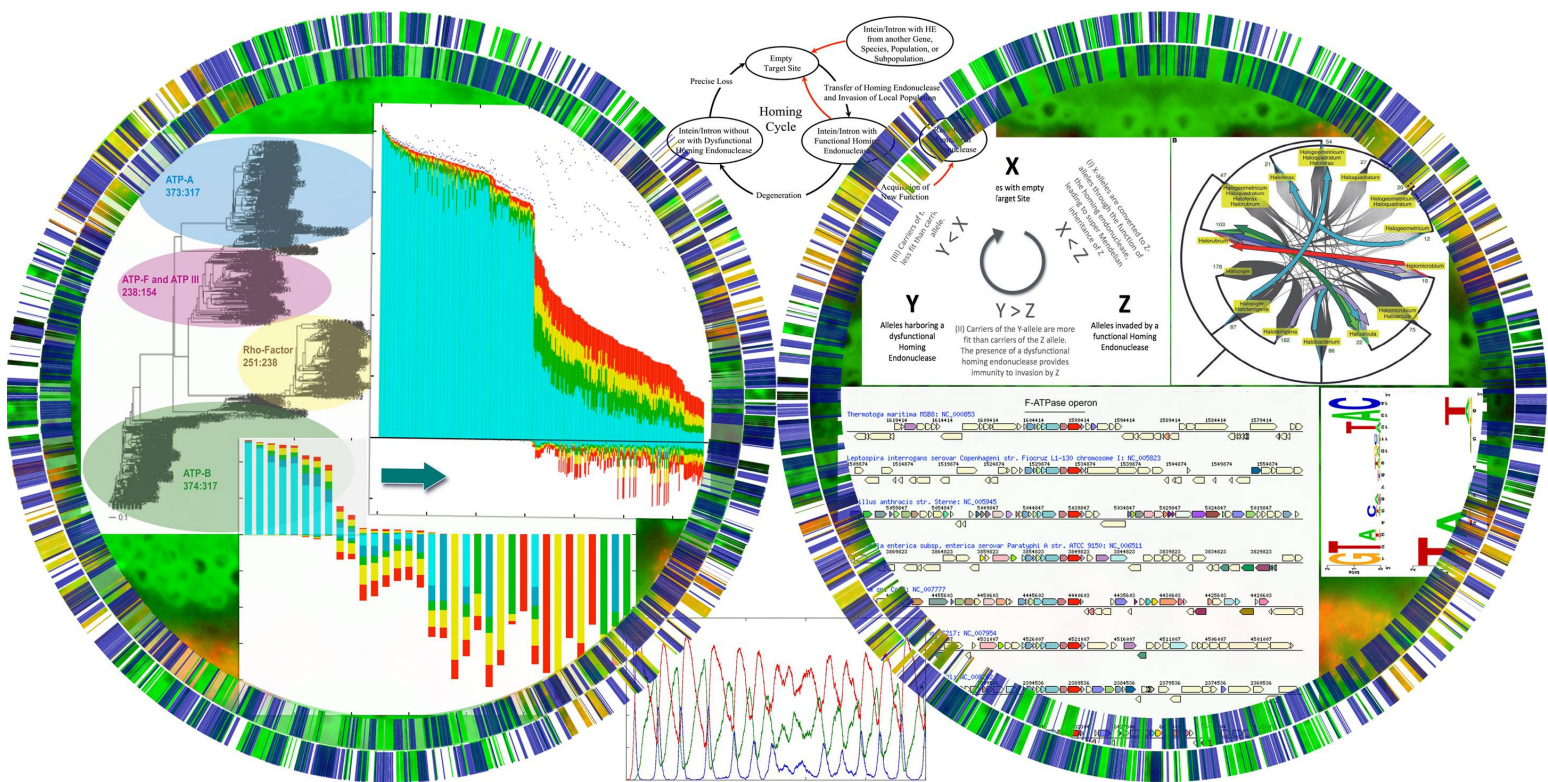Catalog Copy:

3421. Introduction to Molecular Evolution and Bioinformatics

4.00 credits

Prerequisites: Recommended preparation: At least one 2000 level course in MCB.

Grading Basis: Graded

Evolution of biomolecules and application to molecular data analysis and the design of new molecules. Topics include prebiotic chemistry, origin of cells, selfish genes, molecular innovations, data bank searches, alignment of sequence and 3-D protein structures. Course includes lectures, discussions, computer lab exercises, and an independent research project.

Syllabus

MCB 3421: Introduction to Molecular Evolution and Bioinformatics

Prerequisites: Recommended preparation: At least one 2000 level course in MCB.
**Grading Basis:** Graded

Content: Evolution of biomolecules and application to molecular data analysis and the design of new molecules. Topics include prebiotic chemistry, origin of cells, selfish genes, molecular innovations, data bank searches, alignment of sequence and 3-D protein structures.
Course includes lectures, discussions, computer lab exercises and an independent research project.

