Tentative Syllabus
BIOMI/CS 609
Computational Genomics and Bioinformatics
Spring 2022
San Diego State University
Mondays and Fridays from 8 AM – 9:15 AM in LSN 126

Instructor: Dr. Arun Sethuraman
Pronounced: (*Uh-Roon Say-thu-Raa-mun*)
Office: LSS 250
Office phone: 619-594-3979
Office Hours: Tuesday 2 PM-4 PM, **or by appointment**
E-Mail: asethuraman@sdsu.edu

COVID-19 Policies:
Please note that while we will try our best to meet in person through Spring 22, we are still in the midst of a global pandemic, and everyone needs to be aware of the volatility around in-person meetings. In the event of an exposure to COVID-19, we might have to switch to online instruction via Zoom. If this occurs, please watch your Canvas/SDSU emails and I will let you all know about it ASAP. Everyone – that includes me will adhere to Campus/State/County/City masking and vaccine + booster requirements. No exceptions. Please help us all have a successful and fun science-filled semester by strictly adhering to these regulations.
Please refer to the full campus policy regarding COVID-19 here: https://sacd.sdsu.edu/student-rights/covid-policies

Textbook/Readings:
There is no textbook for this class. All lecture notes, readings will be posted on a regular basis on Canvas.

Course Description:
BIOMI/CS 609 Computational Genomics and Bioinformatics

Genomics is shaping everything we do and everything we think about. From the role of genomics in medicine to the ready availability of personal genomics to the consumer, understanding the bioinformatics behind DNA sequence analysis is critical. This course covers computational algorithms in genomics, and their applications to biological questions. Topics covered will include genome assembly, annotation, variant calling, phylo- and population genomics, and genome-wide association studies.

Three hours of lecture+lab. Prerequisite: CS 503 or 514, or enrollment in Master of Science in Biological and Medical Informatics program, or PhD programs in Evolutionary Biology, Computer Science.

Course Learning Outcomes:
By the end of this class, every student should be able to:

1. Describe different types of biological data (genetic, genomic, proteomic) and how to analyze them.
2. Understand the fundamental algorithms/statistics behind genomic data analyses.
3. Use corresponding computational tools (online and offline) to analyze, interpret, and visualize genomic data.
4. Write programs/tools for genomics in Python, Unix, and R.
5. Inquiry-based learning all along in the labs – you will be provided with some background information, but the onus is on you to build a hypothesis, test it during lab, and summarize your results.

Assessments:
Outcomes 1, 3: Bi/triweekly assignments and lab reports.
Outcome 2: In-class problems, take-home assignments, midterm exams, final exam
Outcomes 4, 5: Lab exercises in programming, take-home assignments

Activities:
Outcome 1: Students are expected to practice analyses discussed in class on their own, during lab, and after class on a regular basis. The only way to learn genomics is by doing. While the background reading provides you some basic information, **you will not learn anything if you don't do it yourself**.
Outcome 2: Practice problems, work out assignments, work in groups, if need be.
Outcome 3: Come to class, work on learning methods during lab, work separately on data analyses.
Outcomes 4, 5: **Practice, practice, practice! I can't stress this enough. You are signed up for a class in genomics – you will learn to code. Period. If you don't want to learn coding, or don't care about it, please drop the class. This is not a class on just applying existing tools, you will write some tools yourself.**

Course Grading (Lecture):
There will be two midterm exams, and one final exam, each worth 100 points. The final exam will be comprehensive. All exams will be essay format and will test your ability to (1) understand and assimilate concepts in genomics, (2) apply these to real-life problems. Students will also be assigned weekly/biweekly homework assignments (100 points in all) that will include analyses of real genomic data, and at least one problem that will tie into the "big picture" involving applying concepts learned to real-life problems. You are welcome to type/transcribe/scan/write, but regardless, these problem sets will be due in class, and no email submissions will be accepted. Appropriate rubrics for assignments and exams will be posted regularly.

Course Grading (Lab):
Lab periods (usually Fridays) will be interspersed (bi- or tri-weekly, 6 in all) with programming/data analysis assignments and reports, each worth 50 points. Additionally, there will be weekly quizzes, woth 10 points each. All code written has to be annotated with appropriate comments, and testable with my standard "key" solutions. Each assignment will also be accompanied by a small lab report, that summarizes (1) biological question at hand, (2) hypotheses, (3) methods to test the hypotheses, (4) results, (5) discussion of results. These

assignments will be due every other week, on the morning of class, prior to lab. We will use an online lab notebook to facilitate timely feedback and ease of use. Students are required to follow instructions from me on how to write and submit these reports, details of which are posted on Canvas. In addition to sharing lab reports on a regular basis, students **WILL ALSO HAVE TO SUBMIT THE SAME REPORT THROUGH TURNITIN ON Canvas**. No exceptions will be accepted to this rule. This will ensure accountability on each student against plagiarism. Comments on all lab reports will be provided on a weekly/biweekly basis. Lab grade will be scaled out of 150 to compute the final grade

Your overall grade (lecture + lab = 400 + 150) will be guaranteed as:

A with 95% of the points
A minus with 90% of the points
B with 85% of the points
B minus with 80% of the points
C with 75% of the points
C minus with 70% of the points
D with 65% of the points
D minus with 60% of the points

There will be NO MAKEUP tests given, without prior approval from me. I like relative grading, and so there might be a “curve” at the end of the semester, as need be.

