Big Data in Biology (BDB)

Dr. Melinda A. Yang
In-person: Gottwald B117
Zoom: http://tiny.cc/Yang_OfficeHours

T/R 12:00-1:15 pm EST, Gottwald C200
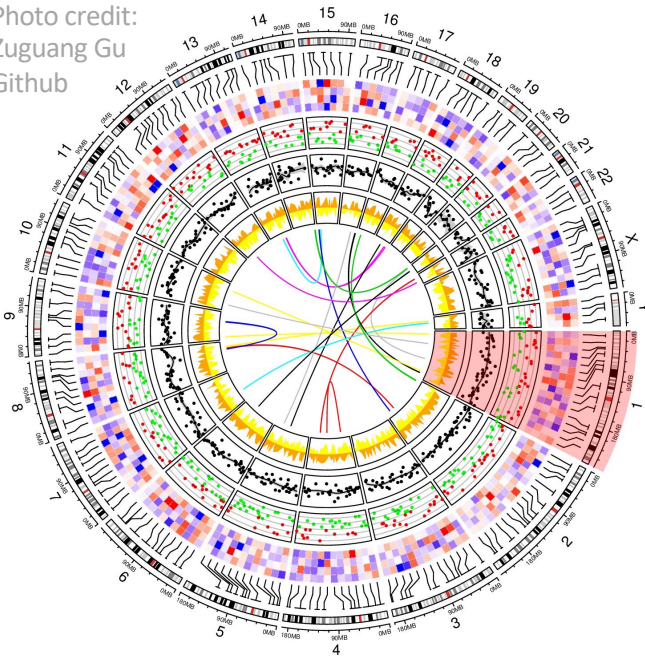W 1:30-4:20 pm EST, Gottwald A201
Office hours by appt at:
https://melyang.youcanbook.me

Photo credit: Zuguang Gu Github

- **What makes a good scientist?**
- **Why are fruit flies so important to human health?**
- **How did computers revolutionize the field of biology?**
- **Who has access to data, and why does that matter?**

In this course, we will use genomic data, computer-based tools, and foundational molecular and evolutionary concepts to research and answer biological questions. We will practice and develop research and communication skills that promote scientific thinking, and hopefully do some cool biological research!

Major Resources
- Online tools: Blackboard, Perusall, Box
- "A Short Guide to Writing About Biology" by Jan Pechenik (older 8th ed. suggested)
- Personal Laptop, Binder/Loose-leaf, Stapler, 3-hole punch, Writing Utensil
- Microsoft Word and Excel (not just Google Drive versions), see http://tiny.cc/word_excel to download

Staying Healthy
We are managing many different physical and mental stressors in our lives. My goal is to be transparent and understanding. To help me do this, I ask that you stay in communication and strive to develop spaces where you can focus on learning.

What you should expect from me
In BDB, all students are welcome regardless of race/ethnicity, gender identities, gender expressions, sexual orientation, socioeconomic status, age, disabilities, religion, regional background, veteran status, citizenship status, nationality and other diverse identities.

As your instructor, I am committed to creating a learning environment that is safe, healthy and fair for all students. I am also committed to fully involving and engaging all students as part of our learning community.

To learn effectively, I encourage you to discuss with me any barriers to learning you may experience. One of the best ways to do so is to come and chat with me about any questions or problems that come up.

Course Format

The current academic environment primarily uses a traditional grading system (e.g., most universities require assignment of a letter grade associated with a numerical score) and is connected to high student anxiety and depression. Our main goal is learning, and studies have shown that a traditional grading system doesn't provide clear feedback, reduces motivation for learning, and often does not accurately test the mastery of skills and concepts emphasized in the class.

To encourage an intrinsic focus on learning, the grading for this course follows a 'specifications (specs) grading' format, which has three major elements: **(1) encouraging student goals focused on mastery of skills and content rather than increasing score; (2) creating tiers based on amount of satisfactory work completed to increase transparency of grading system and facilitate student-driven learning; and (3) connecting assessments directly to course learning goals**.

