Reconstructing the breast cancer genome

Marley Alford[1], Maria Nattestad[2], Michael Schatz[2]
[1]Undergraduate Research Program
[2]Cold Spring Harbor Laboratory

Structural variations in the genome account for some of the most deadly features associated with aggressive cancers such as the HER2 amplification breast cancers. Twenty percent of breast cancers are HER2 amplified. Many studies on HER2 amplification use SK-BR-3, a model cell line from a patient with this often-fatal strain. Though previous analysis has revealed multiple amplifications and deletions, the Schatz lab has fully sequenced the SK-BR-3 genome, and, using the new PacBio long-read technique, intends to map all structural variations and retrace the steps that prompt their creation. This analysis will involve several processes: first the DNA is sequenced; then the structural break-points are ascertained using "split-read" alignments ; from that information, structural variations are identified and analyzed; and finally, genome threading retraces the path each fragment of DNA took in order to create those structural variations, recreating the original SK-BR-3 chromosomes. A full mapping of the SK-BR-3 genome and the processes of structural variation within it will serve as an invaluable resource for breast cancer researchers, allowing them to better understand the overall transformations in cancer genomes and, more specifically, how these changes cause HER2 amplification. Ultimately the information gained from this project will be applied to future research on breast cancer tumors to help defeat this devastating disease.