Mondays & Wednesdays 11.15AM-12.05PM, TLS 301 (lectures)
Fridays 10:10AM - 11:40AM TLS 411 (computer-lab exercises and discussions)
Wednesdays 2:00pm – 5:00pm Open Lab, TLS 411 (independent research project)
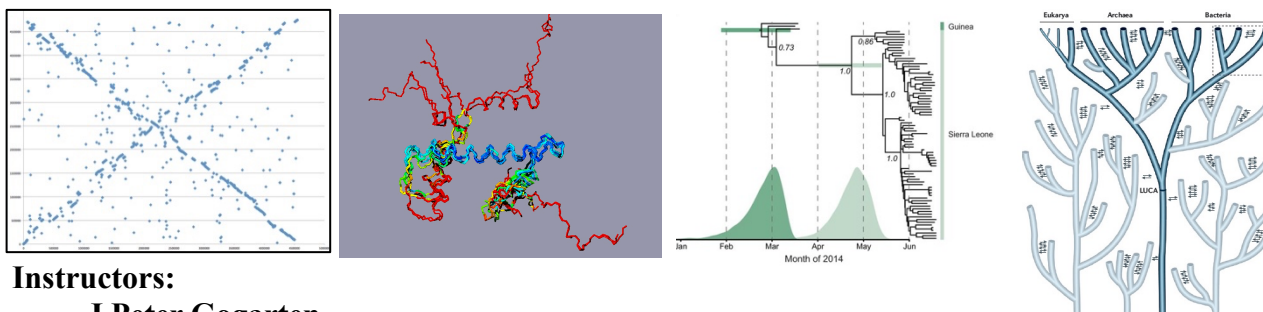
Instructors:

J Peter Gogarten
Dept. Molecular and Cell Biology
University of Connecticut
Storrs, CT 06269-3125
Email: gogarten@uconn.edu ; Phone: 860 465 6267,
https://uconn-cmr.webex.com/meet/jpg02006 , BP 404

Daniel S. Phillips
Email: daniel.s.phillips@uconn.edu
Webex https://uconn-cmr.webex.com/meet/dsp19002

Office hours
> **JPG:** MW after class, or via webex after arrangement: for fast response send an email!
> **DSP** Via Microsoft teams or webex by arrangement, or during open lab hours
> For questions of general interest, use the bulletin board on huskyCT.

Basis for grading:
> **Participation** (huskyct discussion board, in class discussions, in class pop-up quizzes),
> **Computer lab worksheets**,
> **Notebook (1 reflective statement per lecture, detailed notes on labs and research project)**
> **Take-home exams**,
> **Midterm**,
> **Report on conducted research.**

Syllabus

Weights:
25% Report on conducted research
20% Midterm
10% Notebook (detailed notes on each computer lab; 1-3 sentences per lecture; notes on research project)
10% Computer Lab worksheets
10% Participation, bulletin board postings, pop-up quizzes, other in-class assignments:
25% Take-home assignments:
You can **drop the worst 4** grades from the Take Home Assignments – at least 8 will be offered

Grading Scale:

Letter Grade	If your overall percentage is greater than or equal to this	and your overall percentage is less than this
A	93	—
A-	90	93
B+	87	90
B	83	87
B-	80	83
C+	77	80
C	73	77
C-	70	73
D+	67	70
D	63	67
D-	60	63
F	0	60

Expectation: More than 5h reading/studying per week (most will do fine with about 4-6h/week).

Classroom Etiquette
Cell Phones: Cell phone use is only permitted for data recording (pictures, notes, etc) and is never permitted for texting, calls, apps not related to the class, or email.
Computer Use: Eating or drinking is not allowed in th computer lab.
Mask wearing: Please do not come to class if you are sick. If you have a signs of a cold, please wear a mask.

Our Philosophy
We believe in having a dynamic classroom, open to discussion, participation, and inquiry from all members of the class. We strive to establish a fair and balanced classroom where all students can participate. We will try out team based learning in the Friday sessions.

Learning objectives:
By the end of the semester students will be able to
- visualize, analyze, and align protein sequences and structures.
- understand and apply the homology concept to nucleotide and amino acid sequences.
- judge the significance of similarity between sequences.
- perform database searches using NCBI's entrez interface
- connect to the bioinformatics cluster using sftp and ssh
- execute repetitive databank searches using the command line
- create searchable databanks from sequences of interest

Syllabus

- understand the advantages and disadvantages of different approaches to sequence alignment
- perform dotplot analyses to compare genes and genomes and identify recombination events
- reconstruct the evolutionary history of molecules from sequence data
- understand different approaches to taxonomy – appreciate the book title "Why Fish Don't Exist".
- discuss and recognize the different forces acting in sequence evolution
- discuss the role of vertical and horizontal inheritance in evolution
- explain the concept of constructive neutral evolution
- discuss the roles and contributions of introns, inteins and other genetic elements that function as molecular parasites in today's organisms
- keep clear and informative electronic records of computational analyses.

