

## Titus Brown Trains Scientists in Computational Data Analysis

November 10, 2010

The capabilities of technologies used by scientists in the lab have grown exponentially in recent years. These advances have created a problem for biologists who are now struggling with data overload as they wrestle massive amounts of data from genomic sequencers.

“Sequencing technology has exceeded the prediction termed as Moore’s Law, with the amount of data produced more than doubling every 18 months, thereby creating new headaches for scientists as they work with these massive data sets,” says Titus Brown, assistant professor of computer science and of microbiology. “The problem is like generating the amount of data from a particle accelerator and then using basic office computers to conduct analysis.”

Two bioinformatics issues center around the equipment needed to handle large data sets and training biologists with the techniques needed to work with the data. Added to these problems is that often the data involved comes from multiple sources with each data set formatted differently.

Using College of Natural Science support provided by the George E. Leroi Strategic Visioning Fund endowment, Brown developed an intensive two-week program where he trained two dozen scientists in bioinformatics.

“We found a high demand among scientists who were looking for immediate solutions in their ongoing research,” says Brown. “These were biologists who were analyzing and integrating data sets as they are essentially reverse-engineering what nature has created.”

The program attracted scientists from academia and business. While ten of the participants were from MSU, others traveled from universities as far away as Italy and from companies including Monsanto and Pioneer.

Brown has received international attention for the program as it fills a significant need in training scientists in computational data analysis.

“Giving biologists customized training that is normally for computer scientists and physicists is essential for biological scientists working on the latest problems,” says Brown. “The key is focusing on their skills as biologists and giving customized bioinformatics training as it applies to their knowledge base.”

An efficient and cost-effective method he used was partnering with Amazon Web Services. The short burst of cloud computing allowed them to use large amounts of processing power without investing time and money in computing resources for a two-week program. This has earned Brown much attention as it is a creative, low-cost solution which many institutions often overlook.

The program was held in June at MSU’s rural Kellogg Biological Station. The intense day- and night-long session was both a scientific bonding and learning experience for the participants. Brown is considering building this into an annual event given the success of the first workshop, feedback from the participants, and interest he has received from across the country.

Brown notes that, above all, the program was a forward investment as many of the participants will go on to train others. He has included all materials and components of the course along with the datasets on a website ([ged.msu.edu/angus](http://ged.msu.edu/angus)) with a Creative Commons license. Brown hopes that others can benefit and learn the same techniques for managing the data overload confronting scientists in bioinformatics.

## **Titus Brown Trains Scientists in Computational Data Analysis**

Published on College of Engineering, Michigan State University (<http://www.egr.msu.edu>)

---

– Mike Steger

**Source URL:** <http://www.egr.msu.edu/news/2010/11/10/titus-brown-trains-scientists-computational-data-analysis>