1 Course Information

- Number: CMSC 829A
- Section: 0101
- Title: Algorithmic Evolutionary Biology
- Credits: 3
- Instructor: Erin Molloy (she/her/hers)
- Term: Fall 2021 (August 30th to December 21st)
- Times: Mondays and Wednesdays 9–10:15 AM
- Location: IRB-1116 unless otherwise indicated on the schedule
- Office Hours: Thursdays 3:15–4:45 PM (IRB-3256) or by appointment (Zoom)

2 Course Description and Structure

This is a graduate-level course on the design, analysis, and evaluation of algorithms for problems in evolutionary biology, with some detours to other domains (e.g. metagenomics) where related algorithms have utilized to much success. Topics will include multiple sequence alignment, statistical estimation of gene trees, species trees, and admixture graphs, as well as approaches based on parsimony. The algorithmic techniques come from a variety of areas, such as discrete optimization, graph theory, as well as probability and statistics. Therefore, this course is intended for graduate students in computer science, mathematics, statistics, and related disciplines. Familiarity with algorithms, probability, and basic statistics is expected. No prior knowledge of biology is required. Please contact me if you are unsure as to whether you should enroll in the course.

During the semester, you will be expected to

- learn fundamental models and algorithms from the field of phylogenetics / phylogenomics
- read recent literature in phylogenomics and other domains where related algorithms are utilized
- act as a discussion facilitator and contribute to discussions (e.g. ask questions, share observations, identify strengths/weakness of current methodological practices, integrate knowledge across papers and across topics/domains, propose directions for future research).
- seek to contribute to the field of phylogenomics or a related domain, through a final project
A tentative course schedule is available at the end of this document. While the exact schedule is subject to change, we will adhere to the following course structure.

In the first part of the semester, we will operate as a traditional (i.e. lecture-style) course, covering fundamental models and algorithms from the field of phylogenetics / phylogenomics. Course materials for each lecture (e.g. slides, recordings, and recommended homework problems) will be posted to ELMS. I highly encourage you to look at recommended homework problems. Although these questions will not be graded, they will be indicative of the type of questions on the take-home midterm exam (note: this exam will be open book, meaning you will be allowed to use the course materials and your notes—but not google search, etc.). The midterm exam worth is worth 30% of the final grade; this constitutes the “qualifying” aspect of the course.

In the second part of semester, we will operate as a reading course. Each week, we will collectively read and discuss several recent papers (all available on ELMS) on the same topic, within a small-to-large group discussion framework. On Mondays, class sessions will be used for small group discussions, with each small group (max size TBD) having a different assigned reading (typically 1 paper). On Wednesdays, class sessions will be used for large (i.e. class-wide) discussions that begin with short presentations (5–8 minutes) of each small group’s assigned reading. To this end, each small group will have an assigned discussion facilitator for each week of the reading group (schedule TBD). It is their responsibility to

- prepare for small group discussion, not only by doing their assigned reading (which is required of all small group members) but also by submitting reading comprehension and discussion questions to a designated forum ELMS in advance
- lead the small group discussion on Monday
- create/give a lightening talk on their assigned reading on Wednesday

Overall, I hope that these discussions will be lively and fun, giving students the opportunity to ask questions and share observations from individual papers—as well as to make connections between the papers (especially those on different topics) and to identify directions for future research. I also hope that this framework increases student engagement during class and reduces student work load outside of class, compared to a seminar-style course. Please contact me if you have questions or concerns.

Some class time will be dedicated to final projects, which may be conducted individually or in a small group (max size: 3 students). You might consider conducting an experimental study to replicate the results of a recently published study or to address a hypothesis that you have about method performance, improving (e.g. parallelizing) an existing method, or developing a new method. Students should meet with me to agree on the basic direction of their project, and ultimately produce a written plan. More details and suggested topics will be provided later in the semester. In a similar course taught by Prof. Tandy Warnow at the University Illinois of Urbana-Champaign, it is not uncommon for final projects to be published in conferences or journals; see this website. My course project was the basis for this paper.

