

A capstone research experience in evolutionary genomics

Christopher H. Chandler^{1*}

Affiliations:

¹Department of Biological Sciences, SUNY Oswego, Oswego, NY 13126

*Correspondence to: Christopher Chandler, christopher.chandler@oswego.edu, SUNY Oswego address....

Type of Manuscript: *CourseSource* Lesson Manuscript

Article Context Page: To make the submission process easier, you may want to fill out the following form, (you will be asked to select answers during the submission process). Choose all applicable options that effectively describe the conditions **IN WHICH THE LESSON WAS TAUGHT**. Modifications to expand the usability of the Lesson will be addressed later in the text.
 **Please delete this page prior to submission.
 **Not all categories will pertain to your article, in those cases, please select ‘N/A’ when submitting on the website.

- Course
 - Biochemistry
 - Cell Biology
 - Developmental Biology
 - Genetics
 - Microbiology
 - Molecular Biology
 - Introductory Biology
- Course Level
 - Introductory
 - Upper Level
 - Graduate
 - High School
 - Other
- Class Type
 - Lecture
 - Lab
 - Seminar
 - Discussion Section
 - On-line
 - Other
- Audience
 - Life Sciences Major
 - Non-Life Science Major
 - Non-Traditional Student
 - 2-year College
 - 4-year College
 - University
 - Other
- Class Size
 - 1 – 50
 - 51 – 100
 - 101+
- Lesson Length
 - Portion of one class period
 - One class period
 - Multiple class periods
 - One term (semester or quarter)
 - One year
 - Other
- Key Scientific Process Skills
 - Reading research papers
 - Reviewing prior research
 - Asking a question
 - Formulating hypotheses
 - Designing/conducting experiments
 - Predicting outcomes
 - Gathering data/making observations
 - Analyzing data
 - Interpreting results/data
 - Displaying/modeling results/data
 - Communicating results
- Pedagogical Approaches
 - Think-Pair-Share
 - Brainstorming
 - Case Study
 - Clicker Question
 - Collaborative Work
 - One Minute Paper
 - Reflective Writing
 - Concept Maps
 - Strip Sequence
 - Computer Model
 - Physical Model
 - Interactive Lecture
 - Pre/Post Questions
 - Other
- Bloom’s Cognitive Level (based on learning objectives & assessments)
 - Foundational: factual knowledge & comprehension
 - Application & Analysis
 - Synthesis/Evaluation/Creation

- Principles of how people learn
 - Motivates student to learn material
 - Focuses student on the material to be learned
 - Develops supportive community of learners
 - Leverages differences among learners
 - Reveals prior knowledge
 - Requires student to do the bulk of the work

- Vision and Change Core Concepts
 - Evolution
 - Structure and Function
 - Information flow, exchange and storage
 - Pathways and transformations of energy and matter
 - Systems

- Vision and Change Core Competencies
 - Ability to apply the process of science
 - Ability to use quantitative reasoning
 - Ability to use modeling and simulation
 - Ability to tap into the interdisciplinary nature of science
 - Ability to communicate and collaborate with other disciplines
 - Ability to understand the relationship between science and society

- Key Words: List 3 – 10 key words that are relevant for the Lesson (e.g. mitosis; meiosis; reproduction; egg; etc.)
 - _____
 - _____
 - _____
 - _____
 - _____
 - _____
 - _____
 - _____
 - _____
 - _____

Scientific Teaching Context Page

Learning Goal(s):

- Students will understand the utility of next-generation DNA sequencing, especially applications for studying evolution and non-model organisms.
- Students will be able to formulate a scientific question and hypothesis relating to evolutionary genomics, design and carry out a test of this hypothesis (either a novel experiment or with existing data), and analyze and interpret data.
- Students will be able to communicate their results with others.
- Students will be able to analyze and evaluate scientific literature.

Learning Objective(s):

- Write a project proposal clearly describing a novel research question and its significance and the methods (utilizing next-gen sequence data) that will be used to address it.
- Write a clear, complete final paper formatted like a journal article that provides the rationale behind the study and the specific question or hypothesis it addresses, the methods used, the results, and the interpretation and significance of those results.
- Write a detailed, constructive, formal peer review of another student's paper.
- Give a clear 12-15 minute presentation on the project that includes an appropriate amount of detail.

Main Text

1. Introduction:

Rough outline:

- Rationale: Next gen sequencing is revolutionizing the way biological research is conducted. Interdisciplinary skills (especially computational and quantitative skills) are increasingly important for success in a variety of biological careers. Understanding of the scientific method and the ability to communicate effectively is also essential for a scientifically literate society.
- Intended audience: This is a 3-credit capstone research experience (BIO 492 at SUNY Oswego). The students are senior biology and zoology majors who have taken genetics, at least one biological statistics course, and a variety of electives in the biological sciences. There are twelve students per section.
- Pre-requisite student knowledge: Students are expected to understand basic principles of inheritance and molecular genetics as well as essential evolutionary concepts (e.g., mutation, genetic drift, and selection).
- Learning time: 14-week course (+1 final meeting during exam week for presentations); the course meets for a single 3-hour block each week. Students are also expected to commit substantial time outside of class to work on their projects independently.

