

# Massively parallel biology students: More than 900 students co-author genomics research paper

May 11 2015

---



The Genomics Education Partnership takes advantage of the practical advantages of bioinformatics to offer a research experience to many more students. Each horizontal line in this computer display represents a different line of evidence that could be used to argue for or against the presence of gene in a DNA sequence. The student, seeing a discrepancy, must drill down deeper to try to find its source. Credit: Mary Butkus/WUSTL

The list of authors for an article on the comparative genomics of a fruit fly chromosome, published online May 11 by the journal *G3*, runs three single-spaced pages. Large author lists are the norm in high-energy physics, but a novelty in biology. What is going on?

The 1,014 authors include 940 undergraduates from 63 institutions, all working in parallel to solve the mysteries embedded in the DNA sequences of the unusual "dot chromosome" in [fruit flies](#).

"By organizing the efforts of 'massively parallel' undergraduates, we can solve problems that would defeat other methods," said Sarah (Sally) Elgin, PhD, the Viktor Hamburger Professor of Arts & Sciences at Washington University in St. Louis and founder of the Genomics Education Partnership, the group coordinated at Washington University that completed the research project.

Equally important, Elgin said, we are able to give many [students](#) the hands-on research experience traditionally available only to the lucky few who could find slots in laboratories over the summer.

## **From genome to genomics**

Elgin studies genomes, or the genetic material of an organism that consists of both coding and non-coding sequences of DNA. In the past 15 years, high-throughput machines that can spit out sequences have revolutionized the field, turning it from lab-based research that looked at one gene at a time to computationally-based research focused on large-scale data collection and analysis, called "genomics."

"It used to be that the bulk of the employees at Washington University's McDonnell Genome Institute were busy at the wet bench cranking out sequence," Elgin said. "Now, the majority are bioinformatics people who are analyzing data while the machines crank."

When Elgin became a Howard Hughes Medical Institute Professor in 2002 and was tasked with making science more engaging to [undergraduates](#), she realized that the shift from the wet bench to the computer meant that the doors of the lab could finally be thrown open.

"You have low lab costs, you don't have any safety issues, and you can make more use of peer instruction. There are huge advantages to bioinformatics as an introduction to research," she said.

To realize this vision, she founded the Genomics Education Partnership, a large collaboration of colleges and universities coordinated by Washington University's biology department and McDonnell Genome Institute, that provides students with the opportunity to work on large-scale DNA sequencing projects.

## **A gnarly chromosome**

The project that resulted in the *G3* paper was to find and compare the genes on a peculiar fruit fly chromosome, called the Muller F element, or dot chromosome.

"We're particularly interested in this funny little chromosome, because it looks like it's all packaged up as heterochromatin, which is usually inactive, but it has 80 genes that are getting expressed, and they seem perfectly happy," Elgin said.

Chromatin, the complex of proteins that packages the long DNA sequences tightly enough that they can fit inside the nuclei of cells, comes in two varieties. Heterochromatin, made up of DNA tightly spooled around the proteins, is typically gene-poor, and those genes it does contain are typically silenced. Loosely spooled euchromatin, in contrast, is gene-rich and its genes are actively expressed.

The dot chromosome seems to be a weird combination of the two types of chromatin. Elgin would like to know how the 80 genes on the dot chromosome remain active despite being trapped in heterochromatin. The fact that they are expressed means we have more to learn about gene regulation, she said.

The students attacked this problem by comparing the dot [chromosomes](#) of four species of fruit fly that last shared a common ancestor about 40 million years ago. The idea was that the differences and commonalities among the four might point to characteristics of the genes and genome that maintain this balance, helping to define the unique characteristics of the dot chromosome genes, and the features of the surrounding chromatin.

## **How to hack a genome**

The students worked with publicly available draft genome sequences for three *Drosophila* species, plus the high-quality sequence of *D. melanogaster*, the fruit fly commonly used in the lab. Their first task was to correct errors in the published sequences and request additional sequencing to cover gaps.

Their next task was to find the genes in the improved sequence. This is not as easy as it might sound. "The genome is like a novel, say *Moby Dick*, but instead of being a single volume, it's 20 volumes because something has nefariously inserted gibberish at random places throughout the whole text," Elgin said. The students' job is to recover the "sentences"—the genes—within this garbled text.

"For the student, the prediction of the gene is the hypothesis. And the task is to construct an argument from the evidence that your hypothetical gene is actually real, an argument strong enough that you can defend it both to your classmates and in the report form you submit back to Genomics Education Partnership."

The students have many different lines of evidence to work with. "But the lines of evidence often contradict one another," Elgin said with a rueful chuckle. "To locate the genes they must reconcile these contradictions, and this not a trivial problem," Elgin said. "It really

challenges the kids.

"Our goal is to teach students to be inquiring and analytical, but they also come away with an understanding that a genome is a Rube Goldberg device," she said. "It's a mess. It works, but you really have to wonder sometimes."

## **Just Google me on Pub Med**

Students who successfully completed the research they were assigned became co-authors on the *G3* publication. "Students have been apprentices for so long, they jump at the chance to do something real," Elgin said.

The digital labs seem to benefit students in much the same way as wet-bench labs. "Our assessments show GEP students acquire a better understanding of how research is done, and their gains are comparable to those reported by students who have spent a summer in the lab," Elgin said.

She is particularly pleased that the students learned to work with large databases. "Biology used to be the refuge of the scientifically interested and the mathematically challenged," she said. "But modern biology wouldn't exist without computer science, and we really need to help our students improve their math and computer skills."

Not only biology students, however; Elgin is thinking even more broadly. "Ultimately we would like all students to have a hands-on research experience," she said. "We think research is a good thing. It trains the students to think critically, to look at data skeptically, to check one line of evidence against another, and to marshal all their evidence before they reach a conclusion."

"These skills increase the students' confidence that they can make a difference in the world," Elgin said, "and the world also needs more people with the ability to think critically. Now, more than ever."

**More information:** Leung, W. et al. (2015). *Drosophila Muller F Elements Maintain a Distinct Set of Genomic Properties Over 40 Million Years of Evolution*. *G3: Genes| Genomes| Genetics* 5(5):719-740  
doi: 10.1534/g3.114.015966 [www.g3journal.org/content/5/5/719.full](http://www.g3journal.org/content/5/5/719.full)

Provided by Washington University in St. Louis

Citation: Massively parallel biology students: More than 900 students co-author genomics research paper (2015, May 11) retrieved 24 April 2024 from <https://phys.org/news/2015-05-massively-parallel-biology-students.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.