Academic Honesty

Academic misconduct is dishonest or unethical academic behavior that includes, but is not limited to: misrepresenting mastery in an academic area (e.g., cheating), intentionally or knowingly failing to properly credit information, research or ideas to their rightful originators or representing such information, research or ideas as your own (e.g., plagiarism). Examples of misconduct in this class include, but are not limited to: cheating on exams, plagiarism, turning in questions for fellow students, impersonating another student, falsifying data, copying, fabricating or stealing data. For more details on the University of Connecticut's policy on academic integrity, including the instructor's role and procedures you are referred to the following page and references therein: https://community.uconn.edu/the-student-code-appendix-a/

Plagiarism, undisclosed use of AI, and "collaborations" for writing assignments and take-home exams will not be tolerated!
In case of misconduct, **you will receive an F for the assignment, AND the F will count towards the final grade** (*i.e.,* you cannot drop an F received due to academic misconduct). If, to answer a question on a take-home exam, you use information from a website (either verbatim or paraphrased) put in a reference to the website. If you use chatGPT or similar to write a section of an essay, mark the section and put the AI program and your query into the references.

Exam policy:

If you need to reschedule your exam, please let your instructor know ASAP.

The general policy for Syllabi, including links to the expected workload and class attendance is at http://provost.uconn.edu/syllabi-references/

Switch to remote teaching due to covid19 or another outbreak: In case we need to switch to remote teaching due to an outbreak in class, or a worsening overall situation, we will use webex for synchronous lectures (MW, 11.10 am ET, and we will move the computer labs to your personal computers/laptops (10.05 am ET or 2.25 pm ET).

Disability accommodation. The Center for Students with Disabilities (CSD) at UConn provides accommodations and services for qualified students with disabilities. If you have a documented disability for which you wish to request academic accommodations and have not contacted the CSD, please do so as soon as possible. The CSD is located in Wilbur Cross, Room 204 and can be reached at (860) 486-2020 or at csd@uconn.edu. Detailed information regarding the accommodations process is also available on their website at http://www.csd.uconn.edu/.

If you need any additional information on this course, please send an email to gogarten@uconn.edu !

Syllabus

Very Preliminary Schedule

red: computer lab exercises in the TLS 411 (F 10.10 -11.40 or 13.25-14.55);
blue: lectures, demonstrations, and discussions TLS 301 (MW 11.15-12.05)

Class 01: Overview; topics; textbook; reading materials; How will grades be calculated?
Class 02: Homology, basics of protein structure
Computer-lab assignmnet 01: Intro to chimera - binding pocket substrate interactions
Class 03: ATPase structure, function, and evolution; the RNA world concept; discussion of homology
Computer-lab assignment 02: Comparing structures for ancient duplicated genes / following a structure through the catalytic cycle
Class 04: ATP binding sites and convergent evolution,natural selection and definitions of life
Class 05: Divergence of sequences, symbiotic relationships, inteins
Computer-lab assignment 03: Exploring intein structures in chimera
Class 06: The tangled tree of life.
Class 07: Mutationism and the Modern Synthesis ; Mutual Aid, HGT and endosymbiosis
Computer-lab assignment 04: Literature Databanks and Entrez
Class 08: Statistics of sequence comparison
Class 09: Databanks, Blast, Unix
Computer-lab assignment 05: Statistics of Sequence Comparison
Class 10: Dayhoff and the history of bioinformatics
Class 11: Strand-bias and recombination
Computer-lab assignment 06: Blast searches using the command line
Class 12: (A) cladistics (B) early humans
Class 13: Ebola as an example for the usefulness of molecular evolution; review of endosymbiosis in eukaryotes
Computer-lab assignment 07: Creating gene plots
Class 14: Review Session
Midterm
Computer-lab assignment 08: Gene plots, strand bias, and mummer
Class 15: Gene duplication and post mating hybridization barriers, discussion lab 8, sequence alignment
Class 16: Sequence alignment, multidimensional sequence space.
Computerlab assinment 09: Dotlet and Jalview exercises on repetetive motifs, introns and sequence space.
Class 17: Introns late versus introns early
Class 18: Intro to trees and phylogenies
Computerlab assigment 10: Phylogenies, bootstrap, model tests, LBA
Class 19: Discussion of LBA, Why can gene trees be in conflict with one another?
Class 20: selection, deterministic models, genetic drift
Computerlab assignment #11: Simple scripts to handle genome sized data sets
Class 21: Detecting positive and purifying selection
Class 22: dN/dS continued
Computerlab assignmnet #12: Bayesian analysis to determine evolutionary parameters in sequences.
Class 23: Types of selection / GTAs and purifying selection
Class 24: Intro to PSI blast.
Computerlab assignmnet #13: PSI blast exercises
Class 25: Discussion of PSIBlast lab, CNE, early Earth and bottlenecks
Class 26: Discussion of findings from research projects
Computerlab assignmnet #14: Phylogenetic analysis of an intein containing gene
Research project report is due.