

Last Updated Fall 2024

**San José State University Computer
Science Department
CS 224: Next Generation Sequencing & Genome Assembly Fall 2024**

Course Information

Instructor:	Leonard Wesley
Office Location:	MH 212
Telephone:	408.924.5287 (Office, however, I will not be on campus much during the Fall 2024 semester.)
Email:	Leonard.Wesley@sjsu.edu
Office Hours:	Tuesdays 7:00 AM – 9:00 AM, Zoom Link For Fall 2024: https://sjsu.zoom.us/j/82185188270?pwd=UYdaagroNO8GC9lhQB1W1sZh7YzRPO.1 PASSCODE: 279456
Class Days/Time:	Section 01: Tuesdays and Thursdays 4:30 PM – 5:45 PM
Classroom:	ISB 876
Prerequisites:	Biol 135A or equivalent, CS146 or equivalent, Math 161A or equivalent, and be comfortable running UNIX, Windows, or Mac based application software, or instructor consent.

Course Description

Next generation sequencing (NGS) is a high-throughput method used to determine a portion of the nucleotide sequence of a biological organism's genome. NGS techniques utilize DNA and RNA sequencing technologies that are capable of processing multiple genomic sequences in parallel. This course will provide the student with a thorough understanding of the genomic landscape, description of various sequence generation methodologies and technologies (e.g., Illumina, Ion Torrent, Pac Bio, and Oxford Nanopore). The course will also provide instruction on how to perform basic quality control assessment of next generation sequencing data, and how to use next generation sequencing data to perform *de novo* and comparative assemblies of selected genomes and meta-genomes. Students will become familiar with genome annotation techniques, variant calling, and cloud services for bioinformatic analysis of next generation sequencing data.

Learning Outcomes

Upon successful completion of this course, students will:

1. **SLO-1: De Novo Genome Assembly:** The theory, method, and practice of de novo genome assembly. DeBruijn Graphs, and using Spades to assembly genomes from read datasets.
2. **SLO-2: Single Cell Sequencing Methods:** SmartSeq2, SmartSeq, STRT, CEL-Seq...and others
3. **SLO-3: SCS Applications:** Understanding human disease mechanisms, genetic variation, therapeutic screening ...
4. **SLO-4: SCS Computational Challenges:** Computational tools, algorithms, methods used to carry out differential expression, trajectory analysis
5. **SLO-5: Insights into T-cell Receptors via SCS:**

Each SLO above corresponds to a learning module that is described in the course calendar below.

Required Texts

Yanagida, T., Ishii, Y. (eds.) Single Molecule Dynamics in Life Science 2008
ISBN: 978-3-527-31288-7

Miodrag Gužvić et al. Methods in Molecular Biology ISBN 978-1-0716-3620-6 ISBN
978-1-0716-3621-3 (eBook) <https://doi.org/10.1007/978-1-0716-3621-3>

Xiangdong Wang et al., Translational Bioinformatics ISBN 978-94-017-9752-8
ISBN 978-94-017-9753-5 (eBook) DOI 10.1007/978-94-017-9753-5

Guo-Cheng Yuan, Computational Methods for Single-Cell Data Analysis, ISBN 978-1-4939-9056-6
ISBN 978-1-4939-9057-3 (eBook) <https://doi.org/10.1007/978-1-4939-9057-3>

NOTE: The field is advancing so rapidly that the above required textbook will be supplemented with more recent publications as appropriate.

Other Optional Reading Material

A Primer of Genome Science, Greg Gibson, Spencer V. Muse, Publisher Sinauer Associates, 2009, Edition #3, ISBN-10: **0878932364** | ISBN-13: **978-0878932368**

Introduction to Computational Biology: Maps, Sequences and Genomes, Michael S. Waterman, CRC Press. (A statistical oriented view of bioinformatics)

Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins, Andreas D. Baxevanis and B.F. Francis Ouellette, John Wiley & Sons 2nd Ed. (Includes contributions from several authors providing a wide perspective)

Course Requirements and Assignments

Course Logistics

Students should expect to spend approximately nine (9) hours per week (on average) outside of the classroom preparing for and completing the assigned course work. This includes reading papers, viewing videos as appropriate, completing homework and programming exercises, and so forth. The amount of time that a student actually spends will depend on individual skills and the time allocated to the course. The nine (9) hours per week estimate is based on previous experiences of the instructor and students. So please plan and schedule accordingly.

