Modeling and simulation of cancer evolution in single cells

Khanh Dinh[1], Ignacio Vazquéz-García[1,2], Simon Tavaré[1,3]

[1]Irving Institute for Cancer Dynamics, Columbia University
[2]Memorial Sloan Kettering Cancer Center
[3]st3193@columbia.edu

Abstract

Understanding the role of copy number variation in tumor evolution has long been a challenge, in part because identifying subclones from bulk DNA sequencing data is hard. Recent advances in single-cell whole genome sequencing (in our case DLP+), however, enable profiling of copy number aberrations at high resolution in thousands of cells. Single-cell genomics data from these technologies has enabled quantitative measurements of tumor dynamics, and measurements of the rate of chromosomal aneuploidy, whole-genome duplications and replication errors in tumors.

To better understand clonal evolution, we have developed a detailed model for studying single-cell dynamics in a population of cells, incorporating somatic copy number changes, clonal selection of driver mutations and accumulation of neutral passenger mutations. Simulation of the model follows prescribed population dynamics to generate clonal evolution forward in time; clones are defined by their copy number and driver mutation profiles. The phylogeny of a sample is then constructed backward in time. The algorithm is designed to be efficient for large cell populations while maintaining statistical accuracy.

In this talk I will explain the setting of our work and will illustrate some applications of the simulator such as calibrating copy number inference algorithms, generating copy number profiles, and fitting the model to a number of PCAWG datasets. Time permitting, I will also describe a statistical problem arising in the study of scDNA sequencing of a control experiment designed to assess the performance of DLP+.

References

Laks E, McPherson A, Zahn H, Lai D, Steif A, Brimhall J, ... & Shah SP (2019) Clonal decomposition and DNA replication states defined by scaled single-cell genome sequencing. *Cell,* 179, 1207–1221.

Dinh KN, Jaksik R, Kimmel M, Lambert A & Tavaré, S (2020) Statistical inference for the evolutionary history of cancer genomes. *Statistical Science*, 35, 129–144.