2. Scientific Teaching Themes:

- Active learning: The course will utilize in-class lab exercises giving students an opportunity to learn and apply bioinformatic skills, with instructor and peer support. Group discussions of primary literature on problems in genomics will reinforce key concepts. Students (working in pairs) will develop their own research questions within a single research framework that can be answered using existing datasets. They will continue analyses on their own outside of class time, and a portion of class time will be used similar to a lab meeting, in which students will share their progress and challenges they are facing, and provide suggestions to help other groups. Students will also provide peer reviews of rough drafts of papers and on final presentations.
- Assessment: Smaller assessments will be given throughout the semester to test students' understanding of key concepts (e.g., importance of quality control for NGS data). After students complete these assessments individually or in small groups, their responses will be discussed as a class. There will also be multiple large assignments: a final paper in the form of a scientific manuscript describing the students' research project; a formal peer review of another students' final paper rough draft (written as though the draft was being reviewed for publication in a journal); and a 12-15 minute presentation, similar to a conference talk.
- Inclusive teaching: Given the student makeup at SUNY Oswego and that taking a BIO492 is required, students with a variety of backgrounds are expected to enroll. Within the framework of the course, students will have freedom to develop their own specific research question/hypothesis, facilitating the exchange of ideas and viewpoints.

3. Lesson Plan:

Week	Topic
1	Introduction: present students with the broad research framework to be addressed in the course (bring a few specimens of the study organism to class), and the datasets that are available to them. Students will brainstorm some specific research questions and have a discussion about how those questions might be answered. Afterwards, to expose students to some potential applications of NGS, pairs of students will be assigned journal articles to read in which genomic data were used to answer an

	evolutionary question. Students will prepare short presentations on their assigned papers for the following week.
2	Brief student presentations on applications of NGS data Overview of sequencing technologies and how they work Introduction to UNIX—directories and files Assignment: log in and run some basic commands (e.g., cd, pwd, cp, mv, rm, mkdir)
3	Sequence QC (trimming/filtering) Sequence assembly and measures of assembly quality (e.g., N50) Assignment: perform small sequence assembly and explore how varying k affects assembly quality
4	Transcriptomes and brief introduction to annotation Assignment: perform a local BLAST search of a transcriptome to identify homologs of a specific protein
5	ddRADseq and potential applications—identifying SNPs and sex-specific/sex-linked markers Assignment: 1-paragraph project idea utilizing datasets available to students
6	Share and discuss preliminary project ideas Comparative genomics Assignment: Revise project ideas
7	Discuss revised project ideas Introduction to R (using variables) Assignment: R exercise, begin analyses to address research questions
8	R continued (if statements, loops) Analyses—group "lab meeting" Assignment: R exercise, continue analyses
9	R continued (data frames and plotting) Analyses—group "lab meeting" Assignment: R exercise, continue analyses
10	Analyses—group "lab meeting" (by this point, analyses should be nearly complete) Assignment: tie up loose ends in analyses; draft Introduction
11	Due: Introductions In class: Exchange Introduction sections, provide informal peer feedback, discuss writing & revising strategies Assignment: Draft Methods & Results, including a figure
12	In class: Exchange Methods & Results sections, provide informal peer feedback, discuss writing Assignment: A complete rough draft of final paper
13	In class: Discuss peer review process and look at examples of actual peer reviews Assignment: Peer reviews
14	In class: Return and discuss peer reviews; discuss presentations and begin working on presentations
15	Finals week; due: presentations & revised final paper

In the pilot version of the course at SUNY Oswego, student projects will center on the evolutionary genetics of a local species of terrestrial isopod (*Trachelipus rathkei*). They will have access to at least two datasets: a transcriptome dataset (raw Illumina reads from mRNA extracted from head/antenna, leg, ventral muscle/nerve, and gonads, from one male and one female); and a ddRAD-seq dataset (gDNA from 13 males and 13 females).

Potential student projects could include:

- Identifying homologs of specific genes in the transcript data (e.g., look for homologs of *doublesex*), and comparing predicted protein sequences to those of other crustaceans or more distant arthropods (e.g., which regions are highly conserved or divergent?)
- Demographic inferences based on allele frequencies from ddRAD-seq (e.g., inbreeding, polymorphism in coding vs. non-coding regions)
- Exploring how varying parameter values affects assembly outcomes with these datasets

4. Teaching Discussion:

Potential modifications/add-ons:

- The open "lab meeting" weeks provide a buffer in case earlier activities take longer than expected.
- While budget constraints will likely preclude generating new NGS datasets during the course, there may be some money for smaller follow-up work in the molecular lab. Thus, during weeks 7-11, if students finish NGS analyses early, they may be able to perform some follow-up experiments (e.g., use qRT-PCR to test for sex-specific or sex-biased expression of candidate genes)

(Further detail to be added after course is taught)