Specs grading assumes that grades should be based on your mastery of a pre-specified set of learning objectives. If you submit work deemed satisfactory, you are mastering learning objectives. When work is non-satisfactory, feedback and the opportunity for revision will assist you in progressing towards mastery. Partial credit obscures where mastery is not yet achieved and disincentivizes revision, so a key aspect of specs grading is use of a pass/fail system (for us, *a check-based system*). Since UR requires assignment of a final letter grade, this letter grade will be based on demonstrating your learning through satisfactory completion of a pre-specified number of assignments, with more work (i.e. assignments) put in for higher letter grades. See p. 4 for more details.
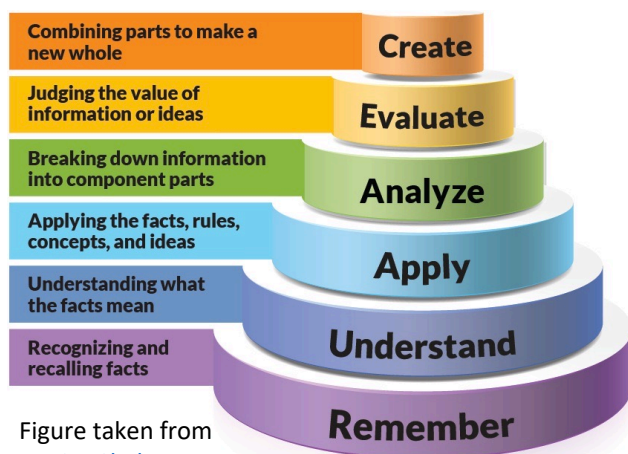

Combining parts to make a new whole	**Create**
Judging the value of information or ideas	**Evaluate**
Breaking down information into component parts	**Analyze**
Applying the facts, rules, concepts, and ideas	**Apply**
Understanding what the facts mean	**Understand**
Recognizing and recalling facts	**Remember**

Figure taken from Jessica Shabatura at U. of Arkansas

LEARNING OUTCOMES

Figure taken from Michigan State University.

✓+ Understanding of material is excellent, with no errors.

✓ High understanding of material is evident w/o need for revision. Has **minor** errors warranting revision, covered by comments.
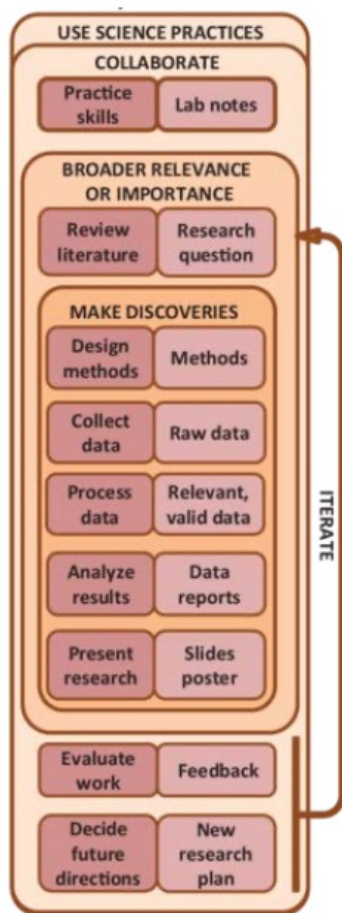
✓- Understanding of material is not clear. Exhibits many minor errors or one or more major errors that necessitate revision.

∅ Little to no understanding of material, or assignment not completed/submitted following instructions

Grading schema adapted from Dr. Rebelsky's CS course at Grinnell College.

Additional Noteworthy Items

Absences*:* I expect you to attend every class/lab and fully participate. Excused absences that are scheduled well in advance (e.g. conflict with religious holiday observance, university-sponsored activities with approval from your Residential Dean) must be shared with me *in the first week of classes if possible*, or as soon as the scheduling conflict arises, otherwise. For excused absences affecting multiple classes (e.g. serious illness, death of family member), I will work on a case by case basis with you to determine a fair makeup policy. If you are feeling ill, please contact the Student Health Center, do not attend class, and let me know asap - following these steps will generally result in an excused absence. If you have an excused absence, it is **your responsibility** to contact me and discuss how to receive participation credit.

