SUMMARY

This course will introduce students to the fundamental problems of comparative genomics and molecular evolution including rates and pattern of mutation, drivers of genome organization, population effects on allele dynamics, phylogenetic inference, ancestral sequence reconstruction, cis-regulatory evolution, functional protein evolution, birth of de novo genes and the detection of natural selection and selective sweeps. Students will also be introduced to DNA sequence databases, the PDB protein data bank, and minimal use of Perl scripting. Students will use existing evolutionary genomics and protein structural modeling software to conduct novel comparative genomic research on a protein coding region of their own choosing. (Prereq-BIOL 201 or BLOG 140). This course will be partly activity based.

TEXTBOOK / MATERIALS

The textbook for the course is
Molecular and Genome Evolution by Dan Gruar (Sinauer Associates Sunderland MA)
ISBN 9781605354699

Throughout the course we will also use two freely available software packages MEGA 7.0 and Chimera 1.1 and the Perl computer language.
MEGA 7.0  http://www.megasoftware.net/

Chimera 1.11  https://www.cgl.ucsf.edu/chimera/download.html

Perl for Windows  (Perl is generally preinstalled on Mac OsX and Linux)  
http://www.activestate.com/activeperl/downloads

Recommended Programming Editor
Komodo Edit 8.0 (later versions are not free)

TEACHING PHILOSOPHY

The mission of the modern university is to produce and disseminate new knowledge to the world. As an instrument towards this goal, I am an active researcher in the areas of molecular evolution and computational structural biology. However, I enjoy my role as a teacher here at RIT just as much as my research. I strive to do both of these tasks well. I ask each of you to take responsibility for the continued development of your own education. We are always here to help you, but ultimately the path of our learning is walked alone in each of our own minds. You must be actively committed towards this personal goal for yourself if you want to get a good grade in this class. Self-study is very important. Simply showing up without actively engaging your mind will not let you achieve success. Please communicate with me early in the semester if you feel you are having trouble understanding any concepts or material. All of us, PhD’s included, have gaps in our understanding of various subjects. There is no such thing as a "dumb" question in my class. Success in academics at every level is often simply a matter of recognizing where missing information is holding us back. As an instructor, I will always try my best to achieve clarity by communicating to the least informed member of my audience, but I may miss the mark at times. During the lecture or lab, please don't feel that you cannot interrupt to ask me questions. While I am with you, my time is always your time.

Unlike some of the other natural sciences which are more mathematically-based, learning concepts in biology is often rooted in specialized vocabulary. I highly recommend that you keep a vocabulary list while reading prior to class and attending lecture.

ATTENDANCE

Welcome to college, where YOU the student are ultimately paying for my attendance in class. Your attendance at lectures and/or studio sessions is hoped for, and probably required for a good grade but not monitored. There will always be information presented that is not otherwise available, i.e. it will not be in the books, and will be needed in projects, exams and/or quizzes. It is a good idea to be proactive if you are going to miss a class, and contact me to discuss material that may be missed.

GRADES

Grades will be based upon 10 assignments, 3 exams and a short end-term report on an analysis of your own gene of interest

HONESTY POLICY

PLEASE REVIEW THE RIT POLICY ON ACADEMIC DISHONESTY AND STUDENT MISCONDUCT AT THE WEBSITE BELOW

http://www.rit.edu/studentaffairs/studentconduct/rr_academicdishonesty.php

WE WILL ABIDE BY THESE POLICIES AND RECOMMENDED PENALTIES...NO EXCEPTIONS.

