

## Gene Prioritization for Breast Cancer

L. jafari<sup>a</sup>, A. Shiralipour<sup>b</sup>, J. zahiri<sup>\*c</sup>, V. Jajarmi<sup>b</sup>

<sup>a</sup> Department of Computer Science and Information Technology, Institute for Advanced Studies in Basic Sciences, Prof. Yousef Sobouti Blvd, Zanjan, 4513766731, Iran

<sup>b</sup> Department of Biotechnology, School of Advanced Technologies in Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

<sup>c</sup> Bioinformatics and Computational Omics Lab (BioCOOL), Department of Biophysics, Faculty of Biological Sciences, Tarbiat Modares University, Tehran, Iran

\*zahiri@modares.ac.ir

**Abstract:** Discovering the most promising genes among the large lists of candidate genes is important. It is scarce that a gene is responsible for a function and an assembly of genes constitutes a molecular pathway or a functional capacity. Therefore, complex diseases are involved multiple genes and are characterized as being multifactorial, so this poses a challenge respecting the search for gene related to them [1]. The candidate genes prioritization prior to experimental testing reduces the associated costs significantly. More reliable techniques that calculate new types of data are essential for the development of recently diagnoses and treatments for many diseases [2]. Prioritization methods for genes are used to test candidate genes and select high-priority genes for more accurate biological tests. There is a demand to explore of new or unknown function of genes for discovering new target genes and pathways in order to diagnosis and treatment of diseases, especially cancers. One of the means to reach to this end is using prioritization tools as a computational approach. We have taken advantage of GPSY, Endeavour and ToppGene among the gene prioritization tools and three ranked lists have been achieved for breast cancer. By analyzing of the obtained results and using R programming language and doing research in previous literature and studies, the important genes were selected that have been studied rarely before. Among these genes, the PRF1 gene is an effective gene for the progression of disorder and irregularities in normal functioning. This gene encodes a protein that it can produce transmembrane tubules and is proficient of lysing non-specifically a multiplicity of many target cells. The PRF1 gene provides instructions for making a protein called perforin. This protein is discovered in immune cells (lymphocytes) called T cells and natural killer (NK) cells, which demolish another cells. Perforin is involved in the process of cell destruction (cytolysis) and the regulation of the immune system. It is a major component of structures called cytolytic granules within T cells and NK cells [3]. Considering the relationship between inflammatory processes and cancer in the beginning of cancer, it seems reasonable to select PRF1 gene for more studies and investigation.

**Keywords:** Microarray Data; Gene Prioritization; Breast Cancer; R programming language; Gene Expression;

### References

- [1] Žitnik, Marinka, et al. "Gene prioritization by compressive data fusion and chaining." *PLoS computational biology* 11.10 (2015): e1004552.
- [2] Simões, Sérgio N., et al. "NERI: network-medicine based integrative approach for disease gene prioritization by relative importance." *BMC bioinformatics* 16.19 (2015): S9.
- [3] Modarresi, Farzaneh, et al. "A case report of novel mutation in PRF1 gene, which causes familial autosomal recessive hemophagocytic lymphohistiocytosis." *BMC medical genetics* 18.1 (2017): 49.