3 Grading Overview

30% – Midterm Exam (due on Wednesday, October 13th at 9:00 AM ET)
35% – Reading Group
- 20% – Discussion Facilitator for small group / Presenter for large group (see reading group schedule!)
- 15% – Participation and Engagement
35% – Final project
- 10% – Project Plan
First draft of project proposal/plan due Wednesday, October 20th at 11:59 PM ET

Revision of project proposal/plan due on Monday, November 1st at 11:59 PM ET

- 10% – Project Presentation (due in weeks 15–16; schedule TBD)
- 15% – Project Report (due on Friday, December 17th at 5:00 PM ET)

4 Course Participation

During the semester, we will be reading and discussing papers together, and learning from one another. Participation is key to this collaborative process, not only for your own learning, but for your fellow classmates’ learning as well. You are expected to attend (in-person) all discussion-based classes, having done the assigned reading and ready to participate in class activities.

That said, students may be absent from class unexpectedly, for example due to sickness, quarantining, or caregiving. Please communicate with me regarding unexpected absences as soon as possible, especially if you are scheduled to present and/or if you able to participate remotely (your small group or I may be able to facilitate this via Zoom or other platforms). The University’s recommended guidelines to address short-term absences related to COVID-19 can be found on this website, and the Graduate School’s course related policies on excused absences can be found on this website.

Recordings

Lectures and class-wide discussions will be recorded and posted to ELMS. This will enable students to access the course in the event of unexpected absences. These recordings will only be made available to students enrolled in the course and are not to be redistributed anywhere (note: it is a FERPA violation to redistribute these videos).

Religious Holidays

Excused absences for religious holidays will be made on an individual basis. You must request accommodations for religious reasons during the first two weeks of classes.

5 Academic Integrity

The University’s Code of Academic Integrity is designed to ensure that the principles of academic honesty and integrity are upheld. All students are expected to adhere to this Code. Please ensure that you fully understand this code and its implications because all acts of academic dishonesty will be dealt with in accordance with the provisions of this Code.

Campus Policies

It is our shared responsibility to know and abide by the University of Maryland’s policies that relate to all courses, which include topics like:

- Academic integrity
- Student and instructor conduct
- Accessibility and accommodations
- Attendance and excused absences
- Grades and appeals
- Copyright and intellectual property

Please visit this website for the Graduate School’s full list of campus-wide policies.
6 Course Guidelines

Communication with Instructor

Email: If you need to communicate with me, please email me at ekmolloy@umd.edu with the subject line starting with CMSC829A or attend office hours.

ELMS: I will send IMPORTANT announcements via ELMS messaging. You must make sure that your email and announcement notifications (including changes in assignments and/or due dates) are enabled in ELMS so you do not miss any messages. You are responsible for checking your email and Canvas/ELMS inbox with regular frequency.

Communication with Peers

With a diversity of perspectives and experience, we may find ourselves in disagreement and/or debate with one another. As such, it is important that we agree to conduct ourselves in a professional manner and that we work together to foster and preserve a classroom environment in which we can respectfully discuss and deliberate controversial questions.

I encourage you to exercise your right to free speech—bearing in mind, of course, that you will be expected to craft and defend arguments that support your position. Keep in mind, that free speech has its limit and that this course is NOT the space for hate speech, harassment, and derogatory language. I will make every reasonable attempt to create an atmosphere in which each student feels comfortable voicing their argument without fear of being personally attacked, mocked, demeaned, or devalued.

Any behavior (including harassment, sexual harassment, and racially and/or culturally derogatory language) that threatens this atmosphere will not be tolerated. Please alert me immediately if you feel threatened, dismissed, or silenced at any point during our semester together and/or if your engagement in discussion has been in some way hindered by the learning environment.

