

**INF 515 Comparative Genomics Syllabus**

Term	Class No.	Section	Units	Days & Times	Room	Mode
Spring 21	INF 515	001	3	MoWeFri 12:40 PM-1:30 PM	Bldg 81 Rm 345	In-person

**Enrollment Requirements**

None.

**Course Website**

<http://bblearn.nau.edu>

**Instructor**

Dr. Marc Tollis

Email: [Marc.Tollis@nau.edu](mailto:Marc.Tollis@nau.edu) MoWeFri 2:00-2:30

**Course Purpose**

The purpose of this course is to familiarize informatics students with methods and theories behind comparative genomics. Biology is now a big-data science that is genomic in scale, and the growing availability of genome sequence data from across the tree of life challenges researchers to provide evolutionary explanations for this diversity. This course focuses on the analysis of genomic sequences, including sequencing technologies, genome assembly, functional and structural predictions, and comparisons at macroevolutionary timescales in a phylogenetic context. Special emphasis is on building practical skills in computational biology. Classes will consist of lectures, student led discussions of primary literature, and computational labs.

**Course Student Learning Outcomes**

Upon successful completion of this course, students will be able to demonstrate the following competencies:

- **LO1:** Describe and explain technologies and resources necessary for comparative genomics analysis from the onset to the conclusion of a study;
- **LO2:** Apply comparative genomic concepts to solve a biological problem or question;
- **LO3:** Analyze results and differentiate between alternative explanations for the data;
- **LO4:** Weigh evidence against various quality-control measures and statistical uncertainty;
- **LO5:** Apply and communicate the above learning outcomes by designing and completing a project.

**Assignments / Assessments of Course Student Learning Outcomes**

**Activity types:** Each class consists of a lecture/discussion led by the instructor, followed by assigned student-led paper discussions (LO1). Five computational labs will be given and students will prepare a written description of their methods and results (LO2, LO3, LO4). There will be two exams (non-cumulative) to assess the first four learning outcomes. A final research project will be proposed early in the semester and each student will work individually or in groups to carry it out, submit a written assignment, and present it to the class at the end of the semester (LO5).

**Course materials:** See [https://github.com/marctollis/INF515-Comparative-Genomics\\_fall22](https://github.com/marctollis/INF515-Comparative-Genomics_fall22). Additional materials will be discussed in class and shared on BB Learn.

**Paper Discussions:** Every student will get an opportunity to lead a class discussion of one of the assigned readings. Additional readings may be added to the course and uploaded to BBLearn, in order to maintain a balanced ratio of students-to-papers over the semester. Expectations for discussion leadership will be made clear in class. A rubric is available for the assignment on BBLearn.

**INF 515 Comparative Genomics Syllabus**

**Assignments:** Assignments will take the form of computational labs, and a written submission for all labs will be due before class exactly two weeks later, submitted to BBLearn. The format of the written assignments will be discussed in class. Students will be able to resubmit their lab results as they may gain better competency with course concepts over time **BUT ONLY IF THEY SUBMIT A FIRST DRAFT ON TIME**. Significant class time will be spent improving the students' performance and interpretation of the computational labs.

**Quizzes:** Each module that contains a reading will also contain a quiz. The quizzes will be short answer form and will address questions/problems/topics covered in comparative genomics and designed to ensure that students understand and synthesize the readings, lectures, and class discussions. They will be ~5 questions in length and randomized to ensure academic integrity. Quizzes will be made available after the class and are due one week later. Each students' two lowest quiz grades will be dropped so only 10 quizzes will count towards their final grade.

**Final Projects:** The final project consists of multiple parts: (i) a proposal (2%); (ii) a written submission that includes Introduction, Goals, Description of the Data Used, Background, Methods, Results, Discussion, and References (15%); and (iii) a 10-minute powerpoint presentation to be given during the last class of the semester (3%). All parts will be submitted to BBLearn to be graded. A rubric for the final project is available on BBLearn. Students may propose to work in groups. Finer details of this assignment will be discussed in class.

