Syllabus for BIOL 475/575: Biodiversity Informatics

Lecture: Wednesdays 2:00pm–4:50pm

Zoom Meeting ID: 924 3442 2002 Passcode: diversity

In-Person: iSELF 204

Instructor: Dr. Althea A. Archer
    Office: 267 Wick Science Building
    320-308-4975 (office)
    Email: althea.archer@stcloudstate.edu
    Twitter: @aaarchmiller
    Virtual Office Hours: Tues & Friday 10:00am–11:00am
    Office Hour Link: https://minnstate.zoom.us/j/98128037816
    Meeting ID: 981 2803 7816 Passcode: Archer

The schedules and policies associated with this course may be subject to revision or change as a consequence of changing circumstances or events. Reasonable notification will be provided to students prior to any major changes in course policies or procedures.

Course Description

Biological collections, collection stewardship, biological collections databases, networks, cybertaxonomy, taxonomic concepts, ontology, specimen digitization, georeferenced specimens, predictive ecogeographic modeling, genomic databases, genomic partitioning strategies, models of molecular evolution, phylogenomics. Prereq.: BIOL 456.

Learning Outcomes

1. Summarize the history of biological collections and their importance to ecological and evolutionary biology

2. Demonstrate good collection stewardship and relevant database information of biological collections

3. Analyze the role of cybertaxonomy in addressing the challenges of modern taxonomy

4. Evaluate methods of specimen digitization and applications of this data for taxonomic, phylogenetic, and evolutionary studies

5. Determine the important of georeferencing biological collections and how to access databases with georeferenced information

6. Apply ecogeographic predictive modeling for ecological and evolutionary studies.

7. Examine genomic databases associated with biological collections.

8. Synthesize genomic partitioning strategies, models of molecular evolution, and fundamentals of phylogenomics.

9. Apply methods and applications of phylogenomic studies.

10. Evaluate and explain the data and conclusions drawn from primary literature, particularly in relation to cybertaxonomy, phylogenomics, and bioinformatics.

Required Textbooks

- Buffalo, Vince. 2015. Bioinformatics Data Skills. O'Reilly publishing

- Each person must sign up for an account with GitHub (free with educational affiliation)

- Recommended: McMillan, V.E. 2012+. *Writing Papers in the Biological Sciences*. Bedford/St. Martin's

CONTACT ME: The best ways to get ahold of me are by visiting my virtual office hours or by emailing me. I will always try to get back to emails within 48 hours. I get a lot of emails, so please begin emails with "BIOL 475" or "BIOL 575" so that I can prioritize your email.

REGULAR ATTENDANCE AND PARTICIPATION IN CLASS IS CRITICAL TO YOUR SUCCESS. This course will be offered in an in-person format with active learning labs required in each class. Each day will begin with introductory lectures followed by tutorial labs on your computer. The lab from each day's work will be graded via D2L (first lab) or GitHub. You will only be able to make up daily labs if you have prior consent.
   In order to have an excused absence, you must notify me prior to the beginning of class of your absence.

ACCOMMODATIONS FOR STUDENTS WITH DISABILITIES: SCSU is an affirmative action, equal opportunity employer and educator. We are committed to a policy of nondiscrimination in employment and education opportunity and work to provide reasonable accommodations for all persons with disabilities. Accommodations are provided on an individualized, as-needed basis, determined through appropriate documentation of need. Please contact Student Accessibility Services (SAS), sas@stcloudstate.edu or 320-308-4080, Centennial Hall 202, to meet and discuss reasonable and appropriate accommodations.

RESPECT FOR DIVERSITY: It is my intent that students from diverse backgrounds and perspectives be well-served by this course, and that the diversity that students bring to this class be viewed as a resource. Please let me know ways to improve the effectiveness of the course for you, personally, or for other students or student groups. As a student in this class, you are required to treat other members of the class with respect and kindness. Diverse perspectives are welcome and disagreeing is fine. However, disrespectful, rude, or exclusive behavior will not be tolerated.

GRADES

Item	Date	Details	points	%
Daily Lab Exercises	Various dates	10 labs x 5pts each	50	50.0%
Journal Club Leader	Sign up for date	With partner	10	10.0%
Journal Club Reader	Various dates	Participation grading	5	5.0%
Data Source Presentation	Feb. 2	With partner	15	15.0%
SDM Abstract	April 20	In groups	5	5.0%
SDM Presentation	April 27 or May 2	In groups	15	15.0%
Total			100	100.0%

DAILY LAB EXERCISES will be completed during class with Program R. Other than the first lab, labs will be handed in by 10pm on each Monday following that class via html documents uploaded to GitHub. You will be graded based on your completion of the day's tasks and your code neatness and readability. For BIOL 575, the grading standards will be held to a higher level than those for BIOL 475.

JOURNAL CLUBS will be held nearly every class period (see schedule) during the final 30 minutes of class.

Percentage	Grade
≥ 99	A+
92-98.9	A
90-91.9	A-
89-89.9	B+
82-88.9	B
80-81.9	B-
79-79.9	C+
72-78.9	C
70-71.9	C-
69-69.9	D+
60-68.9	D
< 60	F

Journal club leaders will be responsible for presenting the main findings of the focal paper as a group and for creating 3-5 questions that will spark discussion. Discussions on the paper will be done in two smaller groups, each lead by one of the Journal Club Leaders for that day. Journal club papers will be chosen by the instructor (for 475 students) or by the leaders (for 575 students) and will be related to that day's lesson topic.