Academic Integrity*: We, the students of the University of Richmond, shall promote and uphold a community of integrity and trust.* Each of you pledge to abide by this Honor Code as a UR student, and I trust each student to act with integrity, clearly separating their own work from that of others. This includes, for example, correctly citing sources, accurately distinguishing between group and individual work, and stopping and reporting dishonest actions when observed. If you are unsure whether an action you or another student is taking violates the Honor Code, asking the instructor for clarification as early as possible will be the best way to avoid unintentional plagiarizing and/or cheating. Any work suspected to have violated the Honor Code is subject to being submitted to the Honor Council.

Token System*:* We will be using a token system to add more flexibility regarding deadlines and absences. You will begin the course with **one token**, which you can use to obtain an extension of up to 24 hours (you must specify new deadline), treat an unexcused absence as an excused absence, or make up a missed cumulative quiz. Note that requests to use tokens must be initiated in writing by you. Occasional opportunities to gain more tokens are detailed here.

Disabilities Accommodation: Any student with disability accommodation should let me know as soon as possible. For me to provide accommodations, you will need a Disability Accommodation Notice (DAN) from the University of Richmond. If you need a DAN, please see the Office of Disabilities Services (804-289-8032) for more information.

1. Readings and videos are scored based on quality of comment (Perusall) or completion. Quality comments must have score 2.5+ for acceptance.

2. Class participation takes many forms dependent on that day's activities. The easiest way to gain credit is to show up on time and participate in provided activities. Credit for excused absences usually require an out-of-class assignment, contact me to request.

3. Lab participation is related to the completion of 1-2 worksheets started in lab. The assignment (usually a worksheet) will be due by a date specified by the instructor.

4. Cumulative quizzes (CQs) are given every 3 weeks in class (5x's) and are based on pre-assigned questions or question topics based on learned material. Two random questions will be given for each CQ, with additional opportunities in final exam week.

How Assessments are Evaluated in BDB

Course Grade Minimum	A-	B-	C-
Assessments (total #)	# Completed Satisfactorily & On Time		
1. Close Readings (1-4/wk)	90%	80%	70%
2. Participation-Class (28)	27	24	21
3. Participation-Lab (14)	14	13	12
4. Cumulative Quiz Questions (16)	10	8	6
5. Challenge Problems (~9)	7	5	4
6. Annotated Bibliographies (3)	3	2	1
7. Weekly Reflection (14)	13	12	11
8. Self-evaluations	3	3	3
9. Paper Write-ups	See additional instructions on BB.		

5. Challenge Problems (CPs) ask students to synthesize their learned knowledge and skills for effective analysis, evaluation, or creation. They are tied to 1+ learning objectives. A new submission, a revision, or both can be submitted each Tuesday at start of class. Revisions are likely, so make time for them. Ideally, you will have completed 3 CPs to satisfaction by Week 8.

7. Weekly reflections involve completing a two question form by week's end to reflect on activities from that week. This activity facilitates a growth mindset for your learning, helps me gauge what to add/edit in the classroom, and increases communication between me and you. Submission of form constitutes a pass.

6. Annotated bibliographies (ABs) are 1-paragraph summaries of scientific articles. Two ABs will be based on class readings, and additional ABs will be based on papers you find in literature searches (vetted by me) based on our Pathways research project.

8. Three self-evaluations (SEs) due in Wks 2, 8, and during finals. These are meant to help you and me see your progress and evaluate your learning. For completion only.

9. Paper write-ups are where you show your mastery of the heart of this course, i.e. your understanding of conducting scientific research and communicating your findings. This primarily involves writing the Results/Discussion sections of a paper and giving peer feedback. A description of what is needed for each course grade tier is described further on Blackboard. You must complete lower tier items to satisfaction first, before moving on to higher tier items.

Final grade within letter tier (e.g. B-/B/B+) is at my discretion, and will be based on responses in Self-Evals 2 and 3, percentage of ✓+ and ✓s overall, observed growth/effort throughout the course, and extra assignments done beyond the required specifications.

Learning Objectives for BIOL199: Big Data in Biology

A. Appreciate the impact of big data in biology and understand how the rise of big data in biology is deeply intertwined with foundational evolutionary concepts
1. Explain how diversity of life evolved over time by processes of mutation, selection, gene flow, and random processes
2. Recognize when evolutionary concepts underlie biological research (e.g. explaining why studying one species can be informative about another species)
3. Explain the difference between homology and similarity
4. Develop example of an evolutionary process
5. Interpret phylogenetic trees using an evolutionary model

B. Understand how comparative approaches combined with large datasets can uncover natural history of living beings
1. Communicate relevant content related to specified question or problem and be able to discuss broader implications of question or problem
2. Analyze data related to a specified question or problem, particularly through the use of online databases, shell scripting, and popular computational tools
3. Relate concepts from multiple biological fields to interpret biological phenomena related to specified question or problem (e.g. molecular, computational, organismal)
4. Propose potential solutions for specified question or problem, both through developing hypotheses and interpreting results

C. Understand how modern biologists derive important and answerable questions and design and execute experiments
1. Define the hallmarks of good experiments and/or studies that address specific hypotheses.
2. Use observational and organizational skills to record and annotate experimental outcomes and/or observations of the natural world
3. Understand, execute, and troubleshoot protocols required for data collection and analysis
4. Make measurements and use quantitative methods to evaluate results
5. Recognize the power and limitations of the scientific process - describe how data is not inherently objective, and they are subject to error, misinterpretation, and bias

D. Learn how to interpret, summarize, and evaluate scientific research, both your own and that of others
1. Create and interpret informative graphs and other data visualizations
2. Communicate experimental results in a concise and clear manner, both in written and oral formats
3. Use online databases and data science tools to find information (e.g. previously published research) relevant to biological questions
4. Distinguish between different types of information sources and develop skills for searching for primary literature
5. Demonstrate conventional and proper use of sources in scientific writing

E. Understand what it means to be an ethical scientist and why this is important
1. Identify and employ the characteristics of research integrity in experimental design, data collection, analysis, and communication
2. Describe why accurate acknowledgement of intellectual contributions from others is important
3. Demonstrate conventional and proper use of sources in scientific writing to acknowledge intellectual contributions from others
4. Describe the ethos of open access in science and its impact on the field of biology, particularly bioinformatics

For detailed listings of assignments/readings, the best resource are the weekly pages on Blackboard

Calendar is subject to change. Revised calendars will be updated on Blackboard and announced through Perusall

Calendar

Week	Date	Topic		Lab	
1T	23-Aug	Introduction to Course		BioMAPS, BioSquare, e-notebooks	NGS Module
1R	25-Aug	Sequencing: DNA Structure, Sequencing Processes, Coverage/Mapping			
2T	30-Aug		CP	NGS1: Read Quality	
2R	1-Sep				
3T	6-Sep		CP	NGS2: Statistical Confidence	
3R	8-Sep		CQ		
4T	13-Sep	Central Dogma (for Gene Annotation) and RNA-Seq	CP	UEG1	UEG Module
4R	15-Sep				
5T	20-Sep		CP	UEG2	
5R	22-Sep	Journal Club 1 + AB		(NGS Paper Due)	
6T	27-Sep	Evolutionary Foundations (Forces, Misconceptions, Phylogenies, Synteny)	CP	BLAST	
6R	29-Sep		CQ		
7T	4-Oct		CP	Pathways A-C	Pathways Module
7R	6-Oct				
8T		FALL BREAK: NO CLASS	CP	PathwaysA-C	
8R	13-Oct	Variation (Mutations), Comparative Genomics, and Selection			
9T	18-Oct		CP	Pathways D-F	
9R	20-Oct		CQ		
10T	25-Oct		CP	Pathways D-F / G	
10R	27-Oct			(Pathways A-C Paper Due)	
11T	1-Nov	Journal Club 2 + AB	CP	Peer Revisions and Finish Pathways G	
11R	3-Nov				
12T	8-Nov	Protein Structure, Function, and Interaction	CP	Primary Literature Research / Wrap up Pathways	Presentation Module
12R	10-Nov		CQ		
13T	15-Nov		CP	Presentation Prep Day	
13R	17-Nov	TBA		(Paper Revisions Due)	
14T	22-Nov		CP	THANKSGIVING: NO LAB	
14R		THANKSGIVING: NO CLASS			
15T	29-Nov	Presentation Prep Day	CP	Final Pres. + Course Evals	
15R	1-Dec	Final Activity	CQ	(Pathways Report Due)	

