Syllabus

Course Description

- This course introduces students to the molecular biological methods by which samples are converted to a state from which DNA sequence information can be produced. When sequence data is produced in a highly parallel fashion across a large fraction of a genome it is the basis of genomics. For historical reasons the sample put on a sequencer is called a library, and the art of genomics lies in library construction. The experimental design and the technical details of library construction will significantly affect the analyses that are appropriate and the conclusions that can be made.
- In the lecture section students will learn to design an appropriate experiment, read and select an appropriate protocol, then carry it out using equipment and reagents correctly in order to take the sample through library preparation and QC, then apply the material to sequencing platforms. The necessary quality control and library validation methods that ensure success in the sequencing step will be explained. Some of the many processing variations will be discussed in terms of the questions they allow scientists to address. The goal is that students will understand published methods and how they relate to proper data analysis and interpretation.
- In the lab section students will carry out a series of experiments that result in sequence data. The unifying concept will be to characterize allelic variants of selected genes from related plant varieties. Students will purify nucleic acid from young leaves of tomato cultivars, and then produce a selected subset of each genome using PCR. Quality control via spectroscopy, gel electrophoresis and quantitative PCR will be performed. Sequencing libraries will be produced and run on the Ion Torrent PGM. We will be using the CLC Genomics Workbench™ software for assessing data quality, assembly and identifying polymorphisms.
- Students are expected to keep laboratory notebooks that allow all aspects of experiments to be reconstructed.

- BINF 6350-001 and BINF 6350L-001 is for Master’s students and BINF 8350-001 and BINF 8350L-001 is for PhD students.

Pre- or Co-Requisites

- The lecture and lab are co-requisites.
- Prerequisite: It is assumed that students have some background in molecular biology and biochemistry, such that the basic properties of cells and their molecular components are understood. It is important that students understand what enzymes are and how they are used in lab protocols, and what DNA is, with some of its chemical properties.
- No lab skills are assumed, but if you do not know how to use micropipetters, microbalances, gel electrophoresis platforms or spectrophotometers you may expect to spend some additional time acquiring skills with those devices and instruments.
Objectives of the Course

Having successfully completed this course, the student will be able to:

- Follow basic safety considerations when working in a genomics lab.
- Prepare working solutions of biological reagents from stocks.
- Extract and purify nucleic acids for genomics experiments.
- Design experiments, including the necessary quality control steps and statistical requirements, to successfully carry out several types of genomics profiling.
- Read and dissect published protocols for making a library from genomic DNA or amplicons.
- Explain the decision-making process in choosing a preparation method for a sequencing library, in terms of the amplicons, the chemistry, the read length and the depth, to answer a particular question.
- Record results of lab processes in a notebook following best practices.
- Using a sequence analysis application, process data obtained from an Ion Torrent PGM library, including the necessary quality control checks.
- Present research results in written form, as Reports in a prescribed format, and be able to defend conclusions in a verbal discussion.
- Address important issues of research ethics (focusing on the ELSI component of the Human Genome Project) that arise from human studies employing genomic approaches.

Instructional Method

The course has two components:

- The lecture will include the following elements as appropriate: presentation of factual material in a standard lecture format, interactive demonstrations of methods to be applied in assignments, and opportunities for student questions, discussion, and presentations by students.
- The lab is carried out in a teaching laboratory in which all of the necessary equipment and materials are present to follow the protocols provided each day. The primary method of teaching is by demonstration and active practice, with students carrying out the operations themselves either at the same time or immediately after a demonstration. Both an instructor and Teaching Assistant will be present to assist the students and to answer questions throughout the process.

Means of Evaluation

Students will be evaluated on their ability to synthesize lecture and lab material effectively.

- Participation in lab - 40%
- Lab Notebook (with multiple interim reviews to give feedback) - 15%
- Method Reviews (2-3) - 15%
- Homework (3), some part of a library preparation protocol - 10%
- Quizzes (8) - 10%
- Experiment Report (1), a final summary of your results in journal format - 10%
• Project: PhD students must propose a sequencing project, with explicit discussion of experimental design, replicates, controls, and any specific analysis modules that will be required. - 10% (so above will be weighted to 90%).

Weekly Class Work

• Papers, blogs, links to demonstrations, etc. about genomics methods and approaches will be assigned weekly. They are used to foster discussions about current limitations of the methods - you won't be able to participate if you don't keep up with the reading.
• It is essential that you come to every lab - the experiments build on each other and falling behind will make it very difficult to understand the demonstrations, or acquire high quality data.
• Students are expected to keep a lab notebook as they work, to write up component parts of the experiment in the notebook at the end of each lab session, and to write a separate condensed report, structured as a publication when the experiment is complete.
• Students are expected to read the lab protocols and explanatory material in advance of the lab, preparing questions so they can be handled expeditiously.
• Students will work in teams of two in the lab, and will have to coordinate schedules to attend to experiments that take longer than the class period.

Method Reviews

• Assignment 1: There are a number of ways to efficiently add primers to sheared DNA fragments. Discuss the advantages and disadvantages of two distinct methods for which there are research reports (a commercial Users Guide is not sufficient, some experiment must have been performed that validates statements about effectiveness)
• Assignment 2: Because ribosomal RNA is such a large fraction of the RNA component of cells it is desirable to remove it. Compare two distinct approaches to such subtraction, and explain how the sequence of the rRNA might affect the method, and whether the amount of reagent is sensitive to the number and/or metabolic activity of the cells
• Assignment 3: Compare methods used to make miRNA-Seq libraries paying specific attention to how they handle contamination from non-miRNA contaminants.

Experiment Report

• An outline of the expected contents and outline of an Experiment Report is given here .
• A sample paper from the journal Biotechniques is given here, as a model for an acceptable technical writing style. link

Homework
Part of the goal of the project is to make sure you can put together a set of methods that are appropriate for a particular purpose with high throughput sequencing as the final assay step, and
that you know how the choices you make in methods can influence the outcome of the experiment. There are 3 homework assignments that have you focus on several aspects of library creation.

- For the GLK2 gene in tomato, create 500bp amplicons that cover the gene from 200bp upstream to 200bp downstream. Include enough overlap to compensate for primer elimination of sequence variation, and explain if you will need bi-directional sequencing.
- For the Mitochondrial genome in Arabidopsis thaliana, create 15,000 bp amplicons that cover the entire genome. Again, discuss any overlap regions, and whether there are regions of high structure or repetitive sequence that will require special handling.
- Provide a protocol we can follow to retrieve the mRNA and miRNA present in the tomato chloroplast and produce one or more libraries suitable for sequencing on the Ion Torrent PGM. If you use a general approach instead of a targeted approach explain how you will verify that you do not get extra-chloroplast contamination.

- Homework write-ups should include a brief Introduction (1-2 paragraphs), Materials and Methods (include the source of sequence data, any PCR design tools you used and any subsequent tools you used to test your products, like BLAST and the nr database at ncbi, and a 1-page Discussion on anticipated problems. For example, long PCR might take optimization - what might that involve? Be sure to include references to papers and books you have consulted as well as Web tools and databases used.

Note: You are encouraged to brainstorm with others, but each student must produce individual reports on method reviews and homework assignments.

Quizzes

- Eight short quizzes will be posted on Moodle, to reinforce concepts covered in lecture that are particularly important. They are open book but you are to complete them yourself, not with anyone else.

Policies that Apply to this Course

- **University Integrity**: All students are required to read and abide by the Code of Student Academic Integrity. Violations of the Code of Student Academic Integrity, including plagiarism, will result in disciplinary action as provided in the Code. Definitions and examples of plagiarism are set forth in the Code. The Code is available from the Dean of Students Office or online at: http://www.legal.uncc.edu/policies/ps-105.html. A set of links to various resources on plagiarism and how to avoid it is available at the UNCC Library website: http://library.uncc.edu/display/?dept=instruction&format=open&page=920.
• **Attendance:** Attendance at lecture and lab is required, although exceptions will be made for reasons such as illness or family emergency. You will be assigned a laboratory partner and are encouraged to maintain good communication with that person so that your samples are processed in step with the class.

• **Grading Policy:** Grades will be assigned using the customary scale
  - A=90-100%
  - B=80-90%
  - C=65-80%
  - U=0-65%

**Additional Policies**

• The use of cell phones, beepers, or other communication devices is disruptive, and is therefore prohibited during lecture.

• However, cell phones equipped with cameras make very good tools for keeping laboratory notebooks interesting and you are encouraged to use them in that way.

• Please note that bags, coats and other personal gear impede safe movement in the laboratory: we will provide secure storage for such belongings during the laboratory session.

• Students are permitted to use computers during the lecture session, but only for note-taking and other class-related work.

• Students are responsible for all assigned material (readings, protocols, methods) and must be prepared to communicate during class, and prepared to begin activities immediately after a brief Q & A session in the lab.

**Textbooks, Learning Resources**

There are currently no textbooks covering this material. Papers, blog entries and Webinars or videos will be assigned on a weekly basis - see the Readings and Links Course pages (hyperlinked in the tabs below). Open Access papers will be posted on the class Web page, others will be available on the course Moodle 2 page. For background on general molecular biology concepts the first two books below are recommended. A book of genomic library preparation research protocols is the third in the list but it is not a textbook.

• "Molecular Cell Biology" by Lodish, Berk, Zipursky, Matsudaira, Baltimore, Darnell (WH Freeman and Co.)
• "Genomics" by Cantor and Smith (John Wiley and Sons, 1999).
• "High Throughput Next Generation Sequencing: Methods and Applications" in the series Methods in Molecular Biology from Springer Protocols, Humana Press 2011 edited by Young Min Kwon and Steven C Ricke.

**Schedule of Topics**

*Note that the labs in particular are subject to change since outcomes can be hard to predict for first-time practitioners.*

Please see the Schedule page for the course