Bio 4342, Research Explorations in Genomics

Spring 2015

M, W 1:30-5:00; F 1:30-2:30 (occasionally 3:30), Writing Intensive group F 2:30-3:30 when scheduled. Meet in LS 311, Danforth Campus; on Friday 1/16 there will be a visit to The Genome Institute, 4444 Forest Park Parkway, fourth floor. Please review our research problem (on the course website at http://www.nslc.wustl.edu/courses/Bio4342/bio4342.html) and read “A Guide to Consed” on the GEP website at http://gep.wustl.edu/curriculum/course_materials_WU/sequence_improvement/tutorials_and_walkthroughs) prior to the first class.

1/12 M  
Course structure; research problem overview (Elgin, 45 min).  
Lecture: The basics on finishing D. elegans hybrid assemblies (30 min) (Shaffer)  
Lab: Intro to computers; UNIX commands to get you started; begin work on Using Consed Graphically (navigation; adding reads) (Shaffer, Trani, Hodges).  
Lab: Pre-course assessment, GEP survey and quiz (45 min) (http://gep.wustl.edu).  
Reading: Please watch Genome Center Virtual Tour and Next Generation Sequencing Video Tour on line prior to class on Wednesday.  
http://gep.wustl.edu/curriculum/course_materials_WU/introduction_to_genomics/

1/14 W  
Lecture: Overview of DNA sequencing (goals); the pipeline; different sequencing strategies (E. Mardis, 45 min plus discussion)  
Lab: Complete Using Consed Graphically; start Finishing a Drosophila Hybrid Assembly (using navigators, assessing quality, making corrections) (Shaffer, Trani, Hodges).  
Hand out HW#1.  
Reading: “Sleeping dogs of the genome,” Gorbunova et al; start Ellison & Bachtrog + commentary.

1/16 F  
Tour of the Genome Institute (be sure to wear long pants, closed shoes) (Cherilynn Shadding, TGI staff)  
NOTE: meet in the elevator lobby on the Fourth Floor, 4444 Forest Park Parkway (West End Metro stop).

1/21 W  
Lecture: The challenges in generating finished sequence; questions on HW #1 (Shaffer).  
Lab: Continue Finishing a Drosophila Hybrid Assembly (corrections, resolving gaps, PCR primer design) (Trani, Hodges)

1/23 F  
Discussion: Ellison & Bachtrog; RR#1 due (Elgin, ~45 min.)  
**WI subgroup:** 1000 word paper due (RR#1); discussion on scientific writing, peer review (~30 min) (Bednarski, SCRE).  
Optional Lab (~1 hr): Work with Consed, HW1. (Danforth Campus, LS311)

1/26 M  
Lab: HW#1 due. Introduce reference materials and finishing report
requirements (including Finishing Checklist), obtain sequence file for own *D. elegans* project, begin analysis (Leung; Trani & Hodges).

1/28 W Lab: Finishing own project with help from TGI finishers available. Discussion of Miniassembly, Consed, and assessment of sequence quality as needed. Review of when to request additional sequencing, designing primers, as needed (Shaffer).

1/30 F Lecture: Eukaryotic genomes/chromatin structure (Elgin); hand out HW #2. **WI subgroup:** Critique of 1000 word paper due. Reading: Review the modENCODE page on chromatin, including the vignette on fly chromatin ([http://modencode.sciencemag.org/chromatin/introduction](http://modencode.sciencemag.org/chromatin/introduction))

2/2 M Lecture: Heterochromatin/euchromatin (Elgin). Lab: Finishing own project (TGI staff). Reading: Start Haynes et al. (RR#2 due 2/13).

2/4 W Lab: Finishing own project (TGI staff); checklist for project submission; final day for consultation **HW #2 due.** Discussion as needed. Short lecture: Bio 4342 alumni present prototypes of oral report.

2/6 F Lecture: Chromatin states (modENCODE); dot chromosome, GEP findings (Elgin) **WI subgroup:** Rewrite of 1000 word paper due.

2/9 M **Lab meeting / reports** (10 min presentation + 5 discussion each) “Finishing my project; problems identified and solved; meeting GEP sequence improvement standards.”


2/13 F Discussion: Haynes et al, Yandim et al (Elgin) **RR#2 due.**

2/16 M Lecture: Gene finding: detecting and interpreting genetic homology **(Buhler)** Lab: Begin CS HW1. **Class begins at 1 pm.**

2/18 W Lecture: Browser-Based Annotation and RNA-Seq Data **(Buhler)** Lab: Begin CS HW2.

2/20 F **Guest lecture,** Ting Wang: How transposable elements shape the genome. **CS HW1 due. WI subgroup:** peer review of finishing papers due.

2/23 M Lecture: Introduction to gene predictors. Gene prediction tutorial (train with chimp Contig95, Genscan). Organize chimp project teams. (Shaffer) **CS HW2 due.**

2/25 W Lab: continue with chimp gene finding lab as group. (Use chimp data sets; work with
partners.)
Brief presentation: Reporting on chimp annotation (Bio 4342 alumni)

2/27 F  Guest lecture, Weixiong Zhang: microRNA
WI rewrites of finishing papers due.