Some of the class will be taught in “flipped” mode where lectures will be available online, and students will be expected to view these lectures **BEFORE** class, as a regular part of their out-of-class work. Classes will concentrate on answering questions and performing exercises that allow students to practice and use the skills, tools and concepts covered in the lectures. Students should consult the Canvas website at least twice weekly and complete assignments by the specified deadlines.

Previously, students have asked for special exception to policies and procedures for this course. An example includes asking the instructor for extra assignments or work to help improve a grade. Even if such a request is reasonable in the opinion of the instructor, no exception will be given to a student unless it can be made available to the entire class, AND does not constitute significant extra work on the part of students, instructors, graders and so forth. Students should have no concern that other students will receive special exceptions that will not be available to the entire class.

NOTE: [University policy F69-24](http://www.sjsu.edu/senate/docs/F69-24.pdf) at <http://www.sjsu.edu/senate/docs/F69-24.pdf> states that “Students should attend all meetings of their classes, not only because they are responsible for material discussed therein, but because active participation is frequently essential to insure maximum benefit for all members of the class. Attendance per se shall not be used as a criterion for grading.” However, attendance will be required in order to complete and submit many in-class exercises, quizzes, and exams.

Success in this course is based on the expectation that students will spend, for each unit of credit, a minimum of 45 hours over the length of the course (normally three hours per unit per week) for instruction, preparation/studying, or course related activities, including but not limited to internships, labs, and clinical practica. Other course structures will have equivalent workload expectations as described in the syllabus.

Classroom Protocol

Instruction will begin at or within several minutes of the official published start time for the course. Please make sure that cell phones, beepers, and texting devices are turned off during the entire scheduled class time. Excessive audible discussions with fellow students is prohibited so that others are not disturbed. If any subject matter is not understood, please do not hesitate to ask for clarification. If an extended response is necessary to remove doubts, then a request to follow up outside of scheduled classroom instruction time might be made.

Quizzes and Exams

There will be three quizzes, one midterm and a final exam all of which will count toward the final grade as specified in the “Grades” section below. During quizzes and exams, communication with other individuals via any means is strictly prohibited without the express permission of the instructor. Violations will be met with the full impact of SJSU’s academic integrity policy and procedures.

Projects

Several life science or genomics-based projects will be described near the start of the course. Projects will involve applying the skills and knowledge learned in the course to the project. Teams of 2-3 students will be formed to work on a selected project topic. Teams will be required to submit a project proposal before starting on a project, and submit a project report along with working code at the end of the course. Individual student scores on a project will be determined by the content and quality of the contribution of each student toward the project. The score on the course project and project presentation will count toward the final grade (percentage wise) as specified in the “Grades” section below.

Reading, Homework, Programming, In-Class Exercises, Participation

Assignments

Graded reading, homework, programming, and class participation and brief course feedback assignments will be given almost weekly, and will count toward the final grade. There will be 4 In-class Exercise sessions. These will typically involve forming teams of 2-3 students that work on assigned exercises in the classroom. They provide an opportunity to get started on homework programming assignments that are to be submitted on a designated due date. Participation is mandatory, and scores will count toward final grade.

Computational Resources

Students are required to make sure that they have access to sufficient UNIX, Windows, or Mac based computational resources (e.g., computers and software) to carryout assignments in the course. An attempt to offer the course in a classroom with sufficient computation resources will be made by the department to support classroom instruction and demonstrations. However, students should be prepared to bring their portable laptops to class.

