Bioinformatics analysis of differentially expressed genes in liver cancer for identification of key genes and pathways

Aisya Fathiya Che Rosli, Siti Razila Abdul Razak and Nurulisa Zulkifle*

Oncological & Radiological Science Cluster, Advanced Medical & Dental Institute, Universiti Sains Malaysia

Abstract

Liver cancer is among the main leading cause of mortality in Malaysia and the world. To identify the key genes and pathways in liver cancer, mRNA microarray dataset GSE84402, GSE64041, GSE60502, GSE29271 and GSE19665 were downloaded from the Gene Expression Omnibus database. Differentially expressed genes (DEGs) from all datasets were obtained by GEO2R web tool and assembled manually with the cut-off criteria set at adj. \( P < 0.01 \) and \(|\log FC| > 1\). Functional and pathway enrichment analysis were performed for the DEGs using DAVID database. STRING database was utilised to construct a protein-protein interaction network. Cytoscape 3.6 software and associated plug-ins were used to visualise and analyse protein-protein interaction (PPI) network. A total of 681, 1564, 1040 and 2265 DEGs were identified from GSE84402, GSE64041, GSE60502, GSE29271 and GSE19665 datasets, respectively. 184 DEGs were screened out in at least four datasets consisting of 70 up-regulated genes and 114 down-regulated genes. These genes were mainly enriched in the terms related to mitotic cell cycle and regulation of nuclear division. Putative PPI network was established with confidence score 0.7 comprising 184 nodes and 1021 edges. Using MCODE plug-in, four modules were detected from the PPI network with module 1 being the most significant. The enriched functions of module 1 included sister chromatid cohesion and mitotic spindle assembly checkpoint. In conclusion, these results identified the key genes and pathways, which could improve understanding of the molecular mechanisms and provide potential targets for liver cancer diagnosis and treatment.

Keywords: bioinformatics; liver cancer; differentially expressed genes; functional enrichment analysis; protein–protein interaction

*Author for Correspondence