**Grading System**

A weighted sum of assessment components is used to determine your final grade in the course:

Assignment	Weight	Assignment Points (1000 total)
Attendance	10%	100
Class Participation	10%	100
Paper Discussion	10%	100
Computational Labs	30%	300
Project Proposal	2%	20
Quizzes	20%	200
Final Project	18%	180

Grades will be assigned using the weighted sum described above using this scale:

**A** ≥ 90%, **B** ≥ 80%, **C** ≥ 70%, **D** ≥ 60%, **F** < 60%.

There is no "curve". Each student's grade is based on their own outcomes assessments and not affected by the grades of other students. Extra credit opportunities may present themselves throughout the semester and will be announced during class meetings. Mistakes in grading do happen, and students are encouraged to discuss such concerns with the instructor during office hours.

**Readings and Materials**

Available on BBLearn.

INF 515 **Comparative Genomics** Syllabus

**Class Outline and Tentative Schedule**

The course topics and a tentative schedule serve as an outline for the class:

Week	Day	Topic/Reading	Assignments (see BBLearn for submission instructions)
8/29	M	<b>Course Introductions</b>	<b>Quiz 1; sign up for discussion leadership; Software Carpentry; Computing Survey</b>
	M	<b>Sequencing Technologies</b>	
	F	<b>Readings:</b> Shendure et al. 2017. DNA sequencing at 40: past, present and future. <i>Nature</i> .	
9/5	M	<b>Computing in Genomics</b> Command line, introduction to Monsoon at NAU	
	W	<b>Computing in Genomics</b> Command line, introduction to Monsoon at NAU	
	F	<b>Computing in Genomics</b> Command line, introduction to Monsoon at NAU	
9/12	M	<b>Genome Assembly</b>	<b>Quiz 2</b>
	W	<b>Readings:</b> 1. Bradnam et al. 2013. Assemblathon 2: evaluating de novo methods of genome assembly in three vertebrate species. <i>GigaScience</i> . 2. Sedlazeck et al. 2018. Piercing the dark matter: bioinformatics of long-range sequencing and mapping. <i>Nature Reviews Genetics</i> .	
	F	<b>Lab 1:</b> Assembling a genome, due 9/23/21	
9/19	M	<b>Genome Annotation 1</b>	<b>Quiz 3</b>
	W	<b>Readings:</b> Yandell and Ence. 2012. A beginner's guide to eukaryotic genome annotation. <i>Nature Review Genetics</i> .	
	F	<b>Lab 2:</b> RepeatMasking a genome	
9/26	M	<b>Genome Annotation 2</b>	<b>Quiz 4; Lab 1; Lab 2</b>
	W	<b>Readings:</b>	

INF 515 **Comparative Genomics** Syllabus

		<ol style="list-style-type: none"> <li>1. Aken et al. 2016. The Ensembl gene annotation system. <i>Database</i>.</li> <li>2. Holt and Yandell. 2011. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. <i>BMC Bioinformatics</i>.</li> </ol>	
	F	<b>TBD</b>	
10/3	M	<b>Orthology Assignment</b>	<b>Quiz 5; project proposal</b>
	W	<b>Readings:</b> <ol style="list-style-type: none"> <li>1. Kriventseva et al. 2019. OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. <i>Nucleic Acids Research</i>.</li> </ol>	
	F	<b>Lab 3:</b> NCBI, Ensembl, OrthoDB,	
10/10	M	<b>Phylogenomics 1</b>	<b>Quiz 6</b>
	W	<b>Readings:</b> <p>Morales-Briones et al. 2020. Disentangling sources of gene tree discordance in phylogenomics data sets: testing ancient hybridizations in Amaranthaceae s.l. <i>Systematic Biology</i></p>	
	F	<b>Lab 4:</b> Tree building	
10/17	M	<b>Phylogenomics 2</b>	<b>Quiz 7</b>
	W	<b>Readings:</b> <p>Donoghue and Benton. 2007. Rocks and clocks: calibrating the Tree of Life using fossils and molecules. <i>Trends in Ecology and Evolution</i>.</p>	
	F	<b>Lab 5:</b> Divergence time estimation	
10/24	M	<b>Comparative methods</b>	<b>Quiz 8; Lab 3; Lab 4; Lab 5</b>
	W	<b>Readings:</b> <a href="https://lukejharmon.github.io/pcm/">https://lukejharmon.github.io/pcm/</a> <b>chapters 3 and 7</b> <b>TBD</b>	
	F	<b>TBD</b>	
10/31	M	<b>Tests for Selection: Codon Models</b>	<b>Quiz 9</b>
	W	<b>Readings:</b> <p>Sackton 2020. Studying Natural Selection in the Era of Ubiquitous Genomes. <i>Trends in Genetics</i>.</p>	
	F	<b>Lab 6:</b> Detecting positive selection with codon models	
11/7	M	<b>Population Genomics and Selective Sweeps</b>	<b>Quiz 10</b>
	W	<b>Readings:</b>	

