

Computational Approach for the Identification of Plausible Biomarkers from Composite Networks and Gene Expression Data Associated with Colorectal Cancer

Manika Sehgal¹, Tiratha Raj Singh^{2,*}

^{1,2} Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology (JUIT),
Waknaghat, Solan, H.P., 173234, India

¹manika.sehgal22@gmail.com, ²tiratharaj@gmail.com

*corresponding author

ABSTRACT

Colorectal cancer (CRC), characterized by an abnormal cell growth at the lining of colon or rectum is one of the leading worldwide concerns. Recently, studies suggest that the disease may be prevented by regular screening and detection of non-cancerous polyps that might possibly turn cancerous. On the whole, early diagnosis for CRC is essential and there is a need to identify biomarkers for the recognition of the disease at early stages. Thus, an integrative *in silico* approach has been applied that proposes a few change in the expression levels for certain genes and their role in the affected pathways implicated in CRC. The DNA microarray analyses revealed the transcription regulation of several genes and also provided clues for certain up and down regulated genes involved in early progression of CRC. These genes were also subjected to manual curation and annotation for their classification in the biological processes, pathways and molecular functions they perform. For examining the intricate biological processes drawn in CRC, a pathway level analysis was also performed for the identification of vital pathway components or network motifs which are functionally significant. In this study, some novel arrangements of network motifs like BiFan, SIM (single input module), MIM (multiple input module) and other variants were also identified. This comprehensive computational analysis for the genes and their association in the biological processes implicated in CRC may assist researchers and provide useful insight for understanding the mechanism underlying the disease. It is believed that the generated information will be of utmost use to biomedical scientists and biotechnologists.

Keywords: Systems biology, gene expression, subgraph, network motif, genetic instability