Course Syllabus

Jump to Today

Syllabus

The overall objectives of the course are: (a) to familiarize students with molecular diagnostic technologies, (b) to increase students' intuition and understanding of computational methods used to analyze molecular diagnostic data, and (c) to build students' abilities to interpret molecular diagnostic testing and to integrate results into clinical decision making. Students will be introduced to principles of molecular diagnostics through video lectures (in preparation for classes), live lectures, case conferences, and computer laboratories. Fundamental computational and quantitative aspects of molecular diagnostics will be emphasized throughout the course. Take-home problem sets and laboratory sessions will provide hands-on practice with bioinformatics methods and computational tools. In class case discussions led by instructors will help students hone the analysis and interpretation skills in molecular diagnostics. Students will apply these same skills on their own while working in teams to prepare, analyze, interpret, and clinically correlate assigned molecular cases for class presentation. Assessment will be accomplished through three take-home problem sets (45%), student group presentations of molecular cases (30%), video lecture quizzes (10%), bioinformatics laboratory (5%), and overall class participation (10%). Students will be expected to have an understanding of human genetics and molecular biology obtained through HST 160/161 or equivalent coursework.

Optional textbooks

Bioinformatics Algorithms: An Active Learning Approach, 2nd Ed. Vol. 1 and Vol. 2
Phillip Compeau and Pavel Pevzner
Active Learning Publishers LLC, 2015

ISBNs 978-0990374619 and 978-0990374626

http://bioinformaticsalgorithms.com **(http://bioinformaticsalgorithms.com)**

These two books were developed for an online course. The books focus on overall concepts as well as algorithmic details, and are quite readable.

Bioinformatics and Functional Genomics 3rd Edition
Jonathan Pevsner
Wiley-Blackwell, 2015
ISBN 978-1118581780

This book focuses a bit more on applications than algorithmic details, although it describes some fundamental algorithms such as BLAST exceptionally well.

Session #1 (October 30, 2018, 2pm-5pm)

1. Preparation before class (Optional): Compeau & Pevzner Vol. 1 Chapters 1 & 5; Pevsner Chapters 3

Bioinformatics Part I: introductory concepts and sequence alignment algorithms (Dr. Gerber). Duration ~90 mins.

Bioinformatics Lecture 1.pptx **Preview the document**

Learning objectives for this lecture are to:

- Understand fundamental concepts of computer algorithms including correctness and run-time notation
- Be able to define the general computational problem of pair-wise sequence alignment, and local versus global alignment variations
- Understand the mathematical framework for evaluating the quality of a sequence alignment, including rationale for substitution/scoring matrices and different gap penalties
- Understand the concept of dynamic programming
- Understand and be able to apply the Needle-Wunsch (global) and Smith-Waterman (local) dynamic programming algorithms

Homework#1 (15% of grade) covering concepts from Bioinformatics Part I – due before Session #2

Problem Set 1.docx **Preview the document**

2. **Molecular Diagnostics Part I (Sanger Sequencing and NGS Technologies~90 minutes total)**:

Preparation before class

- Required: Watch Molecular Video Lectures 04 (see **Assignment page** or **Molecular Pre-Class Video Lectures for October 31st Module** for links to videos). **Molecular Video Lecture 01** is a basic review of concepts from HST 160. Reviewing this video and its quiz are due by the end of the course and may be helpful for the October 31st class.
- **Molecular Video Lecture 04: Nucleic Acid Quantitation, PCR, qPCR, MLPA**
  - Spectrophotometry of nucleic acid
  - Fluorescent-based quantitation of nucleic acid
  - Polymerase chain reaction
  - Quantitative real-time PCR
  - Multiplex ligation probe amplification

Chalk talk on quick PCR review and NGS platforms

Split into 8 groups for end of course case presentations.

Molecular Homework:

Molecular Video Lecture 05: PCR Sizing and Molecular Video Lecture 05 quiz

Session #2 (November 06, 2018, 2pm-5pm)

1. **Molecular Diagnostics Part II (60 minutes total)**:

Preparation before class

- **Molecular Video Lecture 09: NGS Applications, Target Enrichment**
  - Global review of NGS platforms
  - Whole genome sequencing
  - Clinical NGS in practice
  - Whole exome capture by hybridization
  - Haloplex target enrichment
  - Raindance microfluidic-based target enrichment
  - Fluidigm plate-based PCR array
  - Ion Torrent Ampliseq multiplex PCR
  - Illumina TruSeq Amplicon
  - Molecular Inversion Probe technology
- In-class somatic NGS case discussions (links to cases will be posted later) (55 minutes)

2. Review and discussion of Problem Set #3 (Dr. Gerber). Duration 30 mins.

3. Preparation before class (Optional): Compeau & Pevzner Vol. 2 Chapter 9; Pevsner Chapter 4

Bioinformatics Part II: BLAST and short-read mapping (Dr. Gerber). Duration ~90 mins.

