Postgraduate student: **Dimitra Karagouni**

Thesis Title:

Computational and high performance data analysis, aiming at finding and understanding the molecular genetic changes associated with breast cancer.

Abstract:

This thesis is based on tumor data carried out by the Cancer Genome Atlas network (TCGA: Comprehensive molecular portraits of human breast tumours, Nature 2012). These data focused on genome analysis by high performance techniques with aim to better understand the genetic basis of cancer. Our study focuses on the management and analysis of such high throughput data in the main types of breast cancer.

The first part of this thesis refers to the creation of an automated program, which can be used in expression data for finding significant genes which offer the ideal separation of different types of any disease. Specifically, methods for finding gene-transcripts with greater variability of expression in all samples were used and lead to better categorization. Based on this classification the most diverse in their expression regulators were found, and they were characterized as the most representative per group. In the present study we came in groups of genes that exhibit the most variable expression in the major types of breast cancer.

The second part of the thesis deals with the analysis of genes mutations, one or more nucleotides in binding regions of small non-coding RNAs, miRNAs, which play a key role in regulating the expression of most genes. From the above analysis miRNAs, which bind to mutated messenger RNA genes, were found. The mutations referred to MRE regions in quite a large number of samples. SNPs can lead to non-specific binding of miRNAs or even the binding of other small non-coding RNAs in these regions, may contribute to the modulation of expression of these genes to specific types of breast cancer compared with normal samples.

The study of this entire regulatory genes network, mutations and miRNAs can be associated with changes in molecular pathways involved with cancer, and any other under investigation pathological condition.

SUBJECT AREA: Computational Biology

KEYWORDS: breast cancer, high- throughput data, significant genes, miRNA, target prediction, SNPs

Examining Committee

Phd. Artemis Hatzigeorgiou, Professor, Department of Computer & Communication Engineering, University of Thessaly (thesis advisor)

Phd. George Spyrou, Staff Research Scientist (Professor Level), BRFAA

Phd. Evangelia D. Chrysina, Scientific Personnel, Institute of Biology, Medicinal Chemistry and Biotechnology, National Hellenic Research Foundation