1. **What is the course modification?**

This is a new course.

2. **Why is this course being requested or modified?**

Biology is being transformed by genomics, which is the study of the entire coding capacity of an organism's genetic blueprint. Genomics is impacting not only basic biology via the sequencing of model organisms, but also agriculture via the sequencing of crop plant and farm animal genomes. Dozens of bacterial and eukaryotic genomes have already been completely sequenced, and dozens more genomes will be completed shortly. Determination of a genome's DNA sequence reveals the entire complement of genes (between 20,000 and 30,000 for most plants and animals). Methods have been developed that allow the simultaneous study of all of these genes, yielding unprecedented amounts of data for today's biologist. Training on how to manage and analyze these vast data streams has become an essential part of most biologists' education.

3. **How will the content be organized?**

This is mostly a lecture-based course with homework assignments using internet and personal computers.

4. **What other courses at UHM closely parallel the proposed course and in what way will the latter make a distinct contribution?**

Two ICS courses address bioinformatics issues at the undergraduate level:

- ICS 491 SPRING 2005 INTRODUCTION TO BIOINFORMATICS
- ICS 471 SPRING 2005 PROBABILITY AND STATISTICS FOR BIOINFORMATICS

These courses are primarily aimed at Informatics and Computer Science undergraduates, not biologists. MBBE483 is focused on connecting the technology, informatics with biology.

5. **Where or how does the proposed course fit into the current and future curriculum?**

MBBE483 targets mid- to upper-level undergraduates in CTAHR departments (MBBE, Bioengineering, Plant and Environmental Biotechnology, TPSS, PEPS and NREM) and the Biology Program. Students who take this course will already have a solid biology background and a good understanding of Cell Biology and Genetics. This course will give students the ability to understand and work with large data sets that are now common in the biological sciences.

6. **Why is the number of credits and level justified? Explain the prerequisites and the absence thereof.**

The course will meet twice per week for 75 minutes of lecture each.
7. **How will the course assist students to achieve the critical skills and competencies expected of CTAHR graduates?**

**Written communication** – Students will be required to hand in written homework and communicate clearly on exams.

**Oral communication** – Short oral presentations by the students, as well as class participation, will be encouraged. These will account for 10% of the course grade.

**Analytical problem solving** - Bioinformatics involves analytical problem solving involving large data sets.

**Personal characteristics** – Success in the course will require persistence, dependability and hard work on part of the students.

**Human relations skills** – Students will be allowed to work in small groups for some of their homework.

**“Real World” experience** – Students will analyze data collected by genomics experts, using the same tools as are used in industrial and other academic settings.

**Business management skills** - Not applicable.

**Leadership skills** – Familiarity with the vast amounts of biological data available today and the knowledge of how to effectively analyze them will automatically put these students in leadership positions.

**Computer skills** – Students will learn about current and up-to-date analysis methods, all of which involve computers.

**Global perspectives** – Genomics is a global activity as evidenced by the many large and multinational genome sequencing projects. Where applicable, global efforts will be specifically highlighted in the class.

8. **How will students be evaluated?**

Students will work in teams and address real-world problems. They will be evaluated based on their ability to solve the problems posed, on both individual and team basis.

9. **What are the minimum qualifications for teaching this course? Is a qualified instructor now available?**

To be maximally effective, this course requires an instructor with a background in biology.
agriculture and bioinformatics, and ideally hands-on experience in genomics, i.e. genome sequencing and microarray experiments. Dr. Presting headed the public physical mapping effort for the rice genome project, participated in the shotgun sequencing of the rice genome, has led an effort in papaya BAC end sequencing and is currently involved in analyzing the repeat sequences of corn centromeres.

10. **How will the course be financed, assuming no further cutbacks?**

If enrollment stays around 30, this course requires no additional financing. If enrollment exceeds 40 students, a teaching assistant will be required to assist students with homework and help the instructor with grading.

11. **Has the course been offered before? Is there a demand for it?**

This course has not been offered before. Students within CTAHR and the Biology program have indicated a strong interest in taking this course if it was offered. Three undergraduate students have taken the related graduate level course MBBE683 that has a large computer laboratory component. Ten of 12 students who took MBBE683 gave it the highest overall rating. MBBE faculty are fully supportive of this course being offered through the department.

12. **Is the course cross-listed with another department?**

This course will be cross-listed with the Biology Program, as Biology students will benefit as much from this course as their CTAHR counterparts and are expected to present a significant portion of the clientele.
Bioinformatics
MBBE483/BIOL483
Spring 2008

Course Syllabus

Lecture Meeting time and place: 3:00-4:15 T Th, To be determined.
Professor: Gernot Presting
Contact Information:
   email: gernot@hawaii.edu
   office: 515 St. John Plant Science Bldg

Office hours: Tuesday 10-11 AM and Wednesday 2-3 PM, by appointment, or drop by.

Course Description
In this course we explore the major concepts and tools used in Bioinformatics. Generally defined as the application of computational methods to biological data, bioinformatics encompasses a wide spectrum of fields. Although this course will focus primarily on molecular data from DNA, RNA, and protein sequences, given the importance of biogeographic research in Hawaii we will also touch on other important biological data such as museum specimen records, geographic distribution data, and ecological data. We will cover the basics of molecular biology, software tools used by bioinformaticians, sequence alignment, models of sequence evolution, phylogenetic methods, RNA and protein structure prediction. This introduction will provide you with a strong foundation on which to build for more advanced topics or concepts in the field. If there is anything you see or hear during the semester that piques your interest please let me know and I’ll do my best to help you get more information about it.

Course Website:
Relevant instruction materials and exam results will be made available through webCT. You MUST self-register on this site in order to see your exam results (which are only visible to you and the instructor).

