

Classifying transcriptional and genetic heterogeneity in single-cell measurements

Peter Kharchenko

Single-cell assays are making it possible to examine transcriptional states and other genome-wide properties of thousands individual cells. The ability to directly assess cell heterogeneity is particularly critical in the context of cancer therapy, where presence of phenotypically distinct subclonal populations fuels relapse and resistance to treatment. The transcriptional heterogeneity within such tumors and its impact on disease progression is poorly understood. Furthermore, the extent to which genetic and transcriptional subpopulations correspond to each other cannot be currently assessed. To investigate these questions we have developed methods for analysis of single-cell RNA-seq data in concert with other genomic information. To characterize transcriptional subpopulations we identify annotated or newly-discovered gene sets that are linked to statistically significant heterogeneity within the measured collection of cells. To infer genotype information we rely on probabilistic assessment of single nucleotide variants and copy number variation in individual cells, which can be used to distinguish genetically subclonal populations. We apply these methods to examine transcriptional and genetic heterogeneity in samples of multiple myeloma and other tumors.