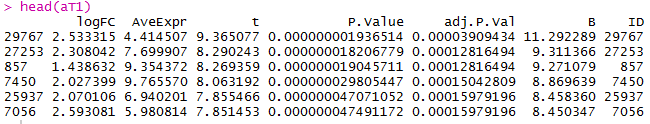
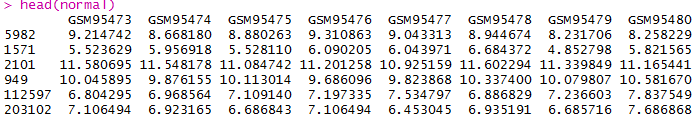
Understanding disease cell behaviors through effector genes using signaling pathways analysis

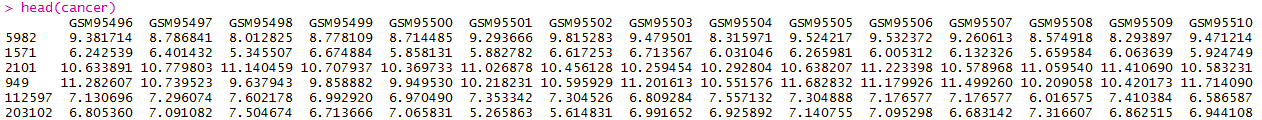
# The input data of SPFA

We demonstrate the functionality of this package using a colorectal cancer dataset obtained using Affymetrix GeneChip technology and available through GEO (GSE4183). The experiment contains8 normal samples and 15 colorectal cancer samples and is described by ([Galamb et al., 2008](#_ENREF_1); [Gyorffy et al., 2009](#_ENREF_2)). RMA preprocessing of the raw data was performed using the affy package, and a two group moderated t-test was applied using the limma package. The data frame obtained as an end result from the function topTable in limma is used as starting point for preparing the input data for SPFA:

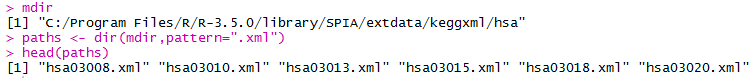


And the original gene expression data is divided into two parts: one contains the expression data from normal samples; one contains the expression data from disease samples:





In the same time, the files path of signaling pathways is also an input of SPFA:

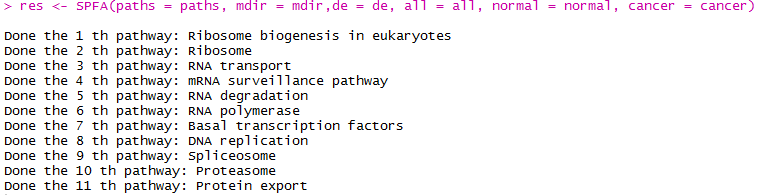


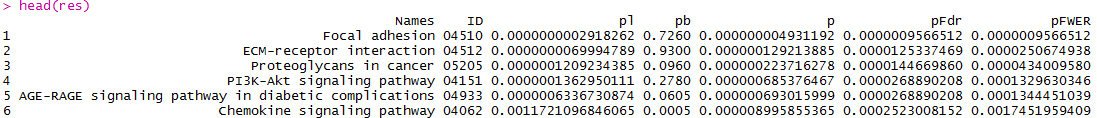
For SPFA to work, we need a vector with all the genes considered to be differentially expressed. The genes must be Entrez gene IDs called de. A vector with the Entrez IDs of all genes profiled on the microarray called all is also need:

C:\Users\Administrator\AppData\Roaming\Tencent\Users\1039919470\QQ\WinTemp\RichOle\YVUDJ80ZWF@[HLPT%VXS8)7.png

# Pathway analysis with SPFA

The SPFA algorithm takes as input the vectors above and produces a table of pathways ranked from the most to the least significant. The SPFA depends on two packages: “KEGGgraph” and “igraph”. The SPFA algorithm can be achieved by calling the SPFA function as follows:

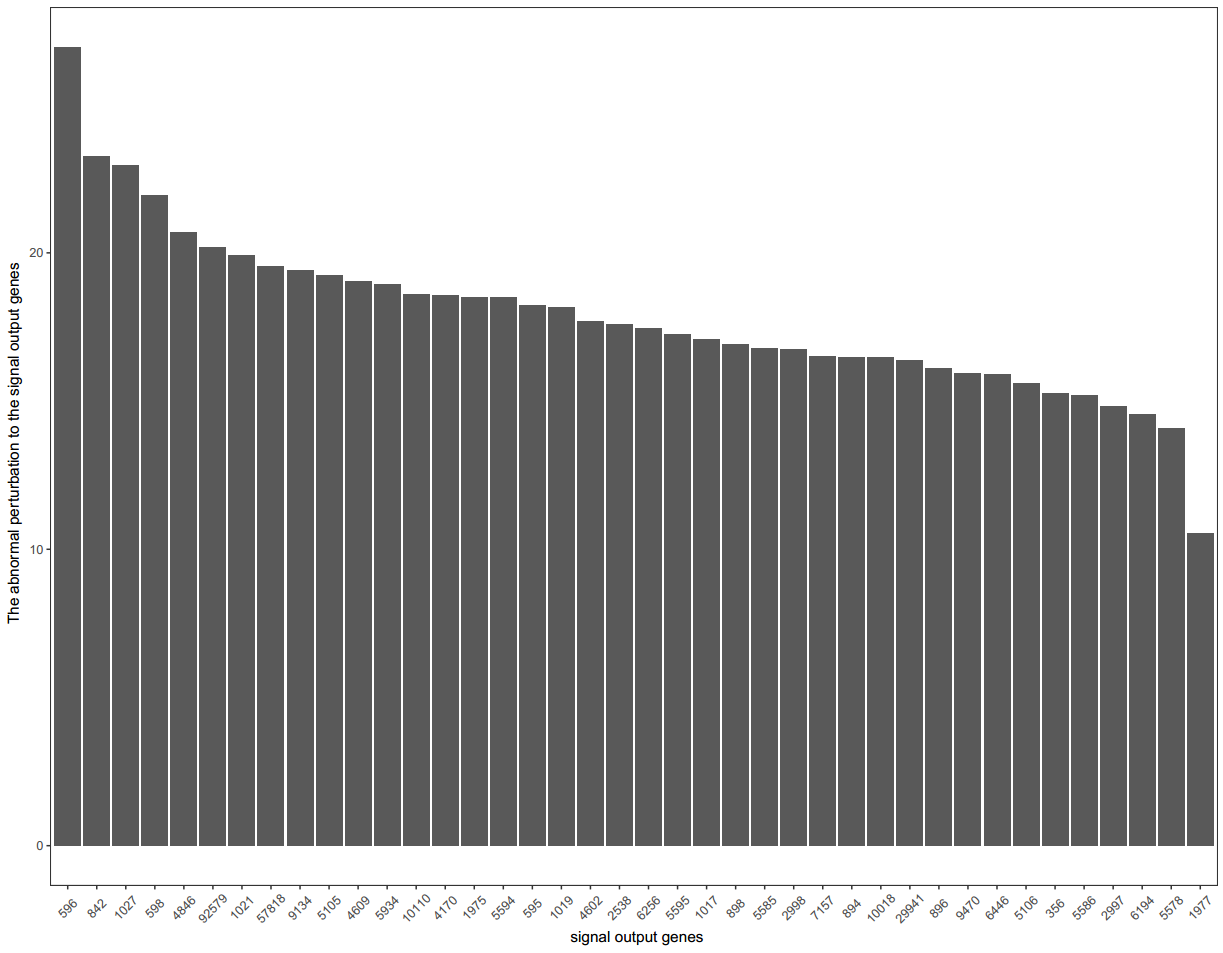




# Analysis of signal output genes in a signaling pathway

To authenticate which signal output gene work in a signaling pathway, we also provide a function called SPFA\_per. The input of SPFA\_per including: mdir, normal, and cancer which mentioned above. And we also need to offer the filename of the signaling pathway. The function will return a figure which contains the signal variations received by effector genes in a signaling pathway. The function SPFA\_per depends on the R package “ggplot2”. The function work in the way as follow:

C:\Users\Administrator\AppData\Roaming\Tencent\Users\1039919470\QQ\WinTemp\RichOle\G@X]P1UPO7GUMM4[OH`_~DA.png



# Reference

Galamb, O., Gyorffy, B., Sipos, F., Spisak, S., Nemeth, A.M., Miheller, P., Tulassay, Z., Dinya, E. and Molnar, B., 2008. Inflammation, adenoma and cancer: objective classification of colon biopsy specimens with gene expression signature. Dis Markers 25, 1-16.

Gyorffy, B., Molnar, B., Lage, H., Szallasi, Z. and Eklund, A.C., 2009. Evaluation of microarray preprocessing algorithms based on concordance with RT-PCR in clinical samples. PLoS One 4, e5645.