444.104 Computational Biology LU - SS 2014

Institute for Genomics and Bioinformatics

Exercise 3: Regulation of gene expression

3.1 MicroRNAs

The following miRNAs have been found to be differentially expressed in an experiment.

miR-124 miR-128 miR-134 miR-9
miR-451 let-7g miR-29b miR-138

3.1.1 Choose one miRNA and gather some information about it like the mature sequence and some references. Use the mirBase database (http://www.mirbase.org/). Explain the nomenclature found.

3.1.2 Find validated target genes of these miRNAs. For this use the miRWalk database (http://www.ma.uni-heidelberg.de/apps/zmf/mirwalk/). Take a look at the file and save it on your PC. Include a small part of it in your protocol.

3.1.3 On the course web site you will find the list of validated target genes resulted from mirWalk as a csv-file. The use of csv-files in R is easier and more stable compared to the use of xls-files. Load this file (be aware that the delimiter used is ";") and create a list with the unique validated target genes. The list provided by MirWalk contains multiple entries for some genes.

3.1.4 Perform a Gene Ontology (GO) analysis on the list of unique validated target genes. We are looking for overrepresented biological processes. As annotation use "hgu133a" (download the hgu133a.db data base). Use a p-value cut-off of 0.05. For this you will need the GOstats package. Install and load the package. How many genes of the validated target genes have annotation in the GO. How big is the universe used for the analysis?

3.1.5 Save your GO analysis to an html file and open this file in a web browser. In this way you can have a look at the overrepresented biological processes. Include a snapshot of the file in your protocol.

3.1.6 The expression level of the miRNAs was measured in patients with traumatic brain injury compared to healthy subjects. The team working on this experiment hypothesized that these miRNA are involved in neuron related biological processes. Can you confirm this?
In order to answer this question you need to read the html file you generated and look for neuron related biological processes. Read the file using the package RCurl. Parse the html content (using the package XML) and find the neuron related biological processes. Include the list of neuron related biological processes into your protocol.