Recent technological advances have allowed generation of high throughput omics data such as single cell RNAseq, genomics, metabolomics, and proteomics. Omics data are of high dimensional and multi-scale nature, and it hold potential to provide new insights to deepen our understanding of intractable diseases such as intra-tumor heterogeneity and neurological disorders. Complex nature of these data (e.g. data size, high dimensionality, and non-linearity of underlying manifold) imposes both computational and memory challenges that would hinder efficient data analysis. Dimensionality reduction is an interesting computational model for omics data analysis that would allow unsupervised data exploratory and provide simpler data representation to avoid the curse of dimensionality issue. Simply, dimensionality reduction methods seek to project these high dimensional data into a smaller subspace to enable data visualization and capturing interesting patterns that may convey biological interest. This lecture will focus on explaining the mathematical concepts underlying some famous dimensionality reduction methods such as PCA, t-SNE, and HSNE
(scalable version of t-SNE). Applications will be demonstrated on mass spectrometry imaging data.