Textbook:

Exams and Grades:
Exams are non-cumulative. However, successfully solving the problems of a later exam may require “common knowledge” acquired earlier in the course. Exam questions will be a mixture of multiple-choice, short answer, fill-in-the-blank, problem solving and/or essay.

<table>
<thead>
<tr>
<th>Exam Type</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st midterm</td>
<td>25%</td>
</tr>
<tr>
<td>2nd midterm</td>
<td>25%</td>
</tr>
<tr>
<td>Final exam</td>
<td>25%</td>
</tr>
<tr>
<td>Class participation</td>
<td>10%</td>
</tr>
<tr>
<td>Homework</td>
<td>15%</td>
</tr>
</tbody>
</table>

Final examination: 13 May 9:45 – 11:45
<table>
<thead>
<tr>
<th>DAY</th>
<th>DATE</th>
<th>LECTURE TOPIC</th>
<th>SUGGESTED READING</th>
</tr>
</thead>
<tbody>
<tr>
<td>Th</td>
<td>Jan. 11</td>
<td>Course Introduction</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>Jan. 16</td>
<td>Introduction to Molecular Biology</td>
<td>Chapter 1</td>
</tr>
<tr>
<td>Th</td>
<td>Jan. 18</td>
<td>Why Bioinformatics?</td>
<td>Reading 1</td>
</tr>
<tr>
<td>T</td>
<td>Jan. 23</td>
<td>Genomics Tools and Techniques</td>
<td>Chapter 1</td>
</tr>
<tr>
<td>Th</td>
<td>Jan. 25</td>
<td>Genomics, Genetic Markers and Breeding</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>Jan. 30</td>
<td>Alignments and Data Searches</td>
<td>Chapter 2</td>
</tr>
<tr>
<td>Th</td>
<td>Feb. 1</td>
<td>Alignments and Data Searches</td>
<td>Chapter 2</td>
</tr>
<tr>
<td>T</td>
<td>Feb. 6</td>
<td>Models of Sequence Evolution</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>Th</td>
<td>Feb. 8</td>
<td>Models of Sequence Evolution</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>T</td>
<td>Feb. 13</td>
<td>Review of Model Selection</td>
<td>Reading 2</td>
</tr>
<tr>
<td>Th</td>
<td>Feb. 15</td>
<td>Phylogenetics – Distance Methods</td>
<td>Chapter 4</td>
</tr>
<tr>
<td>T</td>
<td>Feb. 20</td>
<td>Phylogenetics – Distance &amp; Parsimony</td>
<td>Chapter 4, 5</td>
</tr>
<tr>
<td>Th</td>
<td>Feb. 22</td>
<td>Phylogenetics – Parsimony Methods</td>
<td>Chapter 5</td>
</tr>
<tr>
<td>T</td>
<td>Feb. 27</td>
<td>1st Mid-Term Exam</td>
<td></td>
</tr>
<tr>
<td>Th</td>
<td>Mar. 1</td>
<td>Phylogenetics – Distance vs. Character-Based Methods</td>
<td>Reading 3</td>
</tr>
<tr>
<td>T</td>
<td>Mar. 6</td>
<td>Phylogenetics – Probabilistic Methods (Likelihood and Bayesian)</td>
<td>Sullivan 2005</td>
</tr>
<tr>
<td>Th</td>
<td>Mar. 8</td>
<td>Phylogenetics – Probabilistic Methods</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>Mar. 13</td>
<td>Genome &amp; Gene Recognition</td>
<td>Reading 4</td>
</tr>
<tr>
<td>Th</td>
<td>Mar. 15</td>
<td>Genome &amp; Gene Recognition</td>
<td>Chapter 6</td>
</tr>
<tr>
<td>T-Th</td>
<td>Mar. 20-22</td>
<td>SPRING BREAK!</td>
<td>TBA</td>
</tr>
<tr>
<td>T</td>
<td>Mar. 27</td>
<td>Genome &amp; Gene Recognition</td>
<td>Reading 5</td>
</tr>
<tr>
<td>Th</td>
<td>Mar. 29</td>
<td>RNA Structure Prediction</td>
<td>Chapter 7</td>
</tr>
<tr>
<td>T</td>
<td>Apr. 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Th</td>
<td>Apr. 5</td>
<td></td>
<td>Reading 6</td>
</tr>
<tr>
<td>T</td>
<td>Apr. 10</td>
<td>2nd Mid-Term Exam</td>
<td></td>
</tr>
<tr>
<td>Th</td>
<td>Apr. 12</td>
<td>Protein Structure Prediction</td>
<td>Chapter 7</td>
</tr>
<tr>
<td>T</td>
<td>Apr. 17</td>
<td>Protein Structure Prediction</td>
<td>Chapter 7</td>
</tr>
<tr>
<td>Th</td>
<td>Apr. 19</td>
<td>Protein Structure Prediction</td>
<td>Reading 7</td>
</tr>
<tr>
<td>T</td>
<td>Apr. 24</td>
<td>Proteomics</td>
<td>Chapter 8</td>
</tr>
<tr>
<td>Th</td>
<td>Apr. 26</td>
<td>Review</td>
<td></td>
</tr>
<tr>
<td>M-F</td>
<td>May 4</td>
<td><strong>FINAL EXAM</strong></td>
<td></td>
</tr>
</tbody>
</table>

Document 1

11/19/07
1. Why bioinformatics? What can you do with it and why is it important?

2. Sequence Evolution models

3. Distance and Parsimony

4. Probabilistic methods

5. Genome & gene recognition

6. RNA structure prediction

7. Protein structure prediction