• Three-semester introductory sequence
  – Cell biology, genetics, and biochemistry
  – Wet labs first two semesters
  – Third semester bioinformatics lab (Bio 3055)

• Bio 3055 bioinformatics lab
  – Computer lab
  – 5 weeks
2960 (Cell Biology) Laboratory Projects

- Semester-long experiment to identify an unknown bacterium using two different methods
- Series of labs to introduce molecular biology methods
- Computer tutorials on DNA and protein structure (Protein Explorer-based)
- Exploration of the yeast genome through cloning and sequencing

Identifying Unknown Bacteria

- Biochemical Method
  - Gram-positive vs. Gram-negative
  - Oxidase positive vs. oxidase-negative
  - Enterotube or Oxiferm tube

- Molecular Method
  - Isolate DNA from the unknown
  - PCR amplify, then sequence the 16S rRNA gene
  - Use BLAST to find an organism name that is a high-scoring match to the sequence obtained
Exploration of the Yeast Genome

- Students clone an EcoRI-produced fragment of the yeast genome into \textit{E. coli} using a pBS plasmid with Amp resistance
- Plasmid is isolated from Amp resistant colonies. Then,
  - plasmid is digested with EcoRI and run on a gel to identify the insert size
  - plasmid insert is sequenced
- Insert sequence is examined using the yeast genome website \texttt{(www.yeastgenome.org)}
  - Students determine if they obtained a gene, regulatory sequence, or repetitive element
  - Students answer questions about the yeast genome size, number of genes, etc.

2970 (Genetics) Laboratory Projects

- Semester-long experiment exploring artificial selection in \textit{Brassica rapa}
- Examination of several \textit{Drosophila} crosses
- Preparation and examination of polytene chromosomes
- DNA fingerprinting - PCR amplification of D1S80 locus from cheek cells (D1S80 locus)
- Examination of gene flow in grasshopper populations in the Ozarks
- Investigation of HIV sequences using HIV Problem Space at BioQuest
Investigating Gene Flow in Grasshopper Populations

- Grasshoppers are collected each year from the Missouri Ozarks
- Students isolate mitochondrial DNA from the grasshopper hind legs
- Students PCR-amplify a previously identified variable region, then sequence this region
- Using a statistical method learned in lecture, students conclude whether gene flow or genetic drift is dominating

3050 (Biochemistry) Bioinformatics Laboratory

Each section meets 2 hours/week for 5 weeks:
- **Week 1** - Tutorial
- **Week 2** - LocusLink, BLAST, ClustalW
- **Week 3** - DeepView
- **Week 4** - OMIM, KEGG, PubMed
- **Week 5** - Final presentations

Each section has:
- 2 instructors,
- 20 computers,
- 20 or fewer students
Bio 3055 Website:

http://www.nslc.wustl.edu/courses/Bio3055/bio3055.html
All course materials available on website - no password needed

Links to websites

Glossary - bioinformatics terms and abbreviations

Tutorial
GenBank, SwissProt, LocusLink, ClustalW, Swiss Pdb Viewer

Projects - manual sequences readings

---

Projects

- Lung Cancer
  - K-Ras
  - P450 Enzyme
- Apoptosis
  - Caspase I
  - Superoxide Dismutase I
- Metabolic Diseases
  - Phenylalanine Hydroxylase
  - HPRT-Hypoxanthine-guanine phosphoribosyltransferase
- Cholesterol Homeostasis
  - HMG-CoA Reductase
  - Low Density Lipoprotein Receptor
- Mitochondrial Diseases
  - Cytochrome c Oxidase I
  - ATP Synthase 6

---

Projects

- 250 students
  - 15 - 20 students per section
  - 10 projects
  - 2 students on each project per section

- Projects divided into 5 groups by cellular process
  - Same background reading
  - Present to each other at last lab session
Week 1 - Tutorial

All students worked on the same protein, cyclooxygenase-2, on the first day. This enzyme was a topic in lecture and gave the students a chance to be introduced to the basic tools before starting in on their own projects. Bioinformatics tools used in the tutorial:

- NCBI Homepage
- Entrez Gene
- SwissProt
- Sequence Manipulation Suite
- ClustalW
**Week 2 - Start Project Research**

- Translate mutated sequence (cDNA on website)
- Study entries in **Entrez Gene and SwissProt**
- **BLAST search** to obtain list of homologous sequences
- Align all sequences (including mutated) using **ClustalW**
- Identify mutation
- Map secondary structure predictions (**PSIPRED**)

---

**Glu286Lys**

<table>
<thead>
<tr>
<th>Species</th>
<th>Sequence</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>RPVAGLLSSRDFLGGLAFRVFHTQYIRHGSKPMHYTLPEPDICHELLGHVP</td>
<td>292</td>
</tr>
<tr>
<td>Mutant</td>
<td>RPVAGLLSSRDFLGGLAFRVFHTQYIRHGSKPMHYTLPEPDICHELLGHVP</td>
<td>292</td>
</tr>
<tr>
<td>Mouse</td>
<td>RPVAGLLSSRDFLGGLAFRVFHTQYIRHGSKPMHYTLPEPDICHELLGHVP</td>
<td>292</td>
</tr>
<tr>
<td>Zebrafish</td>
<td>RPVAGLLSSRDFLAGLAFRVFHSQYIRHGSKPMHYTLPEPDICHELGHVP</td>
<td>290</td>
</tr>
<tr>
<td>D. melanogaster</td>
<td>RPVAGLSSRDFLAGLAFRVFHSQYIRHGSKPMHYTLPEPDICHELGHVP</td>
<td>291</td>
</tr>
<tr>
<td>G. cydonium</td>
<td>RPVAGLSSRDFLAGLAFRVFHSQYIRHGSKPMHYTLPEPDICHELGHVP</td>
<td>291</td>
</tr>
<tr>
<td>B. floridae</td>
<td>RPVAGLSSRDFLAGLAFRVFHSQYIRHGSKPMHYTLPEPDICHELGHVP</td>
<td>277</td>
</tr>
</tbody>
</table>

Identify regions of high conservation

---

**Week 3 - Structure Investigation**

- Download pdb file from **Protein Data Bank**, and open pdb file in **DeepView (Swiss-Pdb Viewer)**

- Investigate the protein structure
  - View protein in ribbon form
  - Check secondary structure predictions
  - Show mutated residue side chain in cpk color mode
Predict effect of mutation on structure and function

Wild type residue

Mutant residue

Week 4 - Biological Impact of Mutation

- OMIM database: “Allelic variants” - PubMed articles
- KEGG database
- Examine hypothesis in conjunction with biological information
- Write a 1-page summary on findings from project
Week 5 - Final Reports:
Small group presentations
Joint quiz

Labs available on Elgin website:
[www.nscl.wustl.edu/elgin/genomics]

- Scroll down main page to see "Bioinformatics Lab"
- Click on lab titles to download labs as .pdf files
- Instructor and Student pages included

Bioinformatics Lab

This laboratory exploration, Bio436, challenges students with an inquiry-based research project using web-based bioinformatics tools. This 12-hour computer lab curriculum was developed to accompany a large lecture course in introductory biochemistry, but could be easily adapted to any course size. The students work on their own or in small groups and present their results to each other in small groups on the last day of lab.

Washington University biology students perform several experiments in the introductory lab course in which a critical component is generating and analyzing DNA sequence data. Most of these labs are available to download as pdf files and are described below.

- Investigating a Drosophila Genomic Fragment of yeast: U2016 This lab provides an opportunity for students to randomly clone a fragment of the yeast genome into E. coli, then investigate what part of the genome they have cloned through sequence analysis.

- Identifying Unknown Bacteria Using Biochemical and Molecular Methods: Students use two different protocols, one based on sequence analysis and the other on biochemistry, to determine the identity of an unknown bacterial strain.

- Investigating Gene Flow in Grasshopper Populations: Students use grasshoppers collected from sites in the Midwest to generate and analyze sequence data to determine if the grasshopper populations are becoming fragmented due to the suppression of habitat loss. Sequence data is provided here to students who do not have access to grasshoppers can still perform the sequence analysis portion of the lab.