Week and Class Mtg #	Tue	Thur	Module # & Name	TOPIC	Assignment See Canvas For Module & Weekly Assignment Details and Due Dates
Week 1	N/A	8/22	#1 De Novo Genome Assembly	8/22: <ul style="list-style-type: none"> - Course Intro, Class background survey/skills assessment - Theory, method, and practice of de novo genome assembly. 	Module #1
Week 2	8/27	8/29	#1 De Novo Genome Assembly	8/27: <ul style="list-style-type: none"> - DeBruijn graphs 8/29: <ul style="list-style-type: none"> - DeBruijn graphs 	Module #1
Week 3	9/3	9/5	#1 De Novo Genome Assembly	9/3: <ul style="list-style-type: none"> - DeBruijn graphs 9/5: <ul style="list-style-type: none"> - Quartz, STRT, CEL-Seq 	Module 12
Week 4	9/10	9/12	#1 De Novo Genome Assembly	9/10: <ul style="list-style-type: none"> - Using Spades to assemble genomes 9/12: <ul style="list-style-type: none"> - In-Class Exercise 1 Topics Covered Week-1 to Week-3:	Module #1

Week 5	9/17	9/19	#1 De Novo Genome Assembly	<p>9/17:</p> <ul style="list-style-type: none"> - Using Spades to assemble genomes <p>2/19:</p> <ul style="list-style-type: none"> - Using Spades to assemble genomes 	<p>Module #1</p> <p>September 17, 2024 Last Day To Add/Drop Classes</p>
Week 6	9/24	9/26	#2 SCS Methods	<p>9/24:</p> <ul style="list-style-type: none"> - SCS Methods and Techniques - Challenges To SCS Methods <p>9/26:</p> <ul style="list-style-type: none"> - Quiz 1 (~45mins) Covers Topics rom Week 1 thru Week 5 	Module #2
Week 7	10/1	10/3	#2 SCS Methods	<p>10/1:</p> <ul style="list-style-type: none"> - Challenges To SCS Methods (cont. - Human disease mechanisms <p>10/3:</p> <ul style="list-style-type: none"> - In-Class Exercise 2 <p>Topics Covered Week-4 to Week-6</p>	Module #2
Week 8	10/8	10/10	#3 SCS Applicati ons	<p>10/8:</p> <ul style="list-style-type: none"> - Genetic variations (cont.) <p>10/10:</p> <ul style="list-style-type: none"> - Therapeutic Screening (cont.) 	Module #3

Week 9	10/15	10/17	#3 SCS Applications	<p>10/15:</p> <ul style="list-style-type: none"> - Therapeutic Screening (cont.) <p>10/17:</p> <ul style="list-style-type: none"> - Midterm (Full period): Covers Topics Week 1 thru Week 8 	Module #3
Week 10	10/22	10/24	#3 SCS Applications	<p>10/22:</p> <ul style="list-style-type: none"> - Therapeutic Screening (cont.) <p>10/24:</p> <ul style="list-style-type: none"> - Therapeutic Screening (cont.) 	Module #3
Week 11	10/22	10/24	#4 SCS Computational Challenges	<p>10/22:</p> <ul style="list-style-type: none"> - SCS computational tools and algorithms for differential expression & trajectory analysis (cont.) <p>10/24:</p> <ul style="list-style-type: none"> - SCS computational tools and algorithms for differential expression & trajectory analysis (cont.) 	Module #4
Week 12	11/5	11/7	#4 SCS Computational Challenges	<p>11/5:</p> <ul style="list-style-type: none"> - SCS computational tools and algorithms for differential expression & trajectory analysis (cont.) <p>11/7:</p> <ul style="list-style-type: none"> - Quiz 2 Covers Topics From Quiz 1 thru Week 11 	Module #4
Week 13	11/12	11/14	#4 SCS Computational Challenges	<p>11/12:</p> <ul style="list-style-type: none"> - SCS computational tools and algorithms for differential expression & trajectory analysis (cont.) <p>11/14:</p> <ul style="list-style-type: none"> - In-Class Exercise 3 	Module #4
Week 14	11/19	11/21	#5 SCS T-Cell Receptors	<p>11/19:</p> <ul style="list-style-type: none"> - SCS T-Cell Receptors <p>11/21:</p> <ul style="list-style-type: none"> - SCS T-Cell Receptors (cont.) 	Module #5

