Bioinformatics Tools in Drug Discovery:
Course Outline/Curriculum and Timetable

The four corner stones of bioinformatics are 1) a knowledge of biology (more specifically, molecular biology and genetics) 2) programming 3) databases and 4) bioinformatics-specific tools. It is the fourth component of the bioinformaticians' toolkit, which is the focus of this course.

When you have completed this course, you will have learned approaches and skills that will allow you to analyze DNA and protein sequences with a wide variety of tools. More specifically, you will:

- understand the theory behind sequence alignment and searching and become adept at performing both pair wise and multiple sequence alignment
- be able to run and interpret DNA and protein BLAST with more than just the default parameters
- be capable of applying different approaches to gene prediction
- learn about selected publicly available databases and tools for 2D and 3D protein structure, function prediction and analysis
- access the genome database Ensembl through more than just the web interface

Students are expected to have a background in molecular biology or genetics and while UNIX experience is not essential, it would be useful. Students who have not used UNIX will complete an additional skills training session.

The course will be of most benefit to those who wish to move beyond web-page clicking bioinformatics to gain a better understanding of which programs to use when and why. The course will provide a strong foundation for those interested in exploring more advanced bioinformatics topics.

Course Outline/Curriculum and Timetable

Day 1

0900-1000 Lecture 1: Theory of Sequence Alignments
  - Dot Plots
  - Dynamic Programming
  - FASTA alignment algorithm

1000–1100 Lab 1: Database Similarity Searching - an Introduction to FASTA and BLAST

1100-1200 Lecture 2: BLAST
  - Extreme value distribution
  - Expect values
  - Algorithm

- Segment pairs
- BLAST vs. FASTA
- Parameters
- Matrixes: PAM vs. BLOSUM
- BLAST programs compared
- Optimizing BLAST

1200-1300 Lunch
1300-1400 Lab 2: Standalone (Local) BLAST
1400-1500 Lecture 3: Gene Prediction for Identifying Novel Targets

- Markov models
- Hidden Markov Models
- Prokaryotic tools
- Eukaryotic tools

1500-1600 Lab 3: Target Hunting with GenScan, GenomeScan, and HMMGene
1600-1700 Free lab time to complete assignments

Day 2

0900-1000 Lecture 4: Multiple Sequence Alignment and Molecular Evolution

- Darwin’s theory as a foundation for bioinformatics
- Evolution and characters
- MSA overview
- Tools for MSA

1000-1100 Lab 4 : Multiple Sequence Alignments
1100–1200 Lecture 5: Protein Motifs and Structure Prediction

- Protein Data Bank
- Structure Visualization
- Viewing structures: Rasmol, CN3D
- Homology modeling
- VAST

1200-1300 Lunch
1300-1400 Lab 5 - Proteomics Using the ExPASy Site to Characterize Targets
1400-1500 Lecture 6: Open Source Bioinformatics for High Throughput Target Identification

  - BioPerl
  - Ensembl

1500- 1600 Lab 6 Perl for Bioinformatics
1600-1700 Free lab time to complete assignments

DAY 3: Full-Day of Lab Time