Course Schedule – Lectures

Week 1	Introduction to genomic sequencing technologies, datatypes
Week 2	Base Calling, Quality Control
Week 3	Genome assembly – *de novo*, guided
Week 4	Review + Midterm 1
Week 5	Alignment, BLAST
Week 6	Annotation – *ab initio*, Hidden Markov Models
Week 7	Phylogenomics (Neighbor Joining, Parsimony, Likelihood)
Week 8	mRNAseq Analyses
Week 9	Review + Midterm 2
Week 10	Variant Calling
Week 11	Population Genomics – 1 (HWE, LD)
Week 12	Population Genomics – 2 (Structure, Evolutionary History)
Week 13	Bacterial Community level sequencing – 1 (16s rRNA analyses with MOTHUR)
Week 14	Bacterial Metagenomics – 2 (Metabarcoding, shotgun analyses with Q2)
Week 15	Human genomics - GWAS
Week 16	**Final Exam**

Course Schedule – Labs:

Week 1	Databases, exploring genomic data, programming basics
Week 2	Base calling, FastQC
Week 3	Genome assembly – VELVET, SoapdeNovo/AllPaths-LG
Week 4	Assignment 1 – Assess quality, assemble a bacterial genome
Week 5	BLASTx, Mview
Week 6	Annotation – PGAP, Augustus
Week 7	Phylogenomics – RaxML, ASTRAL
Week 8	RNAseq analyses – Trinity
Week 9	Assignment 2 – Annotate, build species tree from a bacterial genome
Week 10	Variant calling (GATK)
Week 11	GWAS
Week 12	Population genomics – BCFTools, PPP
Week 13	Population genomics – ADMIXTURE, IMa3
Week 14	MOTHUR/QIIME-2
Week 15	Review

Credit Hour Policy Statement:
For the lecture part of the class, each unit of credit corresponds to an "hour" of class-time and two hours of student learning outside of class. Therefore, just for the lecture portion of the class you should expect to spend 3 hours a week in class and 6 additional hours outside of class.

ADA Statement:
If you are a student with a disability and believe you will need accommodations for this class, it is your responsibility to contact Student Ability Success Center at (619) 594-6473. To avoid any delay in the receipt of your accommodations, you should contact Student Ability Success Center as soon as possible. Please note that accommodations are not retroactive, and that I cannot provide accommodations based upon disability until I have received an accommodation letter from Student Ability Success Center. Your cooperation is appreciated.

All-University Writing Requirement:
Will be satisfied through the lab reports that will be due every two or three weeks.

Academic Honesty Statement:
Students will be expected to adhere to standards of academic honesty and integrity, as outlined in the Student Academic Honesty Policy. All assignments must be original work, clear and error-free. All ideas/material that are borrowed from other sources must have appropriate references to the original sources. Any quoted material should give credit to the source and be punctuated accordingly.

Turnitin® or other plagiarism detection software:
Students agree that by taking this course all required papers may be subject to submission for textual similarity review to Turnitin.com or another similar website or software database for the detection of plagiarism. All submitted papers will be included as source documents in the

Turnitin.com reference database (or another similar database) solely for the purpose of detecting plagiarism of such papers. You may submit your papers in such a way that no identifying information about you is included. Another option is that you may request, in writing, that your papers not be submitted to Turnitin.com. However, if you choose this option you will be required to provide documentation to substantiate that the papers are your original work and do not include any plagiarized material. If you do not agree to this, please see Dr. Sethuraman as soon as possible.

Academic Honesty and Integrity:
Students are responsible for honest completion and representation of their work. Your course catalog details the ethical standards and penalties for infractions. There will be zero tolerance for infractions. If you believe there has been an infraction by someone in the class, please bring it to the instructor's attention. The instructor reserves the right to discipline any student for academic dishonesty, in accordance with the general rules and regulations of the university. Disciplinary action may include the lowering of grades and/or the assignment of a failing grade for an exam, assignment, or the class as a whole.
It is recommended that students be referred to the full Academic Honesty Policy at https://newscenter.sdsu.edu/student_affairs/srr/cheating-plagiarism.aspx

The syllabus is tentative and is subject to change depending on how far we get each day. The problems will be due after we have discussed that topic. I will give you updates each week as to what problems will be due.