Bioinformatics Lecture 2.pptx  **Preview the document**

Learning objectives for this lecture are to:

- Understand how the Basic Local Alignment Search Tool (BLAST) algorithm works and the rationale for its development
- Understand the mathematical framework for evaluating BLAST results (*E* values, *p*-values and bit scores)
- Understand why efficient short-read mapping algorithms are needed
- Understand the concept of using data structures to speed up computation
- Understand the suffix tree data structure and how it can be applied to short read mapping
- Understand the Burrows-Wheeler Transform (BWT), be able to apply an algorithm to calculate the inverse BWT, and understand how these techniques are used in short-read mapping algorithms

Homework#2 (15% of grade) covering concepts from Bioinformatics Part II – due before Session #3

Problem Set 2.docx  **Preview the document**

Session #3 (November 13, 2018, 2pm-5pm)

- Required before class: Watch Molecular Video Lectures 02-03 (see **Assignment page** or ▾ **Molecular Pre-Class Video Lectures for November 7th Module** for links to videos)
- **Molecular Video Lecture 02: Variant Sizes, Cytogenetics, FISH**
  - Variant sizes/scale

  - Karyotyping technology
  - Strengths and limitations of karyotyping
  - Fluorescence in situ hybridization technology
  - FISH technique
  - FISH analysis/interpretation
- **Molecular Video Lecture 03: Array CGH**
  - Array technology
  - Array comparative genomic hybridization technique
  - Array CGH analysis
  - Strengths and limitations of array CGH
  - Uniparental disomy
  - Clinical utility of array CGHCase Discussions (Dr. Le)

Chalk talk video lecture review (20 minutes)

In-Class Case Discussion

- **Case 01: Fluorescence In Situ Hybridization (FISH) Teaching Set (20 mins)**
- **Case 02: Autism/Developmental Delay in a Child (20 mins)**

2. Review and discussion of Problem Set #2 (Dr. Gerber). Duration 30 mins

Preparation before class (Optional): Compeau & Pevzner Vol. 2 Chapter 7; Pevsner Chapters 7 & 9

3. **Bioinformatics Part III**: phylogenetic trees (Dr. Gerber). Duration ~60mins.

Bioinformatics Lecture 3.pptx **Preview the document**

Learning objectives for this lecture are to:

- Understand rationale for building phylogenetic trees and learn nomenclature
- Be able to describe the general steps for phylogenetic tree construction
- Understand and apply distance-based algorithms for phylogenetic tree construction
- Understand general concepts behind maximum-likelihood and Bayesian phylogenetic tree construction algorithms
- Understand the strengths and limitations of major phylogenetic tree construction algorithms
- Understand frameworks for evaluating phylogenetic trees and how to interpret evaluation metrics

Homework#3 (15% of grade) covering concepts from Bioinformatics Part III – due before Session #4

Bioinformatics Problem Set #3

Session #4 (November 20, 2018, 2pm-5pm)

1. **Review and discussion of Problem Set #3** (Dr. Gerber). Duration 30 mins.

2. **Next generation sequencing for Clinical Microbiology** (Dr. Lynn Bry). Duration ~50 mins.

The objectives of this lecture are to:

- Understand the rationale for using next generation sequencing assays for infection control and other Clinical Microbiology applications
- Understand the architecture and rationale for a complex clinical bioinformatics pipeline
- Understand challenges specific to Clinical Microbiology in applying next generation sequencing and bioinformatics methods

3. **Molecular Diagnostics Part III** (Dr. Le, 60 minutes total):

- Understand circulating tumor cells vs cell-free/circulating tumor DNA/RNA as substrates for molecular diagnostics
- Clinical utility of ctDNA testing
- Logistics and laboratory techniques for testing cell-free DNA/RNA (ctDNA)
- Molecular barcoding/unique molecular indices
- Bioinformatics analysis for ctDNA

3. **In-class molecular case discussions** (Dr. Le, 30 minutes total, links to cases will be posted later)

Session #5 (November 27, 2018, 2pm-5pm)

1. **Bioinformatics Part IV**: variant calling and bioinformatics file formats (Dr. Gerber). Duration 45 mins.

Bioinformatics Lecture 4.pptx **Preview the document**

Learning objectives for this lecture are to:

- Understand general types of algorithms for finding sequencing variants
- Understand the main concepts behind competing algorithms for single nucleotide variant calling for somatic mutations, and the rationale for combining methods
- Understand important bioinformatics file formats including those for storing nucleotide sequences, quality scores, alignments to a reference, and variant calls

2. **Bioinformatics computer lab** (Dr. Le, Maciej Pacula). Duration ~120 mins.

