Genomics Lab: Methods in Data Collection and Analysis
BSC4445C-RI
Fall 2023 (Savage)
4 credits

Course Description

The field of genomics focuses on understanding the collective function of all components encoded in an organism's genomic blueprint. In the past decade, there has been an explosion of new and cost-effective methodologies to sequence the genetic material of life. Originally, high-throughput sequencing was slow, costly, and used only to sequence the genomes of model organisms. Today genome-scale datasets are essential to most molecular biology research in any taxonomic group, including phylogenetics and population genetics, metagenomic sequencing of entire biological communities, functional genomics of the expressed portion of the genome, and whole-genome assembly and characterization of an organism's complete biological code. An integral, related, and emerging field of study is bioinformatics, which focuses on developing computational tools to analyze these massively large data sets. In this class, we take a hands-on approach to understanding (1) how genomic datasets are generated in the lab, and (2) how they are analyzed computationally using bioinformatics pipelines. We begin, in both lecture and lab, with the fundamental biochemistry of DNA and the tools biologists have developed to isolate and manipulate genomic elements. We then scale up to the "Next Generation" Sequencing (NGS) revolution, including how and why new methodologies have made genome-scale analyses achievable for nearly any organism, and the detailed methodologies and skills necessary to prepare samples for whole-genome, whole-transcriptome, and whole-metagenomic community sequencing. Finally, we spend the last third of the course learning and utilizing bioinformatics pipelines to manipulate and analyze the genomic datasets we have generated. Imparting a working knowledge of practical methods for generating and analyzing genomic datasets is the fundamental goal of this course. By generating, analyzing, and presenting (in both written and oral formats) novel genomic data, students will hone their writing, critical thinking and problem solving skills. The importance of genomics for all facets of life sciences will be emphasized, particularly the impacts of genomic datasets on recent advances in phylogenetics, evolutionary ecology, epigenetics, functional genomics, and heath and medicine.

Course Objectives

- To become a skilled molecular biologist actively engaging in scientific inquiry by conducting, in a group and with guidance, the (1) formation of a scientific hypothesis that utilizes genomic data, (2) maintenance of a laboratory notebook, and (3) successful performance of core molecular protocols, including:
    - purifying nucleic acids
    - quantifying nucleic acids
    - DNA amplification
    - quantitative DNA amplification
    - gel electrophoresis
    - cDNA library construction

    - library preparation for NGS of whole bacterial genomes, whole transcriptomes (RNAseq), and 16S gene transcripts for gene targeted characterization of microbial communities (i.e., metabarcoding).
- To understand the molecular biology behind existing methods for isolating, manipulating and sequencing DNA and RNA
- To understand the NGS revolution and the computational demands of modern molecular analyses
- To understand the types of research and analyses that fall under the umbrella of genomics, including data generation, manipulation, and bioinformatics
- To gain hands-on skills in how to manipulate genome-scale datasets by conducting analyses of group project data using basic computer programming and established bioinformatics pipelines for whole genome assembly and whole transcriptome assembly
- To gain skills in how to present and communicate research by (1) creating figures, tables, and a written narrative in the style of a peer-reviewed publication and (2) presenting results in the format of a 15-minute oral powerpoint presentation using data and analyses from semester-long group research projects

Prerequisites

A grade of B or better in undergraduate genetics or consent of the instructor. Completion of EHS Lab Safety Training online course (EHS102) and practical (EHS202) by end of add/drop (mandatory in order to conduct research in the Genomics Core Facility where the course is held).

Class Meetings

M/W 10:30 am – 2:20 pm
Lecture: 30-60 minutes at the start of class on Mondays, and occasionally on Wednesdays
Lab: remainder of session on Monday and Wednesday (often class will end early, but other weeks it will require additional time outside of class hours)

Research activities often do not fit perfectly into scheduled time blocks due to the trial-and-error nature of conducting protocols on new samples and testing hypotheses where outcomes are unknown. Thus, the Genomics Core Facility will also be open for additional hours every week for students to complete protocols or re-do failed samples, during which time the lab manager, GTA, and/or instructor will be present to answer questions and provide guidance. See office/lab hours listed below.

Course Instructor

Dr. Anna E. Savage
Office: Biology Building 424
E-mail: Anna.Savage@ucf.edu
Office/lab Hours: by appointment. I will typically remain within the Core lab after class officially ends to help students finish protocols or go over course material, and you can always look for me in 414 or my office (424), but to guarantee a 1-on-1 meeting please schedule with me over email.

Genomics Core Facility Manager
Jessica Fulsom
Office: Biology Building 435 (Hoffman Lab)
Jessica will help maintain the lab and reagents for students in the course and allow access to the lab on T/Th/F if needed.

