Molecular Biology on Hyperspheres

Abstract: Mass cytometry is a relatively new technology that combines antibody-based protein recognition with mass spectrometry. It is capable of simultaneously measuring the levels of 40-50 proteins in hundreds-of-thousands of single cells from every sample. Most existing analyses of single-cell data (whether proteomics or RNA) have concentrated on first separating/clustering cell types and then assessing changes in the expression of one protein at a time. Using a published data set containing a total of 58 samples either from normal B cell or from patients with different subtypes of acute myeloid leukemia (AML), we have explored how relationships between collections of proteins change depending on the context. This analysis includes changes both between different kinds of samples and between different cell types within a sample. In this talk, I will first describe a new statistic we have developed to compare changes in correlation between pairs of proteins. Most of the talk, however, will describe applications of topological data analysis to discover higher-dimensional non-linear relationships between sets of proteins.

About the Speaker: Dr. Coombes is a Professor in the Department of Biomedical Informatics at The Ohio State University, He received his BS in Mathematics from Lehigh University in 1977 and his Ph.D. in pure mathematics from the University of Chicago in 1982. After working on problems in algebraic K-theory and arithmetic algebraic geometry for several years, he switched his interests to the field of bioinformatics and moved to the University of Texas M.D. Anderson Cancer Center in 1999. Along with his colleague, Keith Baggerly, he spent some time there performing "forensic bioinformatics" analyses to understand why some published analyses were irreproducible (and often wrong). He has become an expert in the analysis of many kinds of high-throughput data, including gene expression microarrays, mass spectrometry, reverse phase protein arrays, second generation sequencing, and more. He has directed bioinformatics and/or biostatistics cores on Cancer Center Support Grants (at M.D. Anderson and now for the James Cancer Center) and for multiple P01 and SPORE grants. His research combines interests in data quality and reproducibility with a focus on extracting robust signatures from omics datasets to predict clinically relevant outcomes. Today's talk will explain how he may have finally achieved a long-term goal of applying some of the tools from his days as a pure mathematician to current problems in bioinformatics.

James S. Blachly, PhD
Friday, January 29th, 11:00am-12:00pm
Carmen Zoom