Computational methods for characterizing large-scale human genome variations with applications to cancer

**Abstract:** We may be at the cusp of major leaps in personal genome and cancer genome interpretation over the next several years, with profound impact in human health. Large-scale projects based on Next Generation Sequencing technologies have been sequencing thousands of individual genomes and the technologies enabled the sequencing of many cancer genomes and multiple tumor samples from cancer patients. In this talk, I will present some of my major contributions during my Ph.D. and postdoctoral training for characterizing human genomes. In particular, I will present algorithms for detecting various types of large-scale structural variations in sequenced genomes, including NovelSeq, a state-of-the-art method for detecting novel sequence insertions, Next-generation Variation Hunter, a computational method for identifying Mobile Element Insertions (MEI), and CommonLAW, a combinatorial framework for detecting Structural Variations in multiple genomes.

I will also discuss applications of those methods for analyzing cancer genomes and will present computational problems related to the evolution of somatic mutations and intra-tumor heterogeneity. I will present novel algorithms for reconstructing tumor lineage trees and detecting somatic structural variations and will discuss several future directions. These studies are considered first necessary steps for understanding tumor genomics landscape and the challenges in biomarker development.