Graduate Teaching Assistant
George Zaragoza
Office: Biology Building 135 (within the Genomics and Bioinformatics Cluster)
Email: George.Zaragoza@ucf.edu
Office hours: by appointment

Undergraduate Teaching Assistant
Emily Fackler
Email: Emily.Fackler@ucf.edu
Office: Biology Building 135 (within the Genomics and Bioinformatics Cluster)
Office hours: By appointment, and Thursdays 2-3pm on Zoom:
https://ucf.zoom.us/j/98826401021?pwd=KzE5VHloVzFiY2krNlJ0bURISzZVUT09
Meeting ID: 988 2640 1021
Passcode: Genomics23

Genomics Core Facility Lab Hours
The 414 lab manager, the GTA and UTA, and/or Dr. Savage will be available for additional access to 414 and help with protocols on non-class days, as needed. However, you will need to speak with and/or email someone in advance if you plan to work during these lab hours to make sure that someone is available to help you during a particular time. Note that other researchers use the lab on T/Th/Fri so you may need to share space and equipment during those times.

Webcourses Site
There is a course web site available through Webcourses (https://webcourses.ucf.edu) that I will use to post materials for the course, including the syllabus, calendar dates, PowerPoints, and grades.

Required Text
None

High-Impact Practice Course Statement

This section of **BSC4445C** is designated as a Research-Intensive (RI) course. This designation will be noted on your transcripts. Research-Intensive is one of the four High Impact Practice course designations at UCF, along with Service Learning, Integrative Learning, and Global Learning. High Impact Practice courses are some of the most challenging and rewarding at UCF. You will actively engage in research processes and a significant portion of your grade will be derived from course-related project(s) based on original research and/or creative scholarship.

If you have any questions about this designation or HIP designations at UCF, please contact hip@ucf.edu

Class Policies

1. Attendance is required for this hands-on class! A large portion of the grading for this research-based course is based on lab participation, and success in this class requires completion of lab procedures, thus there is minimal room for missing class. That being said, I understand that life happens and you may need to occasionally miss class for valid personal or professional reasons. I am happy to work with you on making up missed work if you do need to be absent for illness or an important reason. Please let me know as far in advance as possible about necessary absences, and if prolonged absences become necessary discuss with me strategies for how to proceed with this course.
2. Proper lab attire must be worn at all times in room 414. Lab coats and safety goggles will be provided as needed, but the ***key requirement is to wear long pants (or a long skirt) and closed-toe shoes during every class.***
3. Quiz and lab make ups are not guaranteed without documentation that is presented prior to the absence or within 24 hours afterwards. I will try to be flexible and accomondating, but the course moves fast, so please communicate with me ASAP if you miss anything.
4. Assigned readings should be completed before attending class and will be provided via webcourses or handed out in class.
5. You are encouraged to discuss any and all portions of the class with me. Please feel free to make an appointment with me to discuss the class or anything else on your mind.
6. Respect should be given to fellow students, the lab manager, the GTA and UTA, and the instructor. Please try to arrive on time and stay until the work is done. You are encouraged to leave the lab space during down time in between procedures to take a break, have a snack or drink, go to the restroom, etc. This is a long course session and there are always times for breaks, but please arrive on time as we do get started right away.
7. Hateful or offensive speech or writing will not be tolerated.
8. Academic dishonesty (cheating and plagiarism) is strictly prohibited and will be taken very seriously and will result at least in an "F" for that assignment (and may, depending on the severity of the case, lead to an "F" for the entire course) and may be subject to appropriate referral to the Office of Student Conduct for further action. See the UCF Golden Rule for further information.

Course Accessibility

It is my goal that this class be an accessible and welcoming experience for all students, including those with disabilities that may impact learning in this class. If anyone believes the design of this course poses barriers to effectively participating and/or demonstrating learning in this course, please meet with me to discuss reasonable options or adjustments. You may also contact SAS (Ferrell Commons 185; 407-823-2371; sas@ucf.edu) to talk about academic accommodations.

Respect for Diversity

It is my intent that students from all backgrounds and perspectives be well-served by this course, that students' learning needs be addressed both in and out of class, and that the diversity that students bring to this class be viewed as a resource, strength and benefit. It is my intent to present materials and activities that are respectful of diversity: gender, sexuality, disability, age, socioeconomic status, ethnicity, race, and culture. Your

suggestions are encouraged and appreciated. Please let me know ways to improve the effectiveness of the course for you personally or for other students or student groups. Specifically:

- If you have a name and/or set of pronouns that differ from those that appear in your UCF records, please let me know.
- If any of our class meetings conflict with your religious events, please let me know so that we can make arrangements for you.
- If you feel like your performance in the class is being impacted by your experiences outside of class, please don't hesitate to come and talk with me. I want to be a resource for you. If you prefer to speak with someone outside of the course, you can contact UCF's Office of Diversity and Inclusion (diverse@ucf.edu or 407-823-6479) or the Biology Department Chair, Dr. Graham Worthy (Graham.Worthy@ucf.edu).
- I am always in the process of learning more about diverse perspectives and identities, and recognize that I have limitations and blindspots. If something was said in class (by me or anyone else) that made you feel uncomfortable, please talk to me about it (or contact the resources listed above if you feel more comfortable).