Names/Pronouns and Self-Identifications

The University of Maryland recognizes the importance of a diverse student body, and we are committed to fostering inclusive and equitable classroom environments. I invite you, if you wish, to tell us how you want to be referred to both in terms of your name and your pronouns (he/him, she/her, they/them, etc.). The pronouns someone indicates are not necessarily indicative of their gender identity. Visit this website to learn more. Additionally, how you identify in terms of your gender, race, class, sexuality, religion, and dis/ability, among all aspects of your identity, is your choice whether to disclose (e.g., should it come up in classroom conversation about our experiences and perspectives) and should be self-identified, not presumed or imposed. I will do my best to address and refer to all students accordingly, and I ask you to do the same for all of your fellow Terps.

Mask Wearing

President Pines provided clear expectations to the University about the wearing of masks for students, faculty, and staff. Face coverings over the nose and mouth are required while you are indoors at all times. There are no exceptions when it comes to classrooms, laboratories, and campus offices. Students not wearing a mask will be given a warning and asked to wear one, or will be asked to leave the room immediately. Students who have additional issues with the mask expectation after a first warning will be referred to the Office of Student Conduct for failure to comply with a directive of University officials.

Other COVID-19 policies and protocols

The University has put forth policies related to COVID-19, many of which can be found on the Provost’s website. For information on COVID-19 protocols, including requirements for coming to campus, visit this
website. It is our shared responsibility to stay up-to-date on the University’s COVID-19 protocols and policies throughout the semester.

7 Course Concerns and Incident Reporting

Please review this information about reporting course concerns as well as incidents of sexual harassment, student contact, and hate bias.

8 Course Evaluation

Please submit a course evaluation through CourseEvalUM in order to help faculty and administrators improve teaching and learning at the University. All information submitted to CourseEvalUM is confidential. Campus will notify you when CourseEvalUM is open for you to complete your evaluations. Please go directly to the Course Eval UM website to complete your evaluations. By completing all of your evaluations each semester, you will have the privilege of accessing through Testudo, the evaluation reports for the thousands of courses for which 70% or more students submitted their evaluations.

9 Resources & Accommodations

Accessibility and Disability Services

The University of Maryland is committed to creating and maintaining a welcoming and inclusive educational, working, and living environment for people of all abilities. The University of Maryland is also committed to the principle that no qualified individual with a disability shall, on the basis of disability, be excluded from participation in or be denied the benefits of the services, programs, or activities of the University, or be subjected to discrimination. The Accessibility & Disability Service (ADS) provides reasonable accommodations to qualified individuals to provide equal access to services, programs and activities. ADS cannot assist retroactively, so it is generally best to request accommodations several weeks before the semester begins or as soon as a disability becomes known. For assistance in obtaining an accommodation, contact ADS via email adsfrontdesk@umd.edu or phone 1-301-314-7682.

Any student eligible for and requesting reasonable academic accommodations due to a disability is requested to provide, to the instructor via email, a letter of accommodation from ADS within the first two weeks of the semester (or within a reasonable timeframe of the disability becoming known).

Student Resources and Services

I encourage you to visit the Student Academic Support Services website to learn more about the wide range of campus resources available to you. In particular, everyone can use some help sharpening their communication skills by visiting the Writing Center. There are a wide range of other resources available to support you. The Student Resources and Services website provides information on

- Resources for Students with Children
- Health Services
- Sexual Assault Resources
- Technology Resources
- University Libraries
- Financial / Resource Assistance
- COVID-19 Information
and more. Other important links include the LGBT Equity Center, International Student Scholars Services, and Veteran Student Life.

Mental Health

Diminished mental health, including significant stress, mood changes, excessive worry, substance/alcohol abuse, or problems with eating and/or sleeping can interfere with optimal academic performance, social development, and emotional wellbeing. If you or someone you know experiences any of the above mental health concerns, it is strongly encouraged to contact or visit any of the University's resources, including UMD Mental Health Center or UMD Counseling Center. Getting help is a smart and courageous thing to do—for yourself and for those who care about you.

Basic Needs Security

If you have difficulty affording groceries or accessing sufficient food to eat every day, or lack a safe and stable place to live, visit the Division of Student Affairs website for information about resources the campus offers.

10 Copyright Notice

Students are permitted to use course materials for their own personal use only. Course materials may not be distributed publicly or provided to others, in any way or format.

11 Acknowledgements

CMSC829A is largely based on courses taught by Profs. Tandy Warnow (at the University of Illinois at Urbana-Campaign), Siavash Mirarab (at the University of California, San Diego), and Luay Nakhleh (at Rice University). The small-to-large group discussion framework is based on Prof. Brent Roberts' course Psych 593 Topics in Personality Psychology, offered at University of Illinois at Urbana-Champaign in Fall 2017. Some policy-related text in this syllabus has been borrowed from Profs. Rachel Rudinger, Chris Metzler, and Leonidas Lampropoulos, and Mohammed El Kebir.

12 Right to Change Information

This is the current version of the syllabus. Unforeseen circumstances arising during the semester could require the adjustment of any material given here. Therefore, I reserve the right to change any information on this syllabus (as well as the course schedule and the course materials) during the semester, given due notice to students.

13 Tentative Course Schedule

A tentative course schedule for CMSC829A Algorithmic Evolutionary Biology in Fall 2021 is on the following page. Reading group topics, readings, and discussion facilitators will be assigned by Friday, October 6th, so that discussion facilitators have sufficient time to prepare. Please send me ideas for topics and/or papers!
# Table 1: This is the tentative schedule for CMSC829A Algorithmic Evolutionary Biology.

<table>
<thead>
<tr>
<th>Week</th>
<th>Date</th>
<th>Agenda</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>08/30</td>
<td>Go through syllabus and give course overview</td>
</tr>
<tr>
<td></td>
<td>09/01</td>
<td><strong>L1 – Phylogenetic Trees and Compatibility</strong>&lt;br&gt;Topics: Rooted and unrooted phylogenetic trees, Newick strings, clades, bipartitions, Robinson-Foulds distance, tree compatibility, compatibility supertrees, triplets, ASSU algorithm, quartets, All Quartets algorithm</td>
</tr>
<tr>
<td>2</td>
<td>09/06</td>
<td>Holiday — No class!</td>
</tr>
<tr>
<td></td>
<td>09/08</td>
<td><strong>L2 – Characters: Compatibility, Parsimony, &amp; Supertrees</strong>&lt;br&gt;Topics: Parsimony score and Fitch’s Algorithm, homoplasy and character compatibility, perfect phylogeny problem, maximum compatibility problem, (large) maximum parsimony, randomized taxon addition, branch-and-bound, heuristic search moves (NNI, SBR, TBR), Matrix Representation with Parsimony (MRP) supertree problem</td>
</tr>
<tr>
<td>3</td>
<td>09/13</td>
<td><strong>L3 – Phylogeny estimation under CFN model - Part 1</strong>&lt;br&gt;Topics: How to generate data under Cavendar-Farris-Neyman (CFN) model (a substitution-based model of binary character evolution), How to compute probability of characters under CFN model, Felsenstein’s Pruning Algorithm, Maximum Likelihood problem, statistical consistency</td>
</tr>
<tr>
<td></td>
<td>09/15</td>
<td><strong>L4 – Phylogeny estimation under CFN model - Part 2</strong>&lt;br&gt;Topics: $p$-distance, CFN-corrected distance, additive (and nearly additive), four point condition, Naive Quartet method, UPGMA, ultrametric, Neighbor Joining</td>
</tr>
<tr>
<td>4</td>
<td>09/20</td>
<td><strong>L5 – Models of Molecular Sequence Evolution</strong>&lt;br&gt;Topics: time-continuous Markov Chains, SHR Assumptions (stationarity, homogeneity, reversibility), Jukes-Cantor model, Generalized Time Reversible (GTR) model, varying rates across sites, heterotachy, insertions/deletions (treated as missing data), taxon sampling</td>
</tr>
<tr>
<td></td>
<td>09/22</td>
<td><strong>L6 – Multiple Sequence Alignment - Part 1</strong>&lt;br&gt;Topics: interpretation of multiple sequence alignment, error quantification (SPFN, SPFP, TC), edit distance and Needleman-Wunsch algorithm, (Generalized) Tree Alignment problem, Sum-of-Pairs (SOP) alignment problem</td>
</tr>
<tr>
<td>5</td>
<td>09/27</td>
<td><strong>L7 – Multiple Sequence Alignment - Part 2</strong>&lt;br&gt;Topics: Sequence profiles, profile Hidden Markov Models (HMMs), aligning alignments, progressive alignment, consistency, Maximum Weight Trace (MWT) problem, iterative refinement, divide-and-conquer, popular MSA methods (Clustal W/Ω, T-Coffee, MAFFT, MUSCLE, PASTA)</td>
</tr>
<tr>
<td></td>
<td>09/29</td>
<td>Review session in preparation for midterm</td>
</tr>
<tr>
<td>6</td>
<td>10/04</td>
<td><strong>L8 – Species tree estimation under MSC model</strong>&lt;br&gt;Topics: Multi-Species Coalescent (MSC) model, probability of rooted and unrooted gene Trees under MSC, pseudo-likelihood method based on triplets model (MP-EST), constrained discrete optimization based on quartets (ASTRAL), distance-based methods (ASTRID), site-based method (SVDquartets), Bayesian co-estimation methods (*BEAST)</td>
</tr>
<tr>
<td></td>
<td>10/06</td>
<td>Hand out take-home midterm exam and discuss logistics for reading group and course projects</td>
</tr>
</tbody>
</table>
Table 2: This is the tentative schedule for CMSC829A Algorithmic Evolutionary Biology, cont.