**INF 515 Comparative Genomics Syllabus**

		<ol style="list-style-type: none"> <li>Kern and Hahn. 2018. The Neutral Theory in Light of Natural Selection. <i>Molecular Biology and Evolution</i>.</li> <li>Jensen et al. 2019. The importance of the Neutral Theory in 1968 and 50 years on: a response to Kern &amp; Hahn 2018. <i>Evolution</i>.</li> </ol>	
	F	<b>NO CLASS: VETERAN'S DAY</b>	
11/14	M	<b>Evolutionary Conservation</b>	<b>Quiz 11</b>
	W	<b>Readings:</b> Sackton et al. 2019. Convergent regulatory evolution and loss of flight in paleognathous birds. <i>Science</i> .	
	F	<b>Lab 7:</b> UCSC Genome Browser, due 11/30/21	
11/21	M	<b>Ancient DNA and Paleogenomics, Species Conservation and Extinction</b>  <b>Readings:</b> Bergström et al. 2020. Origins and genetic legacy of prehistoric dogs	
	W	<b>Readings:</b> Robinson et al. 2018. Purging of Strongly Deleterious Mutations Explains Long-Term Persistence and Absence of Inbreeding Depression in Island Foxes. <i>Current Biology</i> .	
	F	<b>NO CLASS: THANKSGIVING HOLIDAY</b>	
11/28	M	<b>Projects, discussion of issues with computational labs</b>	<b>Quiz 12; Lab 6</b>
	W	<b>Projects, discussion of issues with computational labs</b>	
	F	<b>Projects, discussion of issues with computational labs</b>	
12/5	M	<b>Projects, discussion of issues with computational labs</b>	<b>Lab 7; final projects reports</b>
	W	<b>Projects, discussion of issues with computational labs</b>	
	F	<b>Projects, discussion of issues with computational labs</b>	

**Course Policies**

The following policies will apply to this course:

- Attendance is taken every class because it is required.
- Cheating and plagiarism are strictly prohibited. All academic integrity violations are treated seriously. All work you submit for grading must be your own. You are encouraged to discuss the intellectual aspects of assignments with other class participants. However, each student is responsible for formulating solutions on their own and in their own words. Academic integrity violations will result in penalties including, but not limited to, a zero on the assignment, a failing grade in the class, or expulsion from NAU.
- Electronic device usage must support learning in the class. All cell phones, PDAs, music players and other entertainment devices must be turned off (or in silent mode) during lecture, and may not be used at any time. Laptops or tablets are allowed for note-taking and research (Googling a topic on demand, for instance).

**INF 515 Comparative Genomics Syllabus**

- Grades will be entered in BBLearn but your final grade will be calculated in Excel using the grading system described above and then entered in LOUIE. Your final course grade will **not** necessarily appear in BBLearn. Please check LOUIE for your final grade.
- Visiting the instructor(s) during office hours is encouraged! I am happy to talk about the class, careers, research, and topics related (even loosely) to this course.