Grading

There will be no traditional exams in this class – instead, there will be four in-class quizzes/assignments, each worth 5% of the total grade (20% total); in-class participation in research (30% of total grade for attending and successfully completeing lab protocols); maintaining a lab notebook documenting research to the standards of a professional molecular biologist (19% of total grade); a group research project final report written in manuscript format (19% of total grade) and presented as a 15-minute oral presentation (10%); plus completion of webcourses pre and post assignments (2%).

Grades will be assigned according to the following scale:
A: 90-100%; B: 80-89%; C: 70-79%; D: 60-69%; F: <60%

The grade for this course will be based on six components (presented out of 100% total):

(1) Four semester **quizzes** will be given in class on the dates indicated on the schedule (5% each; 20% total). They will consist of multiple choice (30-50 pts) and short answer questions and problem solving (50-70 pts) based on lecture material. Students will be assessed in their ability to understand genomics lab methodologies, the molecular biology behind how those methodologies work, and how genomic datasets are analyzed and interpreted.

(2) Students will be graded on **lab research participation** (30%), including asking questions, performing protocols, and generating the necessary data for each week to move forward in the next lab session. Extra lab time outside of the scheduled four hours will be necessary to complete protocols for the majority of labs. Make-ups during office hours will be offered for up to two missed lab sessions; additional missed lab sessions will each result in a 5% drop in participation grade.

(3) The **lab notebook** (19%)**,** to be maintained by each group of three students working on a unique project, will be evaluated every third week and given a letter grade. Students will be given feedback on how to improve their note taking and record keeping of all lab activities and results. At any time before lab notebooks must be turned in, students can also check with the instructor to determine whether lab notebook content is sufficient or lacking, and the instructor will provide feedback. The final lab notebook grade will be based on both the biweekly grades and improvement over the course of the semester to emphasize the trajectory of learning gains. The lab notebook record keeper will rotate weekly among the three group members during the nine weeks of molecular lab, and each student will be graded on their own three entries.

(4) Each group will submit a **final report** (19%; **group receives one grade**) detailing the results of their research project. The report must be written in the format of a scientific journal article, including the following sections: Abstract (5 pts) Introduction (10 pts), Methods (25 pts), Results (with figures/tables; 35 pts) and Discussion (10 pts). Groups must turn in drafts of the introduction, methods and results sections on the dates listed in the schedule below. Drafts will be evaluated and the instructor will give feedback on how to make improvements before the final draft is due. Groups will receive 15 pts of their overall final report grade for turning in drafts of these sections (5 pts for each deadline and section).

(5) During the final exam period, each group will give a 15-minute **oral presentation** (10%) on their project, including powerpoint slides illustrating their research. The presentation must include the following sections: overview (10 pts), introduction (20 pts), methods (30 pts), results (30 pts), and conclusions (10 pts). Each student will speak for 5 minutes and be graded individually.

(6) Week one attendance assignment on Webcourses to gather data and determine semester groups (1%) and one post-survey at the end of the semester (1%)

Schedule:

The following schedule is approximate and dates may be changed at any time.

Week 1:
Lecture 1: Intro and overview of course, overview of overarching genomics projects and methods to be used (RNAseq)
Molecular Lab 1: Keeping a lab notebook, lab safety and etiquette, and pipetting.

Week 2:
Lecture 2: Basics of molecular biology; isolating and characterizing nucleic acids
Molecular Lab 2: Polymerase Chain Reaction and PCR visualization with gel electrophoresis.

Week 3:
Lecture 3: Polymerase Chain Reaction (PCR) and Quantitative (q)PCR.
Molecular Lab 3: RNA cleanup and removal of DNA contamination.

Week 4:
Lecture 4: Genomic and cDNA libraries.
Molecular Lab 4: mRNA extractions. First and second-strand cDNA synthesis.

Week 5:
Lecture 5: Overview of Sanger and next-generation sequencing. ***Quiz 1 (Wed).***
Molecular Lab 5: Illumina library prep week one: DNA/cDNA fragmentation; visualization on Tapestation. ***Turn in lab books for evaluation.***

Week 6:
Lecture 6: Overview of Next Generation Sequencing technologies
Molecular Lab 6: Illumina library prep week two: end-repair via blunting; bead-based DNA/cDNA clean-up (group-specific); dC-tailing; bead-based clean-up.

Week 7:
Lecture 7: Overview of Illumina library prep. ***Quiz 2 (Wed)***
Molecular Lab 7: Illumina library prep week three: Ligation of Nextera sequencing adapters using high-fidelity low-cycle PCR with Phusion taq; bead-based clean-up.; bead-based clean-up; gel electrophoresis.

