Activity-by-Contact Model to Predict Enhancer-Gene Connections: A Tool to Increase Our Understanding of Cancer

Petersen, Lillian (School: Los Alamos High School)
Corliss, Maximilian (School: Los Alamos High School)

Gene expression is regulated by proteins known as transcription factors, which bind to specific DNA sequences called enhancers. Enhancers activate nearby genes, but there is still a limited understanding of which genes they regulate. We created an Activity-by-Contact (ABC) model to predict enhancer-gene connections based on the three-dimensional structure of the genome. Predicting enhancer-gene connections is important because it can identify mutant transcription factors causing the up-regulation of oncogenes in cancer patients. First, we conducted a validation in the K562 cell line and found that the ABC model predicted enhancer-gene connections significantly better than the previous method of using linear distance. Next, the model was applied to study 24 B-Cell Leukemia patients. The samples were first grouped into subtypes by comparing principal component analysis of their gene expression data to 2,000 previously identified samples. Differential enhancers, differential genes, and those with high ABC scores to each other were identified within each subtype. In these cases, the differential enhancer likely regulates the differential oncogene. We were able to identify specific enhancers that regulate known leukemia oncogenes such as FOXO4 and HUWE1. This can allow for the development of novel drugs to target these mutant transcription factors and thereby treat the cancer. This model builds a better understanding of the mechanisms of gene regulation and supports the theory that genes are regulated by enhancer activity and enhancer-promoter contact frequency. The ABC model has the ability to illuminate pathways of oncogene activation, identify mutant transcription factors, and lead to the development of new drugs for targeted treatment of cancer.

Awards Won: