

Explore BiGG Data addresses shortage of women in Bioinformatics, Genetics, and Genomic Sciences using undergraduate research and network mentoring

Kimberly D. Gwinn¹, Melissa Cregger^{1,2}, Joseph Donaldson³, Scott J. Emrich¹, Denita Hadziabdic¹, Jessica Moore², Aaron Onufrak¹, Bonnie H. Ownley¹, Shelli Rampold¹, Margaret E. Staton¹, Carrie A. Stephens¹, and Rebecca T. Trout Fryxell¹

University of Tennessee¹, Oak Ridge National Laboratory², and North Carolina State University³



Quotes were collected by program evaluator in Individual Interviews. Unless otherwise noted, figures are from reeu.tennessee.edu. Application materials for the 2023 program will be available at this site in Fall 2022.

Introduction

Discipline Needs

- ❖ Geneticists and genome scientists in Food, Agricultural, Natural Resources and Human Sciences (FANHS) generate large amounts of data that require bioinformatics expertise for analyses and comprehension.
- ❖ Bioinformaticians require biological data for creation of improved computational models, data pipelines, data interpretation, and enhanced machine learning abilities.

Critical Workforce Issues

- ❖ Demand for bioinformatics capabilities greatly outweighs supply^{1,2}.
- ❖ Low diversity in Bioinformatics, Genetics, and Genomics (BiGG)¹.
- ❖ Low numbers of qualified women in computer science programs with training in FANHS².

"While I was already planning on attending graduate school, my research experience reinforced my belief that this was a good choice for me and gave me more confidence in pursuing an advanced degree leading to a career in research." BiGG Participant 2019

REEU Design

The Explore BiGG Data REEU was designed to holistically train women in bioinformatics with leadership being predominantly women (78%). The 2+1+2 Research Team model increased opportunities for interpersonal connections and allowed for greater breadth of research projects. Since the teams were small, students with diverse academic backgrounds needed to collaborate to accomplish project goals.

Students with training in computer science shared and improved upon their scripting skills while gaining tremendous domain knowledge in biological systems. The biology-based students developed basic coding skills, as well as an understanding of the role of big data science in agriculture. Scholars created names for their teams. The REEU was hosted virtually in 2020 and 2021 due to Covid-19.

"The connections made with all of the people I worked with were so important. I feel like I could talk to any of my research mentors or other scholars about anything. I intend to maintain those friendships as I continue my academic career." BiGG Participant 2020

Changes in Second Year

Optional Week of Hands-on Experience

- ❖ Half of the student chose this option while remaining students attended session on how to read and critique a scientific paper

All Projects Were Microbiome-Based

- ❖ Students needed more hands-on training with datasets
- ❖ Developed metabarcoding data analysis tutorial with published work

Communication

- ❖ Changed from infographic (Year 1) to a Weebly website (Year 2) and back to infographic (Year 3) (Fig. 1)
- ❖ Book Club added but eliminated in Year 3

"The biggest challenges we faced were probably the code and all the errors in trying to run code on four to six different computers in six different states and then trying to troubleshoot that it was a lot, but I feel our mentors handled that really well." BiGG Participant 2021

Technical Skills

During the first week, the scholars attended presentations on the fundamentals of biological systems and received R training. They were then presented with a dataset that had been developed by biology-based faculty mentors.

Population Genetics Projects 2020



Team Shade investigated *Geosmithia morbida*, the causal agent of Thousand Cankers Disease. Population genetics data based on microsatellite loci were used to understand co-evolution of the pathogen and its vector, walnut twig beetle.

Team Super Fly focused on the population genetics of an invasive & established pest of cattle. SNP data generated with genotyping-by-sequencing techniques from a unique set of stable flies collected from around the world was used to assess genetic diversity and genetic population structure.