Week 8:
Lecture 8: Genomic analysis 1: Quality control, NCBI, GenBank, and BLAST
Molecular Lab 8: Illumina library prep week four: qPCR on serial dilutions of libraries to precisely quantify the amount of available material for sequencing.

Week 9:
Lecture 9: Genomic analyses 2: Bioinformatics overview
Molecular Lab 9: Pool libraries within working groups for whole-genome and transcriptomics projects, QC using TapeStation; equimolar pooling of libraries, final prep of libraries to be sent to core facility for sequencing. ***Turn in lab books for evaluation.***

Week 10:
Lecture 10: Whole genome assembly, genomes and genome browsers
Computer Lab 1: Intro to Geneious and gene identification using GenBank

Week 11:
Lecture 11: Genome-scale analyses: RadSeq, SNPs and GWAS. ***Quiz 3 (Wed)***
Computer Lab 2: basics of command line and perl; bioinformatics pipeline tutorials (Trinity/Tuxedo or Geneious whole-genome assembly, depending on project) using sample data. **METHODS DRAFT DUE (Wed)**

Week 12:
Lecture 12: Microarrays, transcriptomics, and RNAseq.
Computer Lab 3: Manipulating Illumina data and Quality Control (QC) using Trimmomatic; start running pipeline on lab-generated data (students work in groups).
INTRO DRAFT DUE (Wed)

Week 13:
Lecture 13: *Metagenomics and metabarcoding*
Computer Lab 4: Continue running pipelines on lab-generated data; analyze results/generate summary statistics (students work in groups). **RESULTS DRAFT DUE (Wed)**

Week 14:
Lecture 14: Presenting genomic data. ***Quiz 4 (Wed)***
Computer Lab 5: Finish/troubleshoot lab-generated data analysis; finish analyzing results and work on tables, figures and reports (students work in groups).

Week 15: Work on figures, final report and final presentation (no lecture, but attendance is required)
Computer Lab 6: Work on final report and presentation

Week 16:
Final Exam period: Group presentations of data analysis and results. ***Final reports due***.

Schedule:
The following schedule is approximate and may be changed during the semester

Week	Date	Topic
1	21 Aug M	Class intro ~ lab project intro ~ Pipetting and dilutions
	23 Aug W	Group project assignments ~ pipetting practice
2	28 Aug M	Molecular biology ~ Run PCR
	30 Aug W	Agarose gel visualization of PCR product
3	4 Sept M	**Labor Day – no class**
	6 Sept W	PCR/qPCR methods ~ RNA cleanup.
4	11 Sept M	Genomic and cDNA libraries ~ mRNA capture
	13 Sept W	cDNA synthesis.
5	18 Sept M	Sequencing technologies I ~ Illumina library prep week 1
	20 Sept W	Continue Illumina library prep as needed. ***Quiz 1. Lab notebooks due.***
6	25 Sept M	Sequencing technologies II ~ Illumina library prep week 2
	27 Sept W	Continue Illumina library prep as needed.
7	2 Oct M	Illumina library preparation overview ~ Illumina library prep week 3.
	4 Oct W	Continue Illumina library prep as needed.

8	9 Oct M	Analyzing sequence data and genetic databases ~ Illumina library prep week 4
	11 Oct W	Continue Illumina library prep as needed. ***Quiz 2.***
9	16 Oct M	Bioinformatics overview ~ troubleshooting/catch-up week
	18 Oct W	Troubleshooting/catch-up week. ***Lab notebooks due.***
10	23 Oct M	Genomes, genome browsers, SNPs & GWAS ~ Intro to Geneious & gene identification using GenBank
	25 Oct W	Work on tutorials
11	30 Oct M	Genome-scale analyses: Radseq & sequence capture ~ command line and pipeline tutorials
	1 Nov W	Finish pipeline tutorials. ***Quiz 3.***
12	6 Nov M	Microarrays, transcriptomics and RNAseq ~ QC and manipulation of Illumina data **METHODS DRAFT DUE**
	8 Nov W	Begin analyzing project data
13	13 Nov M	Metagenomics and metabarcoding ~ Continue analysis of project data **INTRO DRAFT DUE**
	15 Nov W	Continue analysis of project data
14	20 Nov M	Presenting genomic data ~ Compete analysis of project data. ***Quiz 4.*** **FULL DRAFT DUE**
	22 Nov W	***No class – Thanksgiving Wednesday***
15	27 Nov M	Work on final report and presentation (no lecture but attendance required)
	29 Nov W	Work on final report and presentation (practice sessions available with instructor / GTA / UTA). **FINAL REPORT DUE**
16	4 Dec M	***FINAL PRESENTATIONS***