

Faculty Candidate Seminar

LIFE IN HIGH-DIMENSIONAL SPACE: MODELING SEQUENTIAL CHANGES IN GENE EXPRESSION

Thursday March 30, 2017 • 11:00AM-12:15PM • HPA II Room 345—Video to Lake Nona BBS 103

Differences in gene expression play a crucial role in determining the tremendous variety of cellular properties within the human body. Therefore, understanding the gene expression differences among cells that differ in function, developmental stage, and disease state is a fundamental problem in biomedical research that has implications for diverse fields including regenerative medicine, developmental biology, cancer research, and neuroscience. Advances in sequencing technologies have enabled high-throughput measurement of all genes expressed within individual cells. Such single cell measurements can be conceptualized as points in a high-dimensional space, where each dimension corresponds to a single gene. These data require new models and new computational tools to reveal hidden structure in the high-dimensional space where cells live.

I will present an approach that reveals key properties of a biological process by positing the process as a nonlinear manifold embedded in high-dimensional gene expression space. The approach constructs a manifold from single cell data in an unsupervised manner and uses the geometry of the manifold to order cells according to position in the process and to discover “branches” and “loops” in the process. In addition, I will show how manifold representations of biological processes can be used to integrate multiple kinds of single cell genomic measurements performed on cells undergoing a common process. These methods produce results consistent with biological prior knowledge and yield new insights when applied to experimental data from a number of interesting biological systems.



Dr. Joshua Welch
University of North Carolina at Chapel Hill

Joshua Welch is a Ph.D. candidate in Computer Science at the University of North Carolina at Chapel Hill. Prior to graduate school, he earned dual undergraduate degrees in Computer Science and Piano from Ohio University. His research focuses on combining methods in computer science and statistics to develop new tools, tailored to rapidly developing high-throughput sequencing technologies, for biomedical research. During graduate school, he received the NSF Graduate Fellowship, the NIH F31 Fellowship, and was a funded trainee of the NIH Big Data to Knowledge (BD2K) program.

Hosted by: Faculty Cluster Initiative, Genomics and Bioinformatics Cluster, and Department of Computer Science