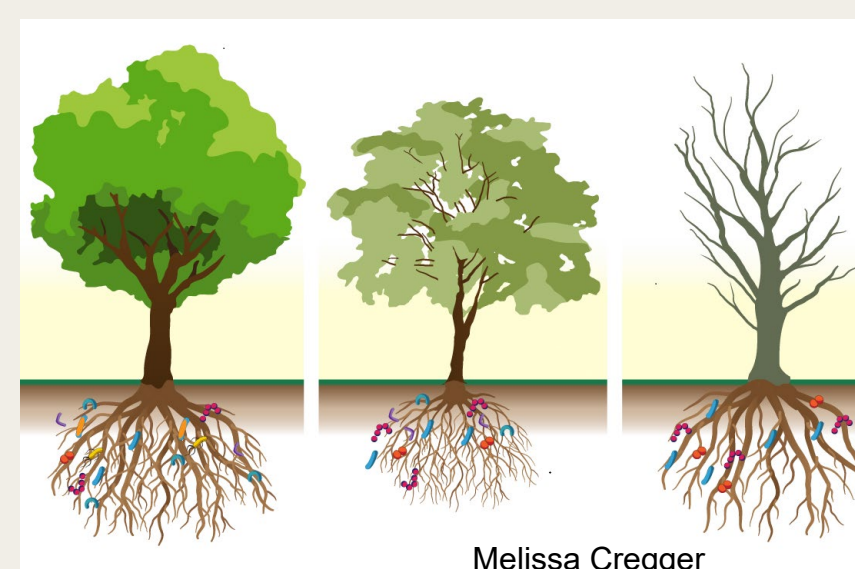
Microbiome-Based Projects 2020 and 2021



Team Cool Ant investigated the impact of myrmecochory on the microbiome of seeds. The question, "Do ants provide seeds with a benefit of reduced plant pathogen loads in the ant nest environment?" was addressed. **2020**

<https://ozarkbill.com/2020/12/07/myrmecochory-seed-dispersing-ants/>

Team Crack-n-Ag investigated soil aggregate inoculum and host factors driving root microbiome assembly and function. Information on unique pockets of microbes in the soil colonizing a root system was investigated. **2020**



Team Plants2 investigated the influence of irrigation and drought tolerance on *Populus* rhizosphere microbiomes. This work was recently published in *Phytobiomes*³. **2021**

Team Rhiz-of-Resistance focused on effects of Companion biofungicide (*Bacillus velezensis* GB03) and plant genotype on *Populus* spp. rhizosphere microbiome. **2021**



Team LunaTicks addressed the questions, "Does the presence of *Rickettsia* within a tick shape the microbial community? If so, how does it differ for each species habitat and sex of the arachnid?" **2021**

21st Century Skills

Development of technical skills is essential in preparation of students for careers and academic programs in BiGG, but development of 21st Century skills (also called professional skills) is critical for student success.

Leadership skill training began with The Clifton StrengthsFinder, an instrument used to identify individual dominant strengths. Programming stressed the importance of these strengths in teamwork. Sessions were designed to increase overall understanding of leadership abilities and how this guides one in future individual and leadership situations. Test results were also used in collaboration with the leadership and research mentors to guide academic/career choices.

"I heard a lot of stories about how there aren't a lot of women leaders. I know that's been true in my personal life. But I really, really loved seeing all the women faculty and women leaders as part of this program. It was very inspiring, and it really helped show me like what I could do. I had trouble picturing myself as a leader before this program, but there were so many good examples of what that could look like. So okay. Now I have a very good one. Awesome." BiGG Participant 2021



Communication training focused on instruction on delivering complex information to the general public. The communications mentor aided students in the development of infographics for their projects (Fig. 1).

Lunch and Learn (LnL). In both years, BiGG Data Scholars indicated that LNL where they learned about the professional journeys of the faculty and graduate student mentors was one of the most beneficial portions in the program. Students became aware of common themes in scientific journeys and that there is no one path to success.

Figure 1. Infographic from Team Crack-n-Ag team.

"This program has been really, very inspiring for me. Women leaders. I've actually never really been involved super much with [understanding] my professors on a personal level. So [in] this program, I got to know again, a lot of very successful women, and get to know more about their life and their history which was my favorite part..." BiGG Participant 2020

Evaluation

- ❖ This study used convergent mixed methods involving three separate assessments:

1. Student Assessment of Learning Gains
2. 21st Century Skills Survey
3. Individual Interviews (~20 minutes)

- ❖ Results were merged, compared, and interpreted using
 1. Quantitative data analysis - IBM Statistics (SPSS)
 2. Qualitative data analysis patterns and themes representative of the BiGG Scholars, consistent with group characteristics sampling were identified⁴

- ❖ Results of the program evaluation were presented at the 2021 Southern Region American Association of Agricultural Educators⁵ and at the 2022 North American Colleges and Teachers of Agriculture⁶.

References

1. Wooley, AW et al. 2010. *Science* 330:686-688.
2. Bonham KS, Stefan MI. 2017. *PLoS Comput Biol* 13(10): e1005134.
3. Kristy, BD et al. 2022. *Phytobiomes*.
4. Patten, M.Q. (2015). *Qualitative research and evaluation methods* (4th ed.). Sage.
5. Donaldson, JL et al. 2021. *Proceedings of the 2021 Southern Region AAE Virtual Conference* (February 8-10).
6. Hunley, L., et al., 2022.. *Proceedings of the 2022 NACTA Conference* (June 20-24, Wooster, OH).