<table>
<thead>
<tr>
<th>Week</th>
<th>Date</th>
<th>Agenda</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>10/11</td>
<td><strong>L9 – Species tree estimation under MSC model, cont.</strong></td>
</tr>
</tbody>
</table>
|      | 10/13| **L10 – Species tree estimation under MSC model, cont.**  
*Topics: Infinite Sites model, Infinite Alleles model, Wright-Fischer model, character parsimony methods (Dollo, Camin-Sokal, and polymorphism parsimony), quartet-based methods (ASTRAL_BP – topology, branch length, and branch support), allele frequency, genetic drift, and f-statistics, admixture graphs* |
| 8    | 10/18| Review midterm exam (solutions) and check-in about preparedness for reading group and course projects |
|      | 10/20| **L11 – Statistical Estimation of Species Trees under GDL model** |
| 9    | 10/25| **RG1 – Multiple Sequence Alignment**  
10/26 Small group discussion  
10/27 Class-wide discussion |
| 10   | 11/01| **RG2 – Plant Phylogeny and Whole Genome Duplication**  
11/01 Small group discussion  
11/03 Class-wide discussion |
|      | 11/08| **RG3 – Bacteria Phylogeny and HGT**  
11/08 Small group discussion  
11/10 Class-wide discussion |
|      | 11/15| **RG4 – Virus Phylogeny and Recombination**  
11/15 Small group discussion  
11/17 Class-wide discussion |
|      | 11/22| Discuss final projects and course so far  
11/24 Holiday — No class! |
| 14   | 11/29| **RG5 – Presenter’s Choice**  
11/29 Small group discussion  
12/01 Class-wide discussion |
|      | 12/06| Project presentations! |
|      | 12/08| Project presentations! |
| 15   | 12/13| (Last day of class) Project presentations! |
| 17   | 12/16| **Project report due at 5:00 PM ET!** **