

## **Genomics**

### **EBIO 4460/5460**

Professor	Nolan Kane office hours:	<a href="mailto:Nolan.Kane@Colorado.edu">Nolan.Kane@Colorado.edu</a> Tuesdays, 2-3, Ramaley N395
TA	Kyle Keepers	<a href="mailto:Kyle.Keepers@Colorado.edu">Kyle.Keepers@Colorado.edu</a> 11:30-12:30 T, Th., Ramaley N395
LA	Dustin Bailey	<a href="mailto:Dustin.Bailey@colorado.edu">Dustin.Bailey@colorado.edu</a>
TA-ish	Cloe Pogoda	<a href="mailto:Cloe.Pogoda@colorado.edu">Cloe.Pogoda@colorado.edu</a>

All of us are available by appointment / drop in.

### **Course description**

(See also <http://www.colorado.edu/ebio/2015/06/15/innovative-courses-genomics-ebio-44605460>)

This course will provide an in-depth introduction to the use of modern, next-generation genomic sequence data to address ecological and evolutionary questions. The emphasis will be on working with real data including working through whole genome assembly, SNP calling, transcriptome assembly, and annotation. I will teach the use of UNIX commands, and the basics of programming in bash, AWK and Sed, which are essential for processing large datasets in novel ways. The broad skills you will learn will be:

1. The basics of handling genomic sequence data -analyzing quality, cleaning and trimming to eliminate bad sequence.
2. Whole genome SNP and indel calling using reference-guided analysis.
3. "De novo" whole genome, genespace and transcriptome assembly and annotation.
4. Downstream analysis after alignment – assessing sequence abundance, SNP variation, heterozygosity, and other analysis
5. Communicating scientific results and resources using figures and text. Reading and writing scientific papers.

We will start the course by each putting together a few small genomes (organellar genomes) using data we have generated that has not yet been assembled for this purpose. We will then annotate them and submit to Genbank, and by the end of the semester we will be comparing those genomes. The final project is a scientific paper, with the aim of submitting it for publication in December.

I am confident that the skills learned will be widely applicable to many purposes, both academic and applied. Additionally, the published genomic resources and published papers provide a demonstrated track record of your scientific accomplishments.

### **Learning Goals**

By the end of the course you will have developed several skills that will serve you well regardless of your career path. The skills include...

- 1) Understand verbal and graphical representations of genomes and comparisons of genomes
  - 1) Graphically, verbally, or quantitatively represent genomic information.
  - 2) Identify, create, and evaluate alternative hypotheses using genomic information
  - 3) Collaborate with people of varying knowledge and skills to accomplish a common goal
  - 4) Summarize complex datasets efficiently and communicate that information effectively
  - 5) Be able to answer the question: "What is a genome?"
  - 6) Learn the process of assembling your findings into a professional publication that explains your research to the public.
    1. Learn how to do background research (Introduction, Methods, References...)
    2. Generate publication quality figures
  - 7) Understand how evolutionary time leads to differences among related organisms and what comparisons can be made between them
    1. Phylogeny
    2. Synteny
    3. Gene loss/Retention
  - 8) Develop hands-on computational skills relevant to genomics
    1. Command-line/AWK/grep/sed/basic scripting etc.
  - 9) Develop working familiarity with tools to annotate and compare genomes
    1. BLAST/DOGMA/Sequin
  - 10) Understand how genomic data is generated, from plant tissue to ASCII files
  - 11) Assemble, annotate and publish a unique rDNA sequence
    1. Learn the different components of this complex and what they do
  - 12) Assemble, annotate and publish a unique mitochondrial genome
  - 13) Understand the different types of genomic content, what they do, and how to find them (proteins, tRNAs, and rRNAs)

## **Content**

Content will consist of a wide variety of primary sources, assigned each class. Expect to have reading or other homework every class, and make sure to complete it before the next class.