Week 15	11/26	11/28	#5 SCS T-Cell Receptors	11/26: - In-Class Exercise 3 Work on Team Projects, Q&A) 11/28: THANKSGIVING HOLIDAY	Module #5
Week 16	12/3	12/5	#5 SCS T-Cell Receptors	12/3: - SCS T-Cell Receptors (cont.) 12/5: Quiz 3 (~45 mins): Covers From Quiz 2 thru Week 16	Module #5
			Final Project Code and Project Report Due To Canvas Dec 11, 2024 By 11:59PM No Final Exam. The Project Takes The Place Of The Final Exam		

SCHEDULE FOOTNOTES:

NONE AS OF Fall 2024

Grades *

WRITTEN HOMEWORK (4 at 10 points each)	40 pts
QUIZZES (3 at 40pts each)	120 pts
MIDTERM	100 pts
IN-CLASS EXERCISES (4 at 50pts each)	200 pts
WEEKLY COURSE FEEDBACK (12 at 5pts each)	60 pts
PROGRAMMING ASSIGNMENTS (2 @ 40pts each)	80 pts
FINAL EXAM & PROJECT REPORT & CODE	400 pts

 Total Course Points = 1,000 pts Total

* The total points for each category might change depending on the number of project teams and assignments. The instructor reserves the right to adjust, with sufficient advanced notice, the above point distribution by ± 5 pts. Such adjustments might be based on the difficulty or simplicity of assignments or quizzes or exams.

Grading Information

Grading Percentage Breakdown

(NOTE: Ranges might change if point totals change)

Grading Percentage Breakdown		
Percent of Total Points	Points	Letter Grade
96.66%	\geq 967	A plus
93.33%	\geq 933	A
90.00%	\geq 900	A minus
86.66%	\geq 867	B plus
83.33%	\geq 833	B
80.00%	\geq 800	B minus
76.66%	\geq 767	C plus
73.33%	\geq 733	C
70.00%	\geq 700	C minus
66.66%	\geq 667	D plus
63.33%	\geq 633	D
60.00%	\geq 600	D minus
59.99%	< 600	F

HOW TO CALCULATE/ESTIMATE YOUR GRADE

If students would like to calculate their numeric grade percentage, the formula is as follows:

Numeric CS 224 Grade Percentage =

$$\frac{\textit{Total points from assignments}}{\textit{Total course points}} \times 100\%$$

There is no guarantee that grades will be curved. If so, it will be done at the end of the semester. The instructor is already aware that graduate students need to maintain an overall GPA of B or better. Just because a student NEEDS a particular grade doesn't mean that the instructor will automatically GIVE the student that grade. Students must EARN a passing grade based on submitted and evaluated course work.

Extra Credit Options, If Available

There are no extra credit assignments in this course except for completing designated "Advanced" assignments. However, homework assignments and exams might contain extra credit options.

Late Assignment Submission

Late assignments will receive a 25% point deduction of a graded assignment for each 24hr period after the submission deadline. For example, if an assignment is worth 10 points, and the grade for the assignment is 8/10, and the assignment is submitted one day late, then the point deduction equals 2.5, and the final grade for the assignment is $\text{MAX}(0, 8 - 2.5) = \text{MAX}(0, 5.5) = 5.5$.

Missed Assignments, In-Class Exercises, Quizzes, and Exams**A. QUIZZES:**

- a. The grade for one missed quiz will be replaced with the average of the remaining two quizzes. The average is calculated as the sum of current quiz grades / the number of quizzes for the semester. For example, if quiz 1 = 85, quiz 2 = 95, and quiz 3 is missed, the quiz 3 grade will be replaced by $(85+95)/3 = 60$.
- b. More than one missed quiz will result in a grade of incomplete provided the total missed points for the semester is less than 20% of the total course points.

