Systems biology approach to obtain significant modules of immune therapy and colorectal cancer

Swapna Vidhur Daulatabad, Meeta Pradhan PhD., Mathew J Palakal PhD.*

School of Informatics and Computing

Indiana University-Purdue University Indianapolis

Colorectal cancer (CRC) is the second leading cause of cancer death in the United States. There has been a lot of research around genes influencing CRC, despite its extensive understanding on the genetic perspective and the emergence of drugs targeting these genes, the tumor progression could be hardly mitigated. However, immune therapy has recently been observed to be effective in CRC treatment and diagnosis. This study focuses on developing a statistically validated multi-feature analytical approach to identify immuno-oncology targets. The features considered in this study were gene expression, DNA methylation, concepts from literature and immuno-cancer pathways. The network algorithm will identify the potentially relevant immuno-oncology modules of CRC. For the study level-3 data (7.2 gigabytes) of gene expression and DNA methylation was obtained from The Cancer Genome Atlas. Around 13000 genes were identified to be significant from the gene expression data analysis and 19000 genes significant in DNA methylation data. The CRC and Immuno-oncology concepts were manually annotated from 50 peer reviewed articles. The output of the preliminary analysis could predict 95 concepts annotated to the 1587 significant genes and were integrated into the network. The top rank concepts in terms of genes associated were '*apoptosis*', '*transforming growth factor*', '*protein arginine methyltransferase*', '*carcinoembryonic antigen*' and '*methyl binding protein*'. The gene annotated with highest number of concepts was '*PRMT5*', '*CSF2*', 'CFLAR' and '*MLH1*'. These genes were observed in the literature as targets of CRC.