Lab Methods in Genomics: What's it *really* like to run a research course with RNAseq?

A. Malcolm Campbell
Davidson College Biology Dept.
James G. Martin Genomics Program

Big Genomic Data Skills Training for Professors
May 26, 2016
Jackson Lab, ME

Student Context

Davidson College

- liberal arts college, NC
- 2,000 undergraduates

Biology Curriculum

- 11 courses required
- 1 chemistry + 10 biology
- most students: O-chem, math, CS
- non-linear courses

My Prior Experiences

Lab Methods in Genomics

- 2006 & 2007 DNA microarrays
- 2008 & 2009 JGI halophile genomes
- 2011 – 2013 NCSU blueberry genome
- 2014 NCSU broccoli genome
- 2016 RNAseq (Burmese python organs)
  - computer lab, 18 students
  - sophomores – seniors
  - no CS experience required

Research Project Structure

- 6 snakes: 3 fasted, 3 fed
- liver and intestine RNAseq
  - (10-20 million reads each sample)
- GCATSeek workshop, June, 2015
  - GCAT/HHMI sever cluster
- Laurie Heyer (math/CS) colleague
  - helped a lot with R

Spring 2016 Learning Goals

Students will:

- experience gene expression analysis of RNAseq data.
- appreciate the amount of subjectivity involved in genomics research.
- understand how genomics is a suite of tools that spans the small/big biology divide.

Spring 2016 Learning Objectives

- Define terms (knowledge)
- Describe a gene (comprehension)
- Report your findings, oral and writing (comprehension)
- Explain transcriptomes (comprehension)
- Demonstrate computer skills (application)
- Examine signaling pathways (analysis)
- Test differential expression (analysis)
- Integrate with published data (synthesis)
- Evaluate data (evaluation)
- Assess objectivity and subjectivity (evaluation)

Weekly Course Structure

- 1 – 3: read papers (graded project outline)
- 4 – 5: learn methods (R, clustering)
- 6 – 7: start DESeq (R coding)
- 8: develop detailed plan
- 9: oral reports (group) + written report (solo)
  + graded peer review
- 10 – 11: continue research
- 12: start writing final report
- 13: oral reports (group) + graded peer review
- 14: rough draft (solo) + graded peer review
- 15: submit final written report (solo)

http://www.bio.davidson.edu/Courses/Bio343/LabMethods_2016.html

Milestones and Grades

- written python summary (solo author; 10% total grade)
- group oral presentation #1 (10% total grade)
- written status report (solo author; 20% total grade)
- group oral presentation #2 (20% total grade)
- peer reviews on presentations (each is 5% total grade)
- term research paper solo author (25% total grade)
- peer review of term paper (review is 5% total grade)

Course Evaluations (pros)

- lit review helpful
- group work good (3/team)
- status reports helpful
- wiki pages helpful (lab notebooks)
- collaboration encouraged
- individualized research paths
- balanced direction with independence
- course structure engaging (self-directed)

Course Evaluations (cons)

- course not structured
- need more outside time (~2 hrs/week?)
- clearer expectations
- rubrics for graded assignments
- R coding was difficult
- frustrated by instructor's lack of R expertise
- little inter-group collaboration
- analysis was rushed

Course Evaluations (quotes)

“For once I felt like I was learning about things that can truly apply after I graduate.”

“Made me change the way I think about science.”

“Challenging because a lot was new to me, but well worth the experience. I felt like I was actually doing something.”

“Was able to use initiative, think and be challenged.”

Changes for Spring 2017

- pipeline will be clearer (to me and them)
- more grades before spring break
- grading rubrics posted
- speed up lit review
- R coding guided tour with mini-datasets
- get to data analysis faster
- informal check-ins with each group (weekly)
- groups meet 2-3 hours outside class
- sustainability of RNAseq projects???

Personal Highlight

During oral presentations, two students **argued** whether a gene was induced or repressed. They discovered two paralogs with the same name, one induced and one repressed. Induced paralog was key insight into organ growth!

Acknowledgements

- Laurie Heyer: Davidson – R coding & data analysis
- Arthur Hunt: UK & GCAT – cDNA library
- Mark Peterson: Viterbo U. & GCAT – R & DESeq
- Christine Walls: Juniata C. & GCAT – HHMI cluster
- Vince Buonaccorsi: Juniata C & GCAT Seek director
- 18 brave, trusting undergraduates