B. MIDTERM:

- a. The grade for a missed midterm exam will be 75% of the average score for quizzes, programming assignments, and homework assignments provided the total missed points for the semester is less than 20% of the total course points. Or, provide acceptable documentation of the reason for missing the midterm as described in version 1 of this course syllabus and a makeup exam will be provided.

C. HOMEWORK ASSIGNMENTS:

- a. The grade for one missed homework assignment will be replaced with the average of the remaining three homework assignments. The average is calculated as the sum of current homework grades / the number of homework assignments for the semester.

- b. The grade for the second missed homework assignments will be replaced with 75% of the average of the remaining two homework assignments.
- c. More than two missed homework assignments will result in a grade of incomplete provided the total missed points for the semester is less than 20% of the total course points. An alternative is to accept zeros for the missed homework assignments, or if acceptable documentation of the reason for missing the homework assignments is provided, makeup assignments will be provided.

D. PROGRAMMING ASSIGNMENTS:

- a. The grade for one missed programming assignment will be replaced with 50% of the remaining programming assignment.
- b. Two missed programming assignments will result in a grade of incomplete provided the total missed points for the semester is less than 20% of the total course points. An alternative is to accept zeros for all missed programming assignments, or if acceptable documentation is provided, makeup assignments can be provided

E. IN-CLASS EXERCISES:

- a. The grade for one missed In-Class Exercise will be replaced with the average of the remaining three In-Class Exercises. The average is calculated as the sum of current in-class exercise grades / the number of in-class exercises for the semester.
- b. The grade for two missed In-Class Exercises will be replaced with 75% of the average of the remaining two In-Class Exercises.
- c. More than two missed In-Class Exercises will result in a grade of incomplete provided the total missed points for the semester is less than 20% of the total course points. An alternative is to accept zeros for all missed in-class exercises, or if acceptable documentation of the reason for missing the IN-Class Exercises is provided, a makeup assignment can be provided.

F. WEEKLY FEEDBACK:

- a. All missed weekly feedback assignments will receive zero points.

G. FINAL PROJECT REPORT & CODE:

- a. The grade for a missed final project report and code will be 75% of the average of all other course assignments, exams, and quizzes provided the total missed points for all other assignments is less than 5% of the total course points.
- b. If the total missed points for all other assignments is more than 5% but less than 20% of the total course points, a grade of incomplete will be given.

H. TOTAL MISSED POINTS MORE THAN 20% BUT LESS THAN 30% OF TOTAL COURSE POINTS AND TOTAL MISSED POINTS MORE THAN 30%.

- a. **Missed between 20% and 30% of total course points:** A course grade that equal to $(100\% - \text{missed points } \%) * \text{Average of remaining assignments, quizzes, exams, and programming assignments}$.
- b. **Missed more than 30% of total course points:** If the percentage of total missed points is greater than 30%, a course grade that is the result of assigning a zero grade for all missed assignments will be assigned. An alternative grade or options can be discussed with the instructor.

Grade Change Policy:

It is a university policy ([S09-7](#)) that "A change of grade request must be submitted by the department office directly to the Office of the Registrar in a timely fashion. Normally, such requests must be received by the drop deadline of the following semester. Requests for exceptions to this policy must be accompanied with a documented and

University Policies:

Per University Policy S16-9, university-wide policy information relevant to all courses, such as academic integrity, accommodations, etc. will be available on Office of Graduate and Undergraduate Programs' Syllabus Information web page at <http://www.sjsu.edu/gup/syllabusinfo/>. Make sure to review these policies and resources

[S90-5](http://www.sjsu.edu/senate/docs/S90-5.pdf) at <http://www.sjsu.edu/senate/docs/S90-5.pdf> and [University Grading System Policy F18-5](http://www.sjsu.edu/senate/docs/F18-5.pdf) at <http://www.sjsu.edu/senate/docs/F18-5.pdf>. More detailed information on a variety of related topics is available in the [SJSU catalog](https://catalog.sjsu.edu/), at <https://catalog.sjsu.edu/>. In general, it is recommended that students begin by seeking clarification or discussing concerns with their instructor. If such conversation is not possible, or if it does not serve to address the issue, it is recommended that the student contact the Department Chair as a next step.

