

CCSU
DEPARTMENT OF MATHEMATICAL SCIENCES

COLLOQUIUM

Friday, October 20

3:00 – 4:00 PM

Maria Sanford, Room 101

**COMPUTATIONAL TOOLS
FOR GENOMIC DATA ANALYSIS
AND DATA SCIENCE EDUCATION**

GARRETT M. DANCIK

EASTERN CONNECTICUT STATE UNIVERSITY

Abstract: In this talk, I will discuss several projects in the area of cancer genomics and data science education. My research focuses on the development of bioinformatics tools for genomic data analysis, with an emphasis on making complex analyses accessible to biologists without expertise in programming or statistics. Increased accessibility promises to lead to a gain in understanding of biological processes and genomic diseases such as cancer. The implementation and applications of several tools will be described, with examples related to biomarker identification and personalized medicine in cancer. The Bladder Cancer Biomarker Evaluation Tool (*BC-BET*) is an online database of genomic profiles from >1500 patients, which allows for the rapid evaluation of whether a gene of interest is associated with clinical features such as stage, grade, and survival. Such a tool allows bladder cancer researchers to quickly assess the clinical importance of potential genomic biomarkers. The web-based tool *shinyGEO* allows users to easily analyze data from the Gene Expression Omnibus (GEO), a public repository of gene expression data that includes >1 million samples. This tool allows researchers to quickly evaluate genomic associations in datasets of interest. Finally, I will describe the educational tool *swirl-tbp*, an *R* package for helping students learn *R* programming, data science, and bioinformatics. *Swirl-tbp* is an extension of the *R* package *swirl* that allows instructors to create "template-based problems" that provide students with an endless supply of practice problems that differ, e.g., with respect to numeric values that are randomly generated. Students can therefore practice as much as needed to reinforce concepts and improve their problem-solving skills.

The tools described in this talk are available from the following link:

<https://gdancik.github.io/bioinformatics/>.

For further information:

gotchevi@ccsu.edu 860-832-2839

<http://www.math.ccsu.edu/gotchev/colloquium/>