CQ	Cumulative Quiz: Doesn't include material covered that week, six additional CQs offered during finals
CP	Challenge Problem: Can submit one new CP, one revised CP, or both each week by Tuesday start of class (up to 7 required for an A-)
AB	Annotated Bibliography: Two are due for articles discussed as a class. Additional ABs can be submitted after Wk 12 based on your Pathways research.

If you experience difficulties in this course, do not hesitate to consult with me. Below are other resources that can support you in creating a learning environment that works for you.

Academic Skills Center (asc.richmond.edu): Academic coaches assist students in assessing and developing their academic and life-skills (e.g., critical reading and thinking, information conceptualization, concentration, test preparation, time management, stress management, etc.). Peer tutors offer assistance in specific subject areas (e.g., calculus, chemistry, accounting, etc.) and will be available for appointments in-person and virtually. Peer tutors are listed on the ASC website**. Email** Roger Mancastroppa (rmancast@richmond.edu) **and** Hope Walton (hwalton@richmond.edu) **for coaching appointments in academic and life skills.**

Boatwright Library Research Librarians: (library.richmond.edu/help/ask/ or 289-8876): Research librarians help students with all steps of their research, from identifying or narrowing a topic, to locating, accessing, evaluating, and citing information resources. Librarians support students in their classes across the curriculum and provide individual appointments, class library instruction, tutorials, and research guides (libguides.richmond.edu). Students can contact an individual librarian(library.richmond.edu/help/liaison-librarians.html) or ASK a librarian for help via email (library@richmond.edu), text (804-277-9ASK), or chat (library.richmond.edu/chat.html).

Career Services: (careerservices.richmond.edu or 289-8547): Can assist you in exploring your interests and abilities, choosing a major or course of study, connecting with internships and jobs, and investigating graduate and professional school options. We encourage you to schedule an appointment with a career advisor early in your time at UR.

Counseling and Psychological Services: (caps.richmond.edu or 289-8119): Assists currently enrolled, full-time, degree-seeking students in improving their mental health and well-being, and in handling challenges that may impede their growth and development. Services include brief consultations, short-term counseling, skills-building classes, therapy groups, crisis intervention, psychiatric consultation, and related services.

Disability Services: (disability.richmond.edu): The Office of Disability Services works to ensure that qualified students with a disability (whether incoming or current) are provided with reasonable accommodations that enable students to participate fully in activities, programs, services and benefits provided to all students. Please let your professors know as soon as possible if you have an accommodation that requires academic coordination and planning.

Speech Center: (speech.richmond.edu or 287-6409): Assists with preparation and practice in the pursuit of excellence in public expression. Recording, playback, coaching and critique sessions are offered by teams of trained student consultants. During scheduled appointments, consultants assist in developing ideas, arranging key points for more effective organization, improving style and delivery, and handling multimedia aids for individual and group presentations. We look forward to meeting your public speaking needs.

Writing Center (writing.richmond.edu or 289-8263): Assists writers at all levels of experience, across all majors. Students can schedule appointments with trained writing consultants who offer friendly critiques of written work.