Dropping and Adding

Students are responsible for understanding the policies and procedures about add/drop, grade forgiveness, etc. Refer to the current semester's [Catalog Policies](https://catalog.sjsu.edu/content.php?catoid=2&navoid=98) section at <https://catalog.sjsu.edu/content.php?catoid=2&navoid=98>. Add/drop deadlines can be found on the current academic year calendars document on the [Academic Calendars webpage](https://www.sjsu.edu/provost/resources/academic-calendars/index.php/) at <https://www.sjsu.edu/provost/resources/academic-calendars/index.php/>. The [Late Drop Policy](https://www.sjsu.edu/aars/forms-resources/late-drops.php) is available at <https://www.sjsu.edu/aars/forms-resources/late-drops.php>. Students should be aware of the current deadlines and penalties for dropping classes. Information about the latest changes and news is available at the [Advising Hub](http://www.sjsu.edu/advising/) at <http://www.sjsu.edu/advising/>.

Consent for Recording of Class and Public Sharing of Instructor Material

[University Policy S12-7](http://www.sjsu.edu/senate/docs/S12-7.pdf) at <http://www.sjsu.edu/senate/docs/S12-7.pdf>, requires students to obtain instructor's permission to record the course and the following items to be included in the syllabus:

- “Common courtesy and professional behavior dictate that you notify someone when you are recording him/her. You must obtain the instructor's permission to make audio or video recordings in this class. Such permission allows the recordings to be used for your private, study purposes only. The recordings are the intellectual property of the instructor; you have not been given any rights to reproduce or distribute the material.”
 - It is suggested that the Syllabus include the instructor's process for granting permission, whether in writing or orally and whether for the whole semester or on a class by class basis.
 - In classes where active participation of students or guests may be on the recording, permission of those students or guests should be obtained as well.
- “Course material developed by the instructor is the intellectual property of the instructor and cannot be shared publicly without his/her approval. You may not publicly share, or upload instructor generated material for this course such as exam questions, lecture notes, or homework solutions without instructor consent.”

Per University Policy S16-9, university-wide policy information relevant to all courses, such as academic integrity, accommodations, etc. will be available on Office of Graduate and Undergraduate Programs' Syllabus Information web page at <http://www.sjsu.edu/gup/syllabusinfo/> and [Syllabus Information web page](https://www.sjsu.edu/curriculum/courses/syllabus-info.php) at <https://www.sjsu.edu/curriculum/courses/syllabus-info.php>. Make sure to review these policies and resources.

Academic Integrity

Your commitment, as a student, to learning is evidenced by your enrollment at San Jose State University. The [University Academic Integrity Policy S07-2](http://www.sjsu.edu/senate/docs/S07-2.pdf) at <http://www.sjsu.edu/senate/docs/S07-2.pdf> requires you to be honest in all your academic course work. Faculty members are required to report all infractions to the office of Student Conduct and Ethical Development. The [Student Conduct and Ethical Development website](http://www.sjsu.edu/studentconduct/) is available at <http://www.sjsu.edu/studentconduct/>.

Campus Policy in Compliance with the American Disabilities Act

If you need course adaptations or accommodations because of a disability, or if you need to make special arrangements in case the building must be evacuated, please make an appointment with me as soon as possible, or see me during office hours. [Presidential Directive 97-03](https://sjsu.edu/president/docs/PD_1997-03.pdf) at https://sjsu.edu/president/docs/PD_1997-03.pdf requires that students with disabilities requesting accommodations must register with the [Accessible Education Center](http://www.sjsu.edu/aec) (AEC) at <http://www.sjsu.edu/aec> to establish a record of their disability.