Towards Accurate Copy Number Calling in High Ploidy Tumors

Frank, Jessica

The goal of this research was to improve cancer diagnostics, specifically copy number detection methods. Accurately determining the abundance of oncogenes and tumor suppressors is necessary to understand the pathogenesis and appropriate treatment of cancer patients. However, high ploidy, low purity, tumor heterogeneity, chromosomal instability and aneuploidy can confound even the most robust per-gene copy number detection methods. The purpose of this research was to code programs to analyze and compare cancer patient data from two copy number detection methods. One was The Cancer Genome Atlas (TCGA) method and the other was FACETS-based (Fraction and Allele Specific Copy Number Estimate from Tumor Sequencing). The hypothesis was that FACETS would be more accurate because it provides an allele-specific integer copy number that accounts for purity and ploidy. Sources of discordance were identified between the algorithms, including differences in segmentation and discretization methods. Subsequently, metrics of accuracy were developed utilizing RNA data with validation by the alternative algorithm. This accuracy test suggested that TCGA undercalls 44.55% of amplifications while FACETS undercalls 45.05% of deletions, indicating the algorithms are not as accurate as expected. To improve copy number detection methods, a new per-gene copy number was derived by using the FACETS total copy number and average sample ploidy. The RNA-based accuracy test suggested that the new copy number improves sensitivity at a cost to specificity. This research lays the foundation for improved copy number detection in complicated tumors to advance diagnostics and patient treatment and progress towards a cure.

Awards Won:
Fourth Award of $500
Air Force Research Laboratory on behalf of the United States Air Force: First Award of $750 in each Intel ISEF Category