GRADE DISTRIBUTION -

Note: grades are scaled as follows
A 93–100 | A- 90–93 | B+ 86–90 | B 83-86 | B- 80 – 83 | C+ 76 – 80 | C 73 – 76 | C- 70 – 73 | D 60-70 | F <60
LECTURE AND STUDIO SCHEDULE ('studio' sessions will require laptop computer)

Molecular Basis of Evolution

Week 1 (chapter 1, paper by Babbitt 2011 on Chromatin, paper by Cobb 2013 on Genes)
A) Molecular evolution from three perspectives (sequence, structure and the dynamics of chromosome)
B) Classifications of chemical mutation - the engine of evolution
C) Basics of evolutionary process – selection and drift

Week 2 (paper by Hayes 1998 on Genetic Code)
A) Structure and conceptualization of the gene
B) Population and genetic variation across the genome
C) (studio) allele dynamics models on MS Excel

Week 3 (chapter 2)
A) Standing variation, recombination and linkage
B) Some details – linkage disequilibrium, genetic hitch-hiking, background selection and selective sweeps
C) (studio) intro to MEGA – BLAST search, sequence download and NCBI

Molecular Phylogenetics

Week 4 (chapter 3 p93-106, see also alignment algorithms in MS Excel)
A) Alignment and molecular homology
B) (studio) intro to MEGA – sequence homology, alignment, and putative orthologs (PSI-BLAST)
C) Exam I

Week 5 (chapter 5 p165-211, chapter 6 p237-243, p257-268)
A) Inferring tree topology via parsimony, evolutionary distance, likelihood, and Bayesian method
B) More details - ancestral sequence reconstructions, codon-based methods, reticulate evolution
C) (studio) tree methods in MEGA

Week 6 (Chapter 3 p79-93, Chapter 5 p207-209, papers on Markov Chain and Bootstrap)
A) Basics of statistical inference and probability applied to evolution
B) Markov chain, Monte Carlo methods, Efron’s bootstrap, empirical p-values
C) (studio) more background on transition matrix models and fitting likelihoods

Functional Evolution of Genome Organization

Week 7 (chapter 10-11, paper by Neme and Tautz 2014 on birth of new genes)
A) Genome diversity – bacteria, Archaea, ciliates, algae, fungi, plants, animals
B) Defining genome complexity – the problem of duplication or de novo genes
C) (studio) measuring mutational tolerance

Week 8 (chapter 12 and papers by Wray in 2003 and 2007 on importance of cis-regulatory evolution)
A) Gene regulatory evolution – cis regulatory motifs and gene expression
B) Interactions within genomes – network topology, robustness, evolvability
C) (studio) motif searching with regular expressions using Perl scripting and PHI-BLAST (studio)
Methods of Detecting Natural Selection

Week 9  (Chapter 4 p 107-139 and paper by Suzuki 2010 on detecting selection)
A) Relative rate and likelihood-based methods
B) Site-specific methods, lineage-comparative methods and mutational tolerance
C) (studio) MEGA – z-test and HyPhy

Week 10  (papers by Hahn et al. 2003 MBE and Zhen&Andolfatto 2012 selection on non-coding DNA)
A) Empirical p-value derived via simulation of neutral evolutionary models
B) Maximal-order Markov methods applied to regulatory motif
C) Exam II

Beyond Sequence-based Bioinformatics

Week 11  (chapter 4 p139-160 and papers by Babbitt et al. 2016 Chouard 2013 Parker&Tullius 2011)
A) Modern mysteries – c-value paradox, selection on syn sites, MA line mutations, ‘missing’ heritability
B) New perspectives – structure-function paradigm, role of regulatory evolution, heritable epigenetics
C) (studio) Chimera and PDB - introduction

Week 12
A) The roots of gene function - molecular bonding/structure, electrostatics, intermolecular forces
B) Molecular modeling – intro to quantum mechanics, molecular mechanics, and molecular dynamics
C) Evolving physics and the boundary between living and non-living matter – potential applications of molecular modeling to molecular evolution

Week 13
A) (studio) Chimera and PDB – surface analysis and molecular dynamics simulation
B) open studio
C) open studio (take-home essay exam III due)

Project Work

Week 14
A) Introduction to scientific report writing and open studio
B) open studio
C) open studio

Week 15
A) End-term papers due