

Genomics Lab: Methods in Data Collection and Analysis
BSC4445C-RI
Fall 2018 (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 health 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 and practical 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 9:30 – 12:20

Lecture: 30-60 minutes at the start of class on Mondays and occasionally on Wednesdays

Lab: remainder of session on Monday and Wednesday

Research activities often do not fit perfectly into scheduled time blocks due to the trial-and-error nature of conducting protocols on novel samples and hypotheses where outcomes are unknown. Thus, the Genomics Core Facility will also be open for an additional 12 hours per 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.

Lecture Instructor

Dr. Anna E. Savage

Office: BL 424, 407-823-4504

E-mail: Anna.Savage@ucf.edu

Office/lab Hours: M/W 12:30 – 3:00 and by appointment. (Anna will typically remain within the Core lab during the scheduled office hours to help students finish protocols or go over course material, but she may also be in her office, so be sure to seek her out in both locations, BL 414 and 424).

Genomics Core Facility Manager

Dr. Anna M. Forsman

Office: BL 414, 407-823-0766

Dr. Forsman will help maintain the lab and reagents for students in the course and allow access to the lab on T/Th (see policies regarding extra lab hours below).

Graduate Teaching Assistant

Ian Will

Office: BL 436

Email: ian.will@knights.ucf.edu

Office/lab hours: T/Th by appointment (see policies regarding extra lab hours below).

Undergraduate Teaching Assistant

Christian Polania

Office: BL 436

Email: cmpolania@knights.ucf.edu

Office/lab hours: By appointment (see policies regarding extra lab hours below).

Genomics Core Facility Lab Hours

Dr. Forsman will be based in 414 on Tuesdays and Thursdays 10:00am – 4:00pm and the GTA will also be available on these days. This will allow students to come and work independently on their projects as needed. However, these times are subject to changes during the semester and 414 may be locked if you do not notify one of us in advance. Please tell or email Dr. Forsman, Dr Savage and/or the GTA in advance if you plan to work during these lab hours to make sure that she is available to help you during that particular time. This will avoid you having to wait for assistance if Dr. Forsman is in a meeting or working with another student.

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

Class Policies

1. Attendance is not strictly required but many studies have shown that students who do not attend class do poorly. In addition, a large portion of the grading for this hands-on methodology course will be based on lab participation, and success in this class requires completion of lab procedures.
2. Quiz and lab make ups will not be given without valid documentation that is presented prior to the absence or within 24 hours of lab or administration of the quiz.
3. Assigned readings should be completed before attending class and will be provided via webcourses or handed out in class.
4. You are encouraged to discuss any and all portions of the class with me. Please feel free to come to my office hours or make an appointment to discuss the class, especially if you are having trouble.
5. Respect should be given to fellow students, the lab manager, and the instructor. Please do not arrive late to class or lab or leave early.
6. Hateful or offensive speech or writing will not be tolerated.
7. Cell phones, iPods, and other electronic devices should be silenced and put

away before class starts.

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.

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 completing lab protocols); maintaining a lab notebook documenting research to the standards of a professional molecular biologist (20% 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 assignments (1%).

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** (20%), 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 determine semester groups (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 two overarching genomics projects (whole-genome sequencing and RNAseq) and methods to be used

Molecular Lab 1: Keeping a lab notebook, lab safety and etiquette, and pipetting. **Serial dilution qPCR test to assess pipetting skill level. Before students are allowed to continue with lab protocols, they must achieve an R^2 value > 0.90 for their serial dilutions.**

Week 2:

Lecture 2: Basics of molecular biology; isolating and characterizing nucleic acids

Molecular Lab 2: DNA extractions (genomics groups) or mRNA extractions (RNAseq groups). Quantitative and qualitative analysis of DNA and RNA using TapeStation. First and second-strand cDNA synthesis.

Week 3:

Lecture 3: Polymerase Chain Reaction (PCR) and Quantitative (q)PCR.

Molecular Lab 3: Polymerase Chain Reaction - all groups amplify viral DNA. **Turn in lab books for evaluation.**

Week 4:

Lecture 4: Genomic and cDNA libraries. **Quiz 1 (Wed).**

Molecular Lab 4: PCR visualization with gel electrophoresis. Quantitative (q)PCR.

Week 5:

Lecture 5: Overview of Sanger and next-generation sequencing

Molecular Lab 5: Illumina library prep week one: DNA/cDNA fragmentation using sonicator; visualization on TapeStation

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. **Turn in lab books for evaluation.**

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:

Monday, Dec.3: **Final Exam**: Group presentations of data analysis and results. **Final reports due.**

Schedule:

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

Week	Date	Topic
1	20 Aug M	Class intro ~ lab project intro ~ Pipetting and dilutions
	22 Aug W	Group project assignments ~ pipette test results and re-dos
2	27 Aug M	Molecular biology ~ DNA and RNA extractions
	29 Aug W	Visualization of DNA/RNA on Tapestation ~ cDNA synthesis
3	3 Sept M	Labor Day – no class
	5 Sept W	PCR/qPCR ~ Run PCR on bacterial genomes. Lab notebooks due.
4	10 Sept M	Genomic and cDNA libraries ~ Agarose gel visualization of PCR product
	12 Sept W	qPCR. Quiz 1
5	17 Sept M	Sequencing technologies I ~ Illumina library prep week 1
	19 Sept W	Continue Illumina library prep as needed.
6	24 Sept M	Sequencing technologies II ~ Illumina library prep week 2
	26 Sept W	Continue Illumina library prep as needed. Lab notebooks due.
7	1 Oct M	Illumina library preparation overview ~ Illumina library prep week 3
	3 Oct W	Continue Illumina library prep as needed. Quiz 2.

8	8 Oct M	Analyzing sequence data and genetic databases ~ Illumina library prep week 4
	10 Oct W	Continue Illumina library prep as needed.
9	15 Oct M	Bioinformatics overview ~ troubleshooting/catch-up week
	17 Oct W	Troubleshooting/catch-up week. Lab notebooks due.
10	22 Oct M	Genomes, genome browsers, SNPs & GWAS ~ Intro to Geneious & gene identification using GenBank
	24 Oct W	Work on tutorials
11	29 Oct M	Genome-scale analyses: Radseq & sequence capture ~ command line and pipeline tutorials
	31 Oct W	Finish pipeline tutorials. Quiz 3.
12	5 Nov M	Microarrays, transcriptomics and RNAseq ~ QC and manipulation of Illumina data
	7 Nov W	Begin analyzing project data METHODS DRAFT DUE
13	12 Nov M	Metagenomics and metabarcoding ~ Continue analysis of project data
	14 Nov W	Continue analysis of project data INTRO DRAFT DUE
14	19 Nov M	Presenting genomic data ~ Complete analysis of project data. Quiz 4.
	21 Nov W	Complete analysis of project data FULL DRAFT DUE
15	26 Nov M	Work on final report and presentation (no lecture but attendance required)
	28 Nov W	Work on final report and presentation (practice sessions available with instructor / GTA)
16	3 Dec M	*FINAL PRESENTATIONS*