Diversifying the genomic data science research community
The Genomic Data Science Community Network

Main Figure image alt text

Figure 1 = Key stakeholders in the genomic data science community, their action items, and support systems. The four stakeholders are Students, Funders plus Underserved Institution (UI) Admin, UI Faculty, and R1 Institution Faculty. Funders plus UI Admin combine with UI Faculty to create Unique UI Support Systems. Likewise, UI and R1 faculty combine to create collaboration and resource sharing. Students are at the top of a pyramid supported by the other stakeholders and systems. Actions are listed below each stakeholder and system. Students should “Recognize the growing importance of genomic data science”. Unique UI Support Systems “Recognize institutional strengths” and “Remove barriers”. The Collaboration and resource sharing system should “Develop Course-based Undergraduate Research Experience (CURE) curriculum” and “Develop accredited degree programs and curriculum”. Funders and UI Admin need to “Support UI faculty development”, “Incentivize research collaborations”, “Recognize student and faculty research”, and “Broaden access to funding for education and research”. UI Faculty should “Support student professional development” and “Get access to infrastructure”. R1 Faculty must “Incentivize project-based learning and open resources” and “Offer UI faculty development and training”. Funders and UI Admin must also “remove financial and social barriers for UI Faculty participation.” “UI Faculty should leverage strengths specific to their institution type: Community College, Historically Black College or University, Hispanic-Serving Institution, or Tribal College or University.” Finally, “R1 Faculty should help facilitate collaboration, sharing, training, and access to infrastructure”.

Figure 2 = Ten actions that can diversify genomic data science. Two of these action items “Provide access and fund the future”. These are “Provide or gain access to infrastructure and data” and “Fund the future by removing financial barriers to UI faculty and student participation”. Next, three action items “Nurture cross-institutional intellectual environments”. These are “Fund or provide UI faculty development opportunities”, “Facilitate fulfilling collaboration among faculty”, and “Recognize/reward faculty research and teaching”. Finally, five action items “Support student pathways”. These are “Foster student awareness of genomic data science”, “Design modular, open-source learning resources”, “Democratize research experience through CUREs”, “Shape and teach curriculum”, and “Ignite student careers through professional development and research”.

Figure 3 = Five testimonials provided by members of the GDSCN community are each listed beneath a checklist item from Figure 2.

Dr. Sourav Roy at University of Texas at El Paso provides a testimonial under “Shape/teach curriculum”. They write, “I joined the Department of Biological Sciences at St. Mary’s University in San Antonio in the Fall of 2018 with a primary mission to

develop a Bioinformatics program. There was high demand for a Bioinformatics program at the undergraduate level, with very few Universities offering such a program. We developed and launched a B.S. in Bioinformatics with three different tracks in the Spring of 2019 with support from the U.S. Department of Education grant for STEM programs at Hispanic Serving Institutions. I moved from St. Mary's University to the University of Texas at El Paso in the Fall of 2019. However, the program has been running successfully, with 43 students enrolled until the Fall of 2021. Also, the students from the first batch are on track to graduate at the end of the Spring semester, this year."

Dr. Brandi Cron Kamermans at Northwest Indian College and Salish Sea Research Center provides a testimonial under "Design modular, open-source learning resources". They write, "With funds from NIH, I am developing a module for a public, free, online genetics course. The objective of the module is to provide students at tribal colleges a curriculum that emphasizes Indigenous Genetic research in relation to (1) sovereignty, (2) stakeholders, and (3) governance. The goals for the student include how to implement both sovereignty and governance in their genetics research with their communities. For this module, interns from the Salish Sea Research Center and I conducted an interview with Lummi Natural Resources hatchery management about the genetics data that is collected on Salmon. The interview questions were proposed by the student interns and reflect the history and the importance of Salmon to the Lummi, as well as the importance of genetics research for hatchery management."

Andrew Lee at Northern Virginia Community College provides a testimonial under "Facilitate fulfilling collaboration among faculty". They write, "The GDSCN has provided a new way of thinking about applying genomics to early college students who lack a coding background. Prior to involvement in the network, I was flailing in my efforts to efficiently develop genomics activities for biology students. Plugging into the resources of NIH and JHU, while connecting with like-minded community college colleagues, has saved countless hours while developing concrete programs for community college students that can be used with little modification. A surprising side benefit has extended beyond the United States. It has been encouraging to introduce our work to European medical schools who instantly recognize the necessity of applying genomics study to their early medical students. Beyond the practical benefits of connecting with experts from across the nation through GDSCN, it has been a tremendously reaffirming, supportive, and encouraging group for advancing genomics teaching."

Dr. Rosa Alcazar at Clovis Community College provides a testimonial under "Democratize research experience through CUREs". They write, "I work to create pathways for every student, no matter their background, socioeconomics, or clout of their home institution. Working at Clovis Community College allows me to help diverse populations reach academic goals that are otherwise impossible. These students are often first-generation college students and/or from minority groups underrepresented in STEM. My collaborators and I are developing a miniCURE, short modules introducing genomics data science to freshmen with a focus on the scientific process. The goal of the miniCURE is to provide access and exposure to the field and its power to transform

biomedical research. We aim to be fun, work collaboratively, and create a sense of belonging in STEM. The capstone is a poster presentation at a local or regional symposium. The miniCURE has been offered for 8 semesters and 300+ students by two different instructors. Students express motivation to take higher level courses and several students have expressed that “Things didn’t click in the Biology for Majors Course until this project”.”

Dr. Xianfa Xie at Virginia State University provides a testimonial under “Democratize research experience through CUREs”. They write, “Trainings in genomics and bioinformatics are keenly needed at Historically Black Colleges or Universities (HBCUs). While the COVID pandemic has disrupted the normal educational and research activities, it has provided a great opportunity to teach students about genomics, bioinformatics, and evolution. Seizing this unique opportunity, I developed a research project for students who worked with me for their capstone Investigations and Research course to study how the SARS-CoV-2 virus genomes have been evolving and how new variants could emerge due to recombination using publically available genomic sequences of the virus and innovative bioinformatic analysis methods, which has led to a manuscript available in public domain and is still ongoing. In collaboration with colleagues at mathematics and computer science, we have also established an interdisciplinary genomics and bioinformatics program at the Virginia State University, a HBCU.”

Raw figures, including all text therein, can be found here: 10 Simple Rules for GDSCN - Figures