Metaheuristic Analysis of Transcriptional Control to Capture Leukemia Cell Line Characteristics

Popescu, Nicholas (School: The Mississippi School for Mathematics and Science)

Cancer is an extremely variant and personalized disease that is characterized by the disruption of regular cell functions. These disruptions can be analyzed by examining regulatory network changes between normal and cancer cell lines. There are two main characteristics that identify a cancer cell: the cancer cell's dysregulation from an otherwise highly structured hierarchical regulatory network and the increase in cancer promoting transcription factor (TF) bindings with a loss of tumor suppressing TF bindings. Using TF binding data from ENCODEc, a cancer-focused database derived from the ENCODE project, an initial regulatory network rewiring was created to display TF binding changes between GM12878, a control healthy blood cell line, and K562, a Chronic Myeloid Leukemia (CML) cell line. To further analyze this network, a genetic algorithm (GA) method was designed and implemented to model the 'evolution' from a healthy to a diseased cell. This method selects the most significant network rewirings in CML as it searches for the most disruptive TF binding changes through thousands of different network evolutions. A fitness function that measures network centralization penalized for the amount of network changes over each evolution. This method models a healthy cell line's evolution towards its cancerous counterpart through TF rewiring crossover and mutation, resulting in a dynamic view of regulatory network changes in CML. Finally, the GA search identifies key regulatory network targets of the driving oncogenes in CML, MYC, JUND, BHLHE40 and STAT5A.