3/2 M  Lab: Annotation of fragments of the chimp genome.
Lab meeting: oral presentations (12 min each) “Genes and pseudogenes in chimp.”

3/4 W  Lecture: Introduction to *ab initio* and evidence-based gene finding (Leung)
Lab: Complete work on chimp annotation.

3/6 F  Chimp analysis paper due (WI exchange papers)
Optional help session
Start reading: F element paper (RR #3)

3/9 – 3/13  Washington University Spring Break

M 3/16  Lecture: A brief introduction to the annotation projects (Elgin)
Lecture: Efficient gene finding in *Drosophila* (Shaffer)
Lab: Annotation of a Drosophila gene (scripted walk-through)

3/18 W  Lecture: Hidden Markov Models (TA)
Hand out CS HW#3 (HMM; read Eddy, S. 2004b).
Lab: begin work with own Drosophila fosmid to identify genes.

3/20 F  Lecture: Detecting genetic variation in humans (E Mardis)
WI group Chimp analysis peer review; critiques due.

3/23 M  Lecture: Review of gene model checker (Leung)
Lab: Gene annotation; check first gene model by end of class.
CS HW#3 due

Lab: continue with annotation, prepare report on first gene.

3/27 F  Reading Reflection #3 due; discussion of F element paper (Elgin).
WI subgroup: Chimp analysis rewrites due

3/30 M  Lab Meeting: oral report on first gene (10’ + discussion)
Lecture: Using multiple sequence alignments Clustal (TA)

4/1 W  Lecture: Searching for transcription start sites (Leung)
Lab: Continuing annotation project; regulatory elements of first gene.

4/3 F  RepeatMasker; other ways to find repeats (Buhler)
Start Reading: Hoskins et al Promoter Architecture paper
Draft report on one gene due (WI exchange papers)
4/6 M  **Guest lecture:** John Edwards: DNA methylation Lecture: design and use of Lab: Continue analysis and annotation of own project; check repeats.

4/8 W  Lecture: Finding repeats in Drosophila (Leung) Lab: Continue analysis and annotation of own project.

4/10 F  **Motif finding (J. Buhler)**  
**WI subgroup:** peer review of annotation draft

4/13 M  Lecture: Dynamic programming (similarity searches) (TA) (CS HW#4; read Eddy, S., 2004a) Lab: Hunting for TSS, motifs; preparation of annotation reports.

4/15 W  Lecture/Lab: Preparation of annotation reports; alumni present prototypes.

4/17 F  Lecture: Targeting heterochromatin formation in Drosophila (Elgin) **WI subgroup:** rewrite of annotation draft.

4/20 M  Discussion of Hoskins et al. **Reading Reflection #4 due.** Lab: Preparation of annotation reports; functions of one gene

4/22 W  Lab: Preparation of annotation report **CS HW#4 due.**

4/24 F  Lab: Final consultations on annotation report

4/27 M am  **Final written and oral annotation reports:** Submit final paper on your project, with a map of genes (including estimates of transcription start sites), repetitious elements, and alignment to *D. melanogaster*, including a Clustal analysis and a discussion of synteny. Complete annotation of all exons, all isoforms. Include results of searches for regulatory elements. As time permits, exploration of one gene on FlyBase, expanding on gene features, regulation, and function. 10’ presentations (9 am – 12 noon in LS 311).

4/28 Tu  Course Assessment (time TBA): Follow-up session on course evaluation (http://evals.wustl.edu; GEP web site; and Bio 4342 surveys/suggestions).

**Reading**

We will read and discuss four papers over the course of the semester, centered on the theme of genome organization and evolution in Drosophila, with an emphasis on the role of repetitious elements. These papers are listed below; for each paper you will turn in a “reading reflection” (~2 pages, double-spaced, typed) that summarizes the big idea and proposes the next experiment. In addition, we have assembled a list of papers that are pertinent to the material we will be discussing, including papers recommended by our guest lecturers. Among these, papers marked “R” are highly recommended background reading. Background material on BLAST and other computer programs can

Required Reading (reading reflections due):


OR (bonus points for doing both!)


Additional References:

Sequencing Technology

Before class starts:

Sequencing a Genome (view on line, including review of chemistry if needed):

“Next generation Sequencing: Genome Center Video Tour” (view all four segments online):
http://gep.wustl.edu/curriculum/course_materials_WU/introduction_to_genomics/nextgen_video_tour


Mardis ER (2010) The $1,000 genome, the $100,000 analysis. Genome Medicine 2:84.


**Chromatin Structure**


**Human Genomics**


The 1000 Genomes Project Consortium (2012) An integrated map of genetic variation from 1,092 human genomes. Nature 491: 56 -


**Finding Genes in Drosophila**


The following papers may be helpful during the second half of the course:


Revised 12/15/2014