

**Class:**            **IST 444**        Bioinformatics  
                  **BSC 483**        Bioinformatics  
                  **BSC 583**        Bioinformatics

**Instructor:**     Dr. Elizabeth Murray  
                  Associate Professor, Integrated Science and Technology  
                  104 Morrow (office)  
                  Science 108 (Cell Central Lab Suite)  
                  [murraye@marshall.edu](mailto:murraye@marshall.edu)

**Office Hours:**   Science 108 7:00-9:30 AM MWF  
                  Morrow Library 104 11-11:50 AM MWF  
                  Science 108 Friday 2:00-4:00 PM  
                  I strongly encourage students to meet me during office hours or contact me via e-mail for discussion about the software. I will be in the Cell Central Lab Area every morning so give me a shout if you don't see me right in 108- I might be back near computers. I have a chalk board there and will leave a location and sign-in so you will know I am in there.

**Class Time:**     Class meets MWF 10-10:50 AM

**Class Location**   Morrow Library 119 computer classroom.

**Format:**         Lectures and hands-on computer time

**Text:**            *Bioinformatics and Functional Genomics*  
                  Jonathan Pevsner  
                  Wiley-Liss 2003  
                  ISBN 0-471-21004-8  
                  I will also have handouts and exercises from Accelrys Workshop and GeneSpring Workshop

**Class Policies:** Attendance and participation is strongly encouraged and will be considered in assigning the final grade. Since we have access to computers during lectures, I will have short in-class demonstrations of software. Some classes will be devoted to running software tutorials or completing assignments. The university attendance policy (p121 online catalog) will apply for excused absences. Please let me know about planned excused absences.

**Computer use should be focused on class work during class time.** I expect that there be a minimum of random surfing, reading/writing of e-mails, playing games or instant messaging during class time. This is distracting to others in the class. Following along with the course slides, performing tutorials and checking out other bioinformatics sites is considered appropriate use of the computer in class time.

**Performing over-due homework for this class is a marginal use of class-time.** Infractions will be reflected in loss of participation points.

Learning disability will be accommodated. Please let me know so I can make appropriate arrangements.

WebCT Vista will be used to post lecture notes and assignments. Additional links and readings will also be posted to WebCT Vista. Many assignments will provide students with individual sequence(s) to analyze using the same software packages. Students may help each other with software difficulties in answering these problems, but they should write independent analyses of the output from the software or to homework questions. While students are encouraged to work with any available resources (friends, classmates, other professors, tutorials) to create their web page or in posting their work to their web page, **the written content of the web page should reflect the creative and intellectual effort of the student.** Material posted to the Web Page from other sources should have appropriate citations or links. **Finally, please, turn off your cell phone during class- this is disruptive to others.**

**Grades:**

Homework/Projects 400 points

You can't learn this information except by doing problems and interpreting the results. Homework will consist of programs to run on your gene of choice. You will answer some basic questions to aid in interpreting results. Work will be posted to the web page you create and maintain. You must provide me with a link to the webpage and I will check that link for grading. If you are struggling getting work posted BUT it is completed, it can be shown to me in hard copy on the due date with a written interpretation. We can then work on the web page issues after the due date. The Web page will be assessed 4 times during the semester and grade returned within 2 weeks. I will announce the dates and which assignments are to be discussed. Please feel free to show me your web progress as the semester rolls along. *Late work will be accepted but will lose 10% for every week it is overdue. You must tell me work is posted if you want to take advantage of this option.*

Quizzes 300 points

There will be 8 quizzes (worth 50 points each) with about 7 objective questions (true/false, multiple choice) and one or two short answer questions. **You will be permitted to drop up to two quizzes. Therefore, there will therefore be NO make-up quizzes.** Quizzes will be announced in advance and will be on the computer during class. Quizzes will include information discussed in lecture **and in the book.** Quiz grades will be returned within a week.

Participation/Attendance/Web Page Quality and Organization 100 points

Adding interesting information, links and graphics to the Web Page can assist students in bringing up grades. *A coherent discussion of your gene of interest and introduction is critical to receiving a good grade on this portion of the course.*

Final Exam 200 points

This comprehensive exam will be based on the earlier 8 quizzes and homework. This exam will take place during the time scheduled. It will be 50% objective questions and 50% short answer.

**BSC 583 Students:**

**As a graduate student, more work is expected.** The 1000 points discussed above to be 80% of your grade. You will have an additional assignment of a short research paper (8-10 pages) or a teaching module discussing an aspect of Bioinformatics in more detail worth 200 additional points. You must submit a one page proposal for the assignment by October 3. This assignment can be one that is useful for your thesis or research, or an analysis of real data.

<b>Grade Scale:</b>	900-1000 points:	A
	800-899 points:	B
	700-799 points:	C
	600-699 points:	D
	<600 points:	F

**Student Outcomes:**

**The student will:**

- Explore the role of bioinformatics and functional genomics in modern research laboratory.
- Develop a Webpage to hold their Bioinformatics portfolio of output and interpretations of significance of results for a single gene they select with the instructor.
- Learn to interpret and edit DNA sequence raw files using various sequence editors. Learn sequence editing to enter and modify sequences in software packages as well as view and edit multiple sequence alignments.
- Search PubMed and other literature databases for on-line journal articles, abstracts and texts.
- Become familiar with these the major DNA and protein sequence databases (GenBank, European Molecular Biology Laboratory (EMBL) and the DNA Database of Japan (DDBJ) and their file structure and organization.
- Master Entrez, the text-based search and retrieval system used at NCBI for the major databases, including PubMed, Nucleotide and Protein Sequences, Protein Structures, Complete Genomes, Taxonomy, and others.
- Learn to enter sequence data into the databases and search for and download data from the databases.
- Learn to use some of the additional bioinformatics tools used to organize these databases (LocusLink, COGs, OMIM, SNP, UniGene, KEGG and others).
- Become familiar with different strategies for genome sequencing projects and analyze and assemble electronic DNA sequence files using the GCG Wisconsin Package programs.
- Understand the dynamic programming algorithm and its application.
- Be able to distinguish between global and local alignment algorithms and apply them appropriately. Understand the significance and applications of the Needleman/Wunsch and Smith/ Waterman algorithms.
- Understand the role of scoring matrices (PAM and BLOSUM) and gaps in sequence alignment.
- Learn to use, display and interpret data generated by the Pairwise and Multiple alignment programs in the Accelrys Package.
- Master the popular search tools BLAST and FASTA and learn to interpret the significance of the statistics and output from these programs.
- Collect a set of related DNA sequences and calculate phylogenetic distances and create a tree using software programs in GCG Wisconsin package (PAUPSearch; PAUPDisplay; GrowTree; Diverge ).
- Investigate the relationships within an aligned set of protein sequences through computation of the pairwise distance, construction of phylogenetic trees, or calculation the degree of divergence of two protein coding regions.
- View 3-D DNA and protein structures using Chime, Cn3d, Mage, or Rasmol.
- Analyze proteins to identify sequence motifs, predict peptide properties, secondary structure, hydrophobicity, and antigenicity, and identify repeats and regions of low complexity.
- Design and evaluate primer pairs for PCR using software.
- Learn the applications of microarrays and software applications to analyze microarray data.

**Tentative Class Schedule and Readings:**

Assignments will be given specific deadlines during the class. There will be four deadlines for review of the web page and assignments, each worth 100 points. These deadlines are September 7, October 12, November 9 and November 30.

<b>Week</b>	<b>Topic</b>	<b>Reading</b>	<b>Assignment</b>
<b>August 22</b>	Overview of Bioinformatics, PubMed, and NCBI	<b>Chapter 1,2</b>	Select Gene of Interest Begin Web Page Design PubMed Assignment Obtain Accelrys passwords
<b>August 29</b>	DNA Cloning, Sequencing and Genomics Projects	<b>Chapter 12 Accelrys handouts</b>	Post DNA and Protein Sequences for Gene of Interest to Web Page and answer questions. Practice chromatograms.
<b>September 5</b>	Pairwise Sequence Alignment	<b>Chapter 3</b>	Practice Pairwise Alignment Matrices. <b>REVIEW</b>
<b>September 12</b>	BLAST	<b>Chapter 4</b>	First BLAST of your gene.  Pairwise alignment with your gene and early BLAST hit.
<b>September 19</b>	Advanced BLAST	<b>Chapter 5</b>	Protein BLAST and PSI Blast of your gene. Identify a group of sequences to be used with your gene of interest in multiple sequence analysis and later comparisons.
<b>September 26</b>	Bioinformatics and Gene Expression	<b>Chapter 6</b>	Plant microarray assignment. Find and EST or cDNA of your gene. <b>REVIEW</b>
<b>October 3</b>	Microarray Data Analysis. Practice with Genespring.	<b>Chapter 7</b>	Find your gene in a microarray and interpret results.
<b>October 10</b>	Protein Analysis and Proteomics	<b>Chapter 8</b>	Compare primary, secondary and tertiary structure of your protein.
<b>October 17</b>	Analysis of Domains and Blocks and	<b>Chapter 8</b>	Analyze Domains and motifs in protein databases such as SCOP, CATH, Blocks, Interpro.
<b>October 24</b>	Protein Structure	<b>Chapter 9</b>	Find your gene in PDB and post it in CHIME to the web page. Discuss its structure/function relationship. If possible, compare its structure with a related protein. <b>REVIEW</b>
<b>October 31</b>	Multiple Sequence Alignment	<b>Chapter 10</b>	Align sequences related to your gene of interest with Pilup, Pretty, and ClustalW
<b>November 7</b>	Molecular Phylogeny and evolution	<b>Chapter 11</b>	Grow Tree, Paup or other tree programs
<b>November 14</b>	Designing PCR Primer	<b>Chapter 12</b>	Design a PCR primer for your gene
<b>November 21</b>	Thanksgiving-catch up	<b>Chapter 14</b>	Find your gene in its genomic context
<b>November 30</b>	Whole Genome Analysis	<b>Chapter 15</b>	<b>REVIEW</b>
<b>December 5</b>	<b>Exam Week</b>		