

TRANSTAGING: Transcriptogram-based staging of cancer

Jose Luiz Rybarczyk-Filho, Marcio Luis Acencio, and Ney Lemke

Departamento de Física e Biofísica, Instituto de Biociências de Botucatu, Unesp - Univ Estadual Paulista

The classification of different tumor types is most important in cancer diagnosis. The cancer classification studies are clinical based and have restricted diagnostic ability. Cancer classification using gene expression data is known to contain the keys for addressing the central problems relating to cancer diagnosis and drug discovery. Analysis of genome-wide expression data poses a challenge to extract relevant evidence. We use computational method that order genes on a line and clusters genes by the probability that their products interact. Protein-protein association information can be obtained from large data bases as STRING. The genome organization obtained this way is independent from specific experiments, and defines functional modules that are associated with gene ontology terms. The starting point is a gene list and a matrix specifying interactions. Considering the Homo sapiens genome, we projected on the ordering gene expression, producing plots of transcription levels for three different tumor types (lung, neuroblastome, breast), whose data are available at Gene Expression Omnibus database. This analysis differentiated normal and tumor tissues. Moreover, the subdivision of the tumor tissues in many classes that were previously inspected with biological process ontologies (Gene Ontology) shown that each class has a set of modified process. This result is the first evidence to find biomarkers for tumor staging by a computational method.