"Structural variation discovery and genotyping from whole genome sequencing: methodology and applications"

A comprehensive understanding about how genetic variants and mutations contribute to phenotypic changes and variations entails experimental technologies and analytical methodologies that are able to detect genetic variants/mutations from various biological samples in a timely and accurate manner. High-throughput sequencing technology represents the latest achievement in a series of efforts to facilitate genetic variants discovery and genotyping and promises to transform the way we tackle healthcare and biomedical problems. The tremendous amount of data generated by this new technology, however, needs to be processed and analyzed in an accurate and efficient way in order to fully harness its potential. Structural variation (SV) encompasses a wide range of genetic variations with different sizes and formed by diverse mechanisms. Due to the technical difficulties associated with SV discovery, their characterization lags behind that of SNPs and indels. In my research I developed two novel computational methods: one for detecting transposable element (TE) transpositions and the other for detecting SVs in general using a local de novo assembly approach. Both methods are able to pinpoint breakpoint junctions at single-nucleotide resolution and estimate variant allele frequencies in the sample. I also applied those methods to study the impact of TE transpositions on the genomic stability, the inheritance patterns of TE insertions in the population and the molecular mechanisms and potential functional consequences of somatic SVs in cancer genomes.

Mentor(s)
Zhiping Weng, PhD

Dissertation Exam Committee
Jeffery Bailey, PhD (Chair)
Daniel Caffrey, PhD
William Theurkauf, PhD
Manuel Garber, PhD
Ekta Khurana, PhD