi@lmu.edu), AIM: dondi2LMU

TR 1:30–3pm, 4:30–6pm

Doolan 106; (310) 338-5782

## Course Objectives

---

This course aims to introduce a broad range of concepts and techniques in bioinformatics, including the underlying biological questions that these techniques are used to answer, and to apply selected computer science topics, such as information modeling, database management, and visualization, to the bioinformatics domain.

Bioinformatics is the application of information technology — the representation, organization, manipulation, distribution, maintenance, and use of digital information — to biological data. Bioinformatics uses databases and analytical tools to answer biological questions. The field is inherently interdisciplinary, involving aspects of biology, computer science, mathematics, physics, chemistry, and engineering. The need for computational methods has recently exploded due to the huge amounts of data produced by genome sequencing projects and other high-throughput technologies. Bioinformatics is moving the field of biology from a “one gene at a time” approach to the analysis of whole systems. Topics may include database management, information modeling, data visualization, sequence alignment and phylogeny, comparative genomics, protein structure prediction, pathways and networks, and the analysis of high-throughput genomic and proteomic data.

## Course Requirements

---

Proficiency in at least one high-level programming or scripting language (Java, C++, Perl, Python, JavaScript); a prior course in data structures and algorithms (LMU CMSI 281 or equivalent); ability to document, demonstrate,

and explain one’s work; willingness to participate actively in class discussions.

## Materials and Texts

---

There is one required text:

- Andreas D. Baxevanis and B. F. Francis Ouellette, editors. *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins*. 3rd edition, John Wiley & Sons, 2005.

The following are optional, and will likely be used sporadically in class or exclusively as an additional resource for background information:

- Jean-Michel Claverie and Cedric Notredame. *Bioinformatics for Dummies*. Wiley Publishing, 2003.
- Larry Gonnick and Mark Wheelis. *The Cartoon Guide to Genetics*. Collins, 1991.

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

## Course Work and Grading

---

Graded coursework consists of accumulated homework (20%), 1 midterm (20%), 1 class paper (20%), 1 class project (20%), and 1 final exam (20%). Letter grades are determined as follows:  $\geq 90\%$  gets an A– or better;  $\geq 80\%$  gets a B– or better;  $\geq 70\%$  gets a C– or better. The instructor 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

Homework consists of questions, exercises, and programming assignments to be given throughout the semester. Homework is where you can learn from your mistakes without grading penalty: if you do the work and submit it on time, you will get full credit, regardless of correctness. What goes around comes around: the effort you put into your homework pays off in the tests and the research paper. The homework submission deadline is always the beginning of class on the designated due date; the due date is encoded in the homework number. Submissions after the deadline receive half credit, period. Extra credit homework may be assigned; fulfilling this is counted *on top of* the 20% allocation of homework to your final grade.

## Class Project and Paper

For this class, we will work together on a bioinformatics software package that we will release as an open source project and for which, at the end of the semester, we hope to have a publication-quality paper that will be submitted to an appropriate journal. The exact domain and function of the project will be determined in the early part of the semester. The versions of the project and paper that will be graded are due at the beginning of our last class, April 27.

To compel you to focus on the content of the paper as opposed to busy work such as formatting and reference management, the paper will be written using LaTeX. If necessary, we can talk about LaTeX in class, and templates to get you started will be provided.

Project grading is based on these criteria:

1. *Design (30%)*: How good is the overall structure of the program? Is it easy to understand, flexible, and easy to maintain? Is it elegant or innovative? Does it follow the “one change, one place” rule of thumb?
2. *Functionality (30%)*: How well does the program work? Does it accomplish the project’s goals? Are its results accurate or correct? Does it perform its tasks in a reasonable amount of time?

3. *Naming (20%)*: Are program entities — classes, subroutines, variables — named clearly and consistently? Do their names correspond to their functions and roles?
4. *Comments (15%)*: Are comments provided where appropriate? Are they clear and well-written? Does the project take advantage of any special comment features in the language or platform (e.g., Javadoc in Java)?
5. *CVS use (5%)*: Do you commit your code at reasonable intervals? Do you provide adequate descriptions of your commits?

Paper grading is based on these criteria:

1. *Content (40%)*: What is the quality of the work? Specific assessment of content will depend on the type of paper or topic chosen.
2. *Organization (30%)*: Is the paper well-structured? Are concepts and the flow of ideas easy to follow? Are distinct sections or topics clearly identified?
3. *Writing (20%)*: Are statements clear and easy to follow? Is the language precise, unambiguous, and grammatically correct?
4. *Polish (10%)*: Is the content properly proof-read? Are there many misspellings, typos, or other formatting faux pas?

## Tests

The midterm is initially scheduled for February 16. The final exam is scheduled for May 4. All tests are open-*paper*-everything; no sharing. “Open computer” *might* also be allowed depending on the scope, subject matter, or circumstances. You may neither solicit nor give help while the exam is in progress. Late and/or missed tests are handled on a case-to-case basis; in all instances, talk to us about them.

## Attendance

We aren’t sticklers for attendance, but we do like having a full class. Remember that the late registration and change of program deadline is January 13. The deadline for withdrawal or credit/no-credit status is March 17.

## University Policy on Academic Honesty

Loyola Marymount University expects high standards of honesty and integrity from all members of its community. Applied to the arena of academic performance, these standards preclude all acts of cheating on assignments or examinations, plagiarism, forgery of signatures or falsification of data, unauthorized access to University computer accounts or files, and removal, mutilation, or deliberate concealment of materials belonging to the University Library.

## Course Schedule

---

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

<b>January</b>	Biological concepts: molecular genetics; systems biology perspective; biological databases
<i>January 13</i>	<i>Last day to add or drop classes for a 100% tuition refund</i>
<b>February</b>	Project startup; working with sequences and structures
February 16	Midterm
<b>March</b>	Functional genomics: analysis of DNA microarray data
<i>March 17</i>	<i>Last day to withdraw from classes or apply for Credit/No Credit grading</i>
<b>April</b>	Pathway reconstruction and visualization
<i>April 27</i>	<i>Project and paper due</i>
<i>May 4</i>	<i>Final Exam, 6:30pm</i>

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.