## **Assessments and Evaluation**

The main assessments are from your projects, as well as class discussion and assisting other students. If you fully complete all of your assignments, contribute to class discussions, and work well with us and with other students, you will get an A in the class. The main part of that is to fully assemble a mitochondrial genome, annotate (identify and describe the function of) all of the genes, and submit it to GenBank (NCBI). That is a huge project, but was fully completed by all but one or two of the class last year. More important and more valuable to you than the grade, though, will probably be the permanent boost to your record from having a very impressive genome you can put on your resume / CV. From that point forward, anyone, anywhere in the world with internet,

can see that you are the first author on a small (but important) genome. And, the goal of a scientific publication is a new addition to the course, which will add another public record of your accomplishments in the course.

### **The "lecture" portion of the class**

The "lecture" portion of the class will consist of various activities, including some lecture, group discussion and problem solving, and peer instruction.

### **Our expectations for you**

We will do our best to make our expectations of you very clear. Here are a few of them:

- 1) Be prepared: read the material before class, do the assignments thoughtfully so you feel that you know the material, bring paper and pencil/pen to class, as well as your laptop.
- 2) Be respectful: come to class on time; treat your peers, and us in a professional manner; refrain as much as possible from side conversations
- 3) Stay on task: an important factor that explains variation in learning gains among students is focus on the problem. Distractions compromise learning.
- 4) Be collaborative: work with your peers and try out your critical thinking and communication skills on your peers.
- 5) Be honest: cheating and plagiarism will not be tolerated and if detected may result in an F for the course (see the University's Honor Code policy below). However, collaboration with other students in this class is allowed and encouraged, as long as each person's efforts are clearly explained and acknowledged.
- 6) Be on time and present in class. Key aspects of class projects are done in class.
- 7) Complete homework on time and keep up with reading and other activities. This class moves fast, so get caught up quickly if you fall behind.

### **What you can expect from us**

We will be here for you and will bring our knowledge of genetics and evolution, our enthusiasm for the subject and for learning, and our experience as an educator to the class every day. We are very committed to teaching and would love nothing more than to have all students earn the highest possible scores in this course.

More specifically, we will, as much as possible:

- 1) be fair
- 2) be transparent about our expectations of you and your work
- 3) be communicative and available to you to talk about the subject or anything you want to talk about
- 4) provide you with cognitive challenges that advance your critical thinking and science-as-a-way-of-knowing skills
- 5) be a resource for you about the study of evolution, and aid you in your professional development

### **Accommodation for Disabilities**

If you qualify for accommodations because of a disability, please submit your accommodation letter from Disability Services to your faculty member in a timely manner so that your needs can be addressed. Disability Services determines accommodations based on documented disabilities in the academic environment. Information on requesting accommodations is located on the [Disability Services website](#). Contact Disability Services at 303-492-8671 or [dsinfo@colorado.edu](mailto:dsinfo@colorado.edu) for further assistance. If you have a temporary medical condition or injury, see [Temporary Medical Conditions](#) under the Students tab on the Disability Services website.

### **Classroom Behavior**

Students and faculty each have responsibility for maintaining an appropriate learning environment. Those who fail to adhere to such behavioral standards may be subject to discipline. Professional courtesy and sensitivity are especially important with respect to individuals and topics dealing with race, color, national origin, sex, pregnancy, age, disability, creed, religion, sexual orientation, gender identity, gender expression, veteran status, political affiliation or political philosophy. Class rosters are provided to the instructor with the student's legal name. I will gladly honor your request to address you by an alternate name or gender pronoun. Please advise me of this preference early in the semester so that I may make appropriate changes to my records. For more information, see the policies on [classroom behavior](#) and the [Student Code of Conduct](#).

### **Honor code**

All students enrolled in a University of Colorado Boulder course are responsible for knowing and adhering to the Honor Code. Violations of the policy may include: plagiarism, cheating, fabrication, lying, bribery, threat, unauthorized access to academic materials, clicker fraud, submitting the same or similar work in more than one course without permission from all course instructors involved, and aiding academic dishonesty. All incidents of academic misconduct will be reported to the Honor Code ([honor@colorado.edu](mailto:honor@colorado.edu); 303-492-5550). Students who are found responsible for violating the academic integrity policy will be subject to nonacademic sanctions from the Honor Code as well as academic sanctions from the faculty member. Additional information regarding the Honor Code academic integrity policy can be found at the [Honor Code Office website](#).

### **Sexual Misconduct, Discrimination, Harassment and/or Related Retaliation**

The University of Colorado Boulder (CU Boulder) is committed to fostering a positive and welcoming learning, working, and living environment. CU Boulder will not tolerate acts of sexual misconduct (including sexual assault, exploitation, harassment, dating or domestic violence, and stalking), discrimination, and harassment by members of our community. Individuals who believe they have been subject to misconduct or retaliatory actions for reporting a concern should contact the Office of Institutional Equity and Compliance (OIEC) at 303-492-2127 or [cureport@colorado.edu](mailto:cureport@colorado.edu). Information about the OIEC, university policies, [anonymous reporting](#), and the campus resources can be found on the [OIEC website](#).

Please know that faculty and instructors have a responsibility to inform OIEC when made aware of incidents of sexual misconduct, discrimination, harassment and/or related retaliation, to ensure that individuals impacted receive information about options for reporting and support resources.

### **Religious observances**

Campus policy regarding religious observances requires that faculty make every effort to deal reasonably and fairly with all students who, because of religious obligations, have conflicts with scheduled exams, assignments or required attendance.

See the [campus policy regarding religious observances](#) for full details.

### **Course schedule:**

Week 1: Intro to Genomics, Bioinformatics, and Unix (HW=play Zork, complete Unix tutorial chapters)

Week 2: More Command Line Basics (HW=SNP table)

Week 3: Identification of sequence differences(HW=Scripting Tutorial)

Week 4: Reference-guided genome assembly (bwa, samtools) (HW=\$Variable Script)

Week 5: De Novo Assembly: concepts and intro (Fieldtrip with Erin Tripp?) (HW=paragraph about lichen species)

Week 6: Trimming and cleaning next-gen sequence reads. Introduce rDNA (HW=pre/post fastqc reports—due Thursday, rDNA Chapter)

Week 7: Intro to rDNA/command line blastn. Isolating rDNA contig/annotating in BankIt (HW=submitting rDNA)

Week 8: Finding your mitochondrial contig. Web blast. Error correction and orientation. Background to mitochondrial genomes. (*Mitochondrial Annotation* Chapter 1, HW=vcftable of error corrected, oriented, circularized genome using script from week 4)

Week 9: Intro to annotation, DOGMA and command line tblastx (*Mitochondrial Annotation* Chapter 2, HW=screen shot of tRNA scan, screen shot of tRNAs in DOGMA and turn in tRNA portion of the checklist)

Week 10: Intro to Sequin (*Mitochondrial Annotation* Chapter 3, HW=First OGDraw)

Week 11: Sequin continued (All imported CDS features properly by end of week) (HW=Intermediate OGDraw)

Week 12: More Sequin (gradHW=outline of paper following [Mitochondrial DNA Journal](#) formatting)

Week 13: Even more Sequin (ORFfinder) (gradHW=methods section written (start thinking about comparisons) HW=submit your genome/Final OG Draw (Optional: make a GIF/post to social media for internet points))

Week 14: Synteny (genoPlotR). Phylogenetic Tree. Compiling table of statistics. Help graduate students write the paper. Each group should have a Google Doc of the paper.

Week 15: Editing and Polishing the paper (Citations)

Week 16: Finals Week – submitting the paper (completely formatted, with author lists, figures, citations)

**Figure 1.** *Sorghum timorense* chloroplast, complete genome. Final project of Kyle Keepers, 2014.

<https://www.ncbi.nlm.nih.gov/nucore/KF998272.1>