Learning objectives are to:

- Understand NGS file formats: fasta, fastq, bam/sam, vcf
- Perform sequence alignment and generate a sam file
- Visualize alignment pileup and perform variant calling
- Generate a vcf file
- Annotate variants according to HGVS format
- Interpret variants based on knowledge base/literature search

Bioinformatics Lab Quiz

Session #6 (December 04, 2018, 2pm-5pm)

1. **Molecular Diagnostics Part III**: Advanced Topics (Dr. Le). Duration ~60mins.

The objectives of this lecture are to:

- Understand the limitations of short-read NGS technology and the importance of phasing
- Molecular signatures: total mutation burden, microsatellite instability
- Become familiar with big data analytics in healthcare (if time permits)

2. Molecular Diagnostics Part III: Lab management (Dr. Le). Duration ~60 mins.

Learning objectives are to:

- Understand clinical validity vs. clinical utility
- Understand insourcing vs. outsourcing testing
- Understand laboratory capital costs, operating costs, fixed costs, variable costs
- Understand clinical sensitivity, clinical specificity, analytical sensitivity, analytical specificity, PPV, NPV
- Understand assay design, validation, proficiency
- Understand quality management
- Understand LDT vs IVD regulation and coverage

3. Bioinformatics pipelines in practice (Dr. Le). Duration ~50 minutes

Recommended Pre-Class Bonus Video Lecture: **Molecular Video Lecture 10**

- Clinical vs research analysis
- Software engineering
- Algorithm development vs open source tools
- Quality system

Session #7 (December 11, 2018, 2pm-5pm)

1. Group presentations and case discussions (led by Drs. Long Le and Georg Gerber). Duration ~180 mins. 8 groups of ~4 students will present (30% of grade)

Learning objectives are to:

- Deepen knowledge and understanding of applications of molecular diagnostics in the context of clinically relevant scenarios
- Increase understanding of clinical laboratory diagnostic modalities and technologies
- Build skills in appropriately choosing and evaluating diagnostic tests for patients

Required Molecular Diagnostics Videos and Quizzes (due by Dec 11, midnight)

- **Molecular Video Lecture 01: Variants, Nomenclature, Nucleic Acid Extraction**

- **Molecular Video Lecture 02: Variant Sizes, Cytogenetics, FISH**
- **Molecular Video Lecture 03: Array CGH**
- **Molecular Video Lecture 04: Nucleic Acid Quantitation, PCR, qPCR, MLPA**
- **Molecular Video Lecture 05: PCR Sizing**)
- **Molecular Video Lecture 09: NGS Applications, Target Enrichment**

Other Molecular Diagnostics Assignments

Video quizzes due by midnight on last day of course, December 11, 2018:

- **Molecular Video Lecture 01: Variants, Nomenclature, Nucleic Acid Extraction**
  - Germline vs somatic testing
  - Different types of variants
  - HGVS variant nomenclature
  - Preanalytical sample assessment
  - Nucleic acid extraction
  -
- **Molecular Video Lecture 05: PCR Sizing**)
  - Capillary electrophoresis
  - Microsatellite/STR genotyping
  - Microsatellite instability testing in HNPCC/Lynch syndrome
  - STR genotyping for chimerism testing in bone marrow transplant

Bonus videos and associated quizzes for extra points:

- **Molecular Video Lecture 06: Classical Genotyping Techniques**
  - Allele specific PCR
  - Methylation-specific PCR
  - Restriction fragment length polymorphism (RFLP)
  - Sanger sequencing
  - Single base extension genotyping (SNaPshot)
  - Other genotyping techniques (single-strand conformation polymorphism (SCP), melting curve analysis, invader assay)
- **Molecular Video Lecture 07: The Human Genome Project**
  - Project
  - Technology/techinque
  - Cost/time
  - HGP (Sanger sequencing) vs NGS

- **Molecular Video Lecture 08: NGS Platforms**
  - Massive parallelization
  - Emulsion PCR (emPCR)
  - Whole-genome library construction
  - Illumina Y-adapter library construction
  - 454 pyrosequencing
  - Illumina clonal cluster generation (in-situ PCR)
  - Illumina sequencing by synthesis w/ reversible dye terminators
  - SOLiD sequencing by ligation (deprecated, skip if time limited)
  - Ion Torrent semiconductor pH-based sequencing
  - Pacific Bioscience single molecule sequencing by synthesis w/ reversible dye terminators
  - Helicos single molecule sequencing by synthesis w/ reversible dye terminators (deprecated, skip if time limited)
  - Oxford Nanopore single molecule sequencing
- **Molecular Video Lecture 10: NGS Informatics, Analysis**
  - Computation/storage infrastructure to support clinical NGS
  - FastQ format and Phred score
  - Sequence read duplicates in shearing-based libraries vs. PCR amplicon sequencing
  - Demultiplexing, trimming
  - De novo sequencing vs resequencing
  - Sequence alignment overview
  - Variant calling
  - Variant annotation

Course Summary:

Date	Details