Journal club readers will be responsible for reading the paper ahead of time and coming to class prepared to discuss the paper. To get 100% participation as a journal club reader, you must participate in the discussion during each journal club. BIOL 542 Exams will be slightly different than BIOL 442 Exams.

DATA SOURCE PRESENTATIONS will take place during class on February 2. Each group of students (2 per group) will choose and sign up for an online data source of biodiversity data. I will present a possible list of data sources; however if your group would like to choose a different data source, please double check with me. In class on February 2, each group will present on the data source, including information such as data formats, data extent (geographical, chronological), data access issues or costs, etc. Basically, you should answer questions such as: who would use this data source, how could one use it, and how are data added to this source?

BIOL 575 presentations will be graded to a higher standard than BIOL 475.

SPECIES DISTRIBUTION MODELING (SDM) ABSTRACT AND PRESENTATIONS will take place during the last weeks of class (see schedule). Each group of students (3 per group for 475 students, 2 per group for 575 students) will choose and sign up for a focal species to analyze. Each species' data will be downloaded, processed, and mapped, and then analyzed with SDM methods, including cross-validation and mapping.

The abstract will be written as though you were presenting the results of your research at the Ecological Society of America. I will provide specific guidance in class. This abstract writing exercise will strengthen your ability to write concisely and scientifically.

The presentations will be held during the last class period and during the final exam period. Each group's SDM presentation will be graded on your professionalism, use of time, use of visuals, results of your analysis, and your overall story cohesion.

BIOL 575 abstracts and presentations will be graded to a higher standard than BIOL 475.

Academic Integrity

As a student at St. Cloud State University and as a student in this class, you are expected to fully and properly acknowledge the work of others. Every instance of plagiarism will be reported, as per the policies of the college, but please do not hesitate to ask me in advance if you think something might be questionable or if you are unsure about what is considered to be plagiarism. I am happy to help, as long as you inquire in advance!

Academic misconduct includes but is not limited to:

- cheating: using a resource other than one's own work to answer questions;
- plagiarism: misrepresenting another's ideas as one's own or not giving credit to the creator of a work;
- falsification: submitting falsified or fabricated information;
- facilitating others' violations: knowingly permitting or facilitating the dishonesty of others;
- impeding: placing barriers in the way of others' academic pursuits'

Instances of academic dishonesty will result in either a failing grade for that activity or for the course, according to the perceived intent and extent of the instance(s) of academic dishonesty. All academic integrity violations will be reported.

Course Schedule (version dated January 3, 2022)

Monday	Tuesday	Wednesday
Jan 10th	11th	12th **Topic:** Introduction to biodiversity informatics, introduction to reproducible research *Bioinformatics Data Skills Ch 1* *Lab 1: Molecular Evolution*
17th *Lab 1 due on D2L by 10pm*	18th	19th **Topic:** Introduction to Project Management, git/GitHub, RStudio *Bioinformatics Data Skills Ch 2, 5* *Lab 2: Set up GitHub/RStudio*
24th *Lab 2 due on GitHub by 10pm*	25th	26th **Topic:** Data Sources & Downloading data *Bioinformatics Data Skills Ch 6* *Journal Club 1*
31st	Feb 1st	2nd **Data Source Presentations**
7th	8th	9th **Topic:** Introduction to R *Bioinformatics Data Skills Ch 8* *Lab 3: R analysis code* *Journal Club 2*
14th *Lab 3 due on GitHub by 10pm*	15th	16th **Topic:** R continued *Bioinformatics Data Skills Ch 8* *Lab 4: R graphing code* *Journal Club 3*
21st *Lab 4 due on GitHub by 10pm*	22nd	23rd **Topic:** Working with Range data *Bioinformatics Data Skills Ch 9* *Lab 5: Basic Range Arithmetic*
28th *Lab 5 due on GitHub by 10pm*	Mar 1st	2nd **Topic:** Working with Range data *Bioinformatics Data Skills Ch 9* *Lab 6: Finding Overlaps in Range data*

MONDAY	TUESDAY	WEDNESDAY
7th *No class*	8th *No class*	9th *No class*
14th *Lab 6 due on GitHub by 10pm*	15th	16th **Topic:** Working with Range data *Bioinformatics Data Skills Ch 9* *Lab 7: Advanced Range Operations* *Journal Club 4*
21st *Lab 7 due on GitHub by 10pm*	22nd	23rd **Topic:** Metabarcoding *Read metabaR article* *Lab 8: metabaR package* *Journal Club 5*
28th *Lab 8 due on GitHub by 10pm*	29th	30th **Topic:** Environmental DNA *No lab* *Journal Club 6*
Apr 4th	5th	6th **Topic:** Species Distribution Modeling *Lab 9: SDM analysis* *Journal Club 7*
11th *Lab 9 due on GitHub by 10pm*	12th	13th **Topic:** Species Distribution Modeling *Lab 10: SDM validation*
18th *Lab 10 due on GitHub by 10pm*	19th	20th **Topic:** SDM Work day *Journal Club 8*
25th	26th	27th **SDM Presentations**
May 2nd **SDM Presentations** Final Exam Period: 12:20pm - 2:35pm	3rd	4th