Ruli Gao, PhD
Assistant Professor
Center for Bioinformatics and Computational Biology
Department of Cardiovascular Sciences

Clonal Substructure and Copy Number Inference of Human Tumors Using Single Cell transcriptomes

The Gao laboratory has significant contributions in delineating tumor evolution using single cell DNA and RNA sequencing technologies. Dr. Gao and colleagues developed the punctuated copy number evolution model in primary triple-negative breast cancer challenging the Darwinian Gradualism. They proposed a bi-mode chemoresistance evolution model where adaptive selection of genotypes and inductive reprogramming of phenotypes are activated in parallel in triple-negative breast cancer patients. They also developed a nanogrid single nuclear RNA sequencing method that opened the avenue of analyzing single cell transcriptomes from archival frozen tissue samples. Moving forward, her research is focusing on developing novel computational tools to integrate single cell multi-omics and establishing new experimental pipelines to perform third generation of single cell long read and spatial sequencing to unveil both somatic and mosaic genetic abnormalities that underly human cancers, neurodegenerative disorders and heart diseases.

Abstract:
High-throughput single cell transcriptomics analysis is widely used to study human tumors, however a major challenge is distinguishing the stromal cells from the malignant cancer cells, as well as clonal substructure within tumors. To address this challenge, we developed an integrative Bayesian segmentation approach, COPYKAT to estimate genomic copy numbers at 5MB resolution from high-throughput single cell RNA-seq data. We applied COPYKAT to 39,709 single cells from 16 tumors across 4 cancer types, including premalignant and triple-negative breast cancers, pancreatic ductal adenocarcinomas, and anaplastic thyroid cancer. From these data we could accurately (98%) classify tumor cells from stromal cells. In three TNBC tumors COPYKAT resolved multiple clonal subpopulations of genotypes that differed in expression of breast cancer genes and enrichment of cancer hallmarks including EMT and hypoxia. These data show that COPYKAT can accurately resolve clonal copy number substructure in tumors and classify tumor and normal cells in a variety of human cancers.