The Genomics Education Partnership
An Introduction to Bioinformatics in a Course-based Undergraduate Research Experience

Sarah C R Elgin
Sara Anderson, Shan Hays, Nighat Kokan, Nathan Mortimer, Chris Shaffer, Wilson Leung, David Lopatto, and the students and faculty of The Genomics Education Partnership
Poster P0381 Monday 3:00 pm – 4:30 pm

GEP Funding:

I hear and I forget.
I see and I remember.
I do and I understand.
- Confucius

PCAST recommendation #2:
“Advocate and provide support for replacing standard laboratory courses with discovery-based research courses.”

Goals: analytical thinking, beginning bioinformatics skills, and building scientific community

• Integration of genomics into the undergraduate biology curriculum
• Integration of research thinking into the academic year curriculum
• Creation of dynamic student-scientist partnerships
• Publication of research in genomics & in science education

Advantages of bioinformatics:
– Low laboratory costs (computers, internet connection)
– Web-based tools
– No lab safety issues; open access 24/7
– Lends itself to peer instruction
– Large pool of publicly accessible raw data
– Currently >100 faculty from >100 colleges & universities; last year 70 schools claimed projects, >1000 students involved.

Goals:

• Most GEP members are faculty at PUIs

Drosophila comparative genomics:
to understand the organization, evolution and gene function on the 4th (dot) chromosome

Phylogenetic tree produced by Thom Kaufman as part of the modENCODE project

Genomics Education Partnership

Most GEP members are faculty at PUs

Partners join by attending one-week workshop.
Shared work organized on GEP website (http://gep.wustl.edu).

Phylogenetic tree produced by Thom Kaufman as part of the modENCODE project

Drosophila Muller F Elements Maintain a Distinct Set of Genomic Proportion Over 40 Million Years of Evolution
Varied responses!

- GSA journals blog: "Undergrads power genomics research" - good description of undergrad role.
- GSA officers, journal editors, CUR, others supportive
- Wash U press release: "Massively parallel biology students"
- BUT Neuro DoJo blog: When does authorship stop meaning anything useful? Tweets: "I'm flabbergasted..." "Basically, if you're not on this author list you have probably never seen a fruit fly." "The author list is over 1,000 people - seems longer than the chromosome they were studying." "But how do you list it on your CV?"
- News item by Nature fairly balanced; several students responded
- Article level metrics: Score 274, top 5% all papers, highest ever for G3! (measuring "the social impact of scholarly literature")
- Strategy: divide and conquer!
- Students improve the sequence & carefully annotate the genes
  - Use publically available genome sequences from the web
  - Divide into 100 or 40 kb projects, which different schools claim & then submit
  - Each project done at least twice independently; should pass "Gene Model Checker;" results reconciled (~75% complete congruence for D. biarmipes).

Recent publication built on student assemblies and gene annotations from four species (30 million years)

Students improved 3.8 Mb, closed 72/86 gaps, added 44,468 bases; annotated 878 genes (1619 isoforms); only 58% agree with GLEAN.

![Diagram of G3 Sequence Improvements and Gene Annotations](image-url)
The class was very intellectually challenging for me. It taught me to think in a way that I had never thought before. It was a great team-building and learning experience. It was independent work, but you were going through the same things as the other students and built camaraderie.

Everything we gained from the class... was either found by desperately messaging around the various websites or by talking with other students.

I know if I could survive this class then I could survive just about anything.

Pool student data allows us to analyze patterns of evolution of different chromosome domains

SURE survey: GEP students show results that are the same or better than summer-in-the-lab students (green) given sufficient time

The class was very intellectually challenging for me. It taught me to think in a way that I had never thought before.

It was a great team-building and learning experience. It was independent work, but you were going through the same things as the other students and built camaraderie.

Everything we gained from the class... was either found by desperately messaging around the various websites or by talking with other students.

I know if I could survive this class then I could survive just about anything.

Causative evidence for the impact of CUREs

- University of Texas – Austin
  Freshman Research Initiative
- Three semesters; research methods course, literature; two semesters in lab, ~9 hr/wk
- Each stream led by Research Educator and PI
- Used propensity score matching (FRI n = 2648; non-FRI ~50,000)
- Effects similar for diverse students

Goal: custom annotation workflow: G-OnRamp

Partnership with NIH BD2K funding

4217 Genome Annotation with Galaxy (4 pm Tuesday, California Room)
**G-OnRamp can facilitate annotation projects**

- Investigator identifies genes, sequences, or regions of interest
- Use gene predictors, RNA-seq data, comparison to other species to create genome browser; plan to make project packages.
- Target: new species (eukaryotes) of interest
- Students can use the Gene Model Checker to check their work when a reference species is available
- Recommend “parallel” projects
- Minimum of two independent determinations
- GEP looking for partners with new projects

![Gene Model Checker](image)

**Conclusions:**

- **Genomics** is an excellent topic for CUREs (course-based undergraduate research experiences)
- **Evidence** points to knowledge gains, gains in understanding of science, improved retention in STEM & improved graduation rates.
- **Massively parallel undergraduates** can accomplish research that could not be done in any other way!
- **G-OnRamp** should make it easy to create good genome browsers for annotating new genomes (contact selgin@wustl.edu for info on beta-users workshop or go to [http://gonramp.wustl.edu/signup](http://gonramp.wustl.edu/signup)).