Analysis and Mining Onco-targets Breast through Ontology

Edgar Lacerda de Aguiar[1], Lissur Azevedo Orsine[2], Carlos Alberto Xavier Gonçalves[2],

José Miguel Ortega[2], Marcos Augusto dos Santos[1]

1 Bioinformatics and Systems Laboratory, Federal University of Minas Gerais, Brazil , 2 Biodata Lab,

Federal University of Minas Gerais, Brazil

Studies indicate that by the end of 2016 more than 20 million patients will develop some type of cancer. Among the various types of cancer, breast cancer is the most impact among women and high mortality rate. Breast cancer has high biological heterogeneity, which implies a high diversity of molecular forms which are associated with distinct subtypes and distinct drug targets. This high range of variations in the biological entities involved in disease pathology impacts directly on diagnosis and treatment. Because of these facts this work aims to mine the possible genes related to breast cancer, from different databases (DBs), Cancer Genome Atlas (CGA), COSMIC and to relate the genes to database Gene Ontology (GO) with its molecular functions, biological processes, cellular components, thus inferring the main ontologies associated with breast cancer. Initially there was an interpolation of genes between BDs CGA and COSMIC after curated genes were mined and crossed with the top mutated genes of breast cancer. The genes were cured with the BDs UniProt and NCBI. With the UniProt valors Id of cured Genes there was a search for Ontologies in the GO database. The initial results were few satisfactory due to high specificity and high granularity of ontological terms. Better treatment of the data and a new methodological approach to the Ontological terms was necessary. The ontological terms of GoSlim, which are terms and healed with a median specificity were used. Through the analysis it was observed that many biological processes encountered are certainly associated with cancer, such as cell motility, cell adhesion, cell death, response to stress, immune system process and biosynthetic process. The molecular functions found are mainly engaged in activity in the DNA (DNA binding, nucleic acid binding transcription factor activity, transcription factor binding) and protein activity (enzyme binding, kinase activity, phosphatase activity, atpase activity, enzyme regulator activity). The analyzes show a significant correlation between the biological processes and molecular functions encountered: for example, signal transduction (biological process), and signal transducer activity (molecular function). This work proved to be essential to deepening of breast Oncogenes, as well as a better understanding of biological processes, molecular functions and their ontologies. These results will come to confirm the selection of the most critical processes for future studies of Oncogenes and Onco-targets.