

Comparative Genomics
Bi 407/507, Fall 2003

Time: Tuesdays 4:30-6:20 pm

Dates: 29 Sep - 12 Dec. 2003

Location: Room 389 Neuberger Hall

Instructor: Dr. Susan Masta

Office: M103 (inside Room 103), Science Building 1

Email: smasta@pdx.edu

Phone: 725-8505

Office Hours: Mondays 4:00 - 5:00 pm, and by appointment

Course Description: The completion of the sequencing of the human genome was announced this spring. But how do we interpret what the sequence means? This course will introduce you to the approaches and techniques used to sequence and analyze genomes. We will examine how comparing genomes within and between species is vital to understanding and interpreting genomic data. The course will combine discussion of current literature with hands-on computer analyses to familiarize you with the burgeoning field of comparative genomics.

Course Objectives: You will become familiar with:

Where to find genomic information and resources

How to find and annotate genes

How to make phylogenomic comparisons

What we have learned from comparing genomic data

The history of the Human Genome Project: who, what, when, where, and how

What we can do, can't do, and may soon be able to do, with genomic data

Assignments: There will be weekly assigned readings, that will often be available via WebCT or email. You will need to prepare two questions or comments based on these readings in advance of each class, and we will begin our discussions by each presenting our questions. All students will be expected to participate in the classroom discussions. For one of the classes, each student will have an individual research/reading assignment they will need to summarize for the class. There will be computer-based assignments to familiarize you with some of the databases and analyses used in genomics. These will be explained and illustrated during class. Graduate students will prepare for and lead a portion of one of the classes. There will be no quizzes, tests, or final exam.

Grading: Grades will be based upon attendance, class participation, the quality of the preparation for the discussion section the student helps lead, and the computer assignments.

Syllabus (tentative):

30 September: an introduction to comparative genomics

What is "comparative genomics"?

Timeline of comparative genomics developments

Overview of syllabus, expectations

Databases for genomics resources: NCBI

Assigned readings:

- Pennisi, E. 2001. The Human Genome. *Science* 291:1177-1180.
Roberts, L. 2001. Controversial from the start. *Science* 291:1182-1188.
Baltimore, D. 2001. Our genome unveiled. *Nature* 409:814-816.
Wade, N. 2003. Scientists say human genome is complete. *The New York Times*.

Computer assignment: NCBI website questions

7 October: The Human Genome Project.

Discussion of above papers:

- The history of the Human Genome Project. Why, when, who and how.
- The public versus the private sector
- Why sequence the human genome?
- Ethical and societal issues: Iceland's Genomic Database as a case study

Assigned reading: read pages 863-875, beginning with 'Hierarchical shotgun sequencing'

- I. H. G. S. Consortium. 2001. Initial sequencing and analysis of the human genome. *Nature* 409:860-921.

14 October: Obtaining and Assembling Sequences.

Discussion of assigned reading on hierarchical shotgun sequencing

- Large-scale sequencing methods: cloning and BAC library creation
- Sequence quality scoring, vector screening (phred)
- Sequence assembly into contigs and skeleton frameworks (phrap)
- Introduction to evolutionary thinking:
- Phylogenetic analyses: tree terminology and parsimony
- Concept of homology

Assigned readings on genome annotation:

- Wolfsberg, T. G., J. McEntyre, and G. D. Schuler. 2001. Guide to the draft human genome. *Nature* 409:824-826.,
- Birney, E., A. Bateman, M. E. Clamp, and T. J. Hubbard. 2001. Mining the draft human genome. *Nature* 409:827-828.

21 October: Gene Identification and Annotation I:

Computer lab (322 Cramer Hall)

- Sequence alignment: global versus local
- Databases and tools for annotating sequence (FASTA, BLAST)
- Modifying search strategies; searching different databases
- Inferring gene function from relatedness to other genes
- Finding Open Reading Frames (ORFs)
- distinguishing introns from diverged exons

Computer assignment: gene annotation I(BLAST)

28 October: Gene Identification and Annotation II:

Computer lab (322 Cramer Hall)

CLUSTALX

Sequence alignment based on predicted secondary and tertiary structures
tRNA folding introduction (tRNAscan)
Mapping to known regions
Other gene finding methods

Computer assignment: gene annotation of a mitochondrial genome (BLAST, finding ORFs, tRNAscan)

Assigned reading: Saccone et al. 1999. Evolutionary genomics in Metazoa: the mitochondrial DNA as a model system. Gene 238:195-209.

4 November: Genome Comparisons I: Organelles

Phylogenetic analyses: introduction to programs
Discussion of mitochondrial genome paper
Genome size, content, and gene order
The minimal genome
Mapping characters onto a phylogenetic tree

Assignment: Survey of bacterial genomes

Assigned reading: Salzberg et al, 2001. Microbial genes in the human genome: lateral transfer or gene loss? Science 292: 1903-1906.

11 November: University closed for Veterans Day - No Class

18 November: Genome Comparisons II: Bacteria

Student summaries of bacterial genomes
Phylogenetic analyses to determine relationships and interpret character evolution
Class discussion of "Microbial genes in the human genome: lateral transfer or gene loss?"

Assigned reading: parts of Waterston et al. 2002. Initial sequencing and comparative analysis of the mouse genome. Nature 420:520-562.

25 November: Genome Comparisons III: Vertebrates

Vertebrate evolution based on genome comparisons
Class discussion of the mouse genome paper
Graduate student presentation: Human and chimpanzee genome comparisons

Assigned reading: Rokas, A., B. L. Williams, N. King, and S. B. Carroll. 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature 425:798-804.

2 December: The Future of Comparative Genomics.

In-class assembly of a mitochondrial genome
Class discussion of "Genome-scale approaches to ... phylogenies".
Discussion of the future of genomics