MiRComb: an R package for analyzing miRNA-mRNA interactions

Maria Vila-Casadesús[1,2], Jan Graffelman[3], Meritxell Gironella[2], Juan José Lozano[1,2]

[1]Bioinformatics Platform, CIBERehd (Barcelona, Catalonia, Spain). [2]Gastrointestinal & Pancreatic Oncology Team / CIBERehd / IDIBAPS/ Hospital Clínic de Barcelona (Barcelona, Catalonia, Spain). [3]Department of Statistics and Operations Research, UPC (Barcelona, Catalonia, Spain).

Background and aim

MicroRNAs (miRNAs) are small RNAs that regulate the expression of target mRNAs by specific binding on the mRNA 3'UTR and promoting mRNA degradation in the majority of cases. It is often of interest to know the specific targets of a miRNA in order to study them in a particular disease context.

In that sense, some databases have been designed to predict potential miRNA-mRNA interactions based on hibridization sequences. However, one of the main limitations is that these databases have too many false positives and do not take into account disease-specific interactions.

Our aim was to design an R package able to combine miRNA and mRNA expression data with hibridization information, in order to find potential miRNA-mRNA targets that are more reliable to occur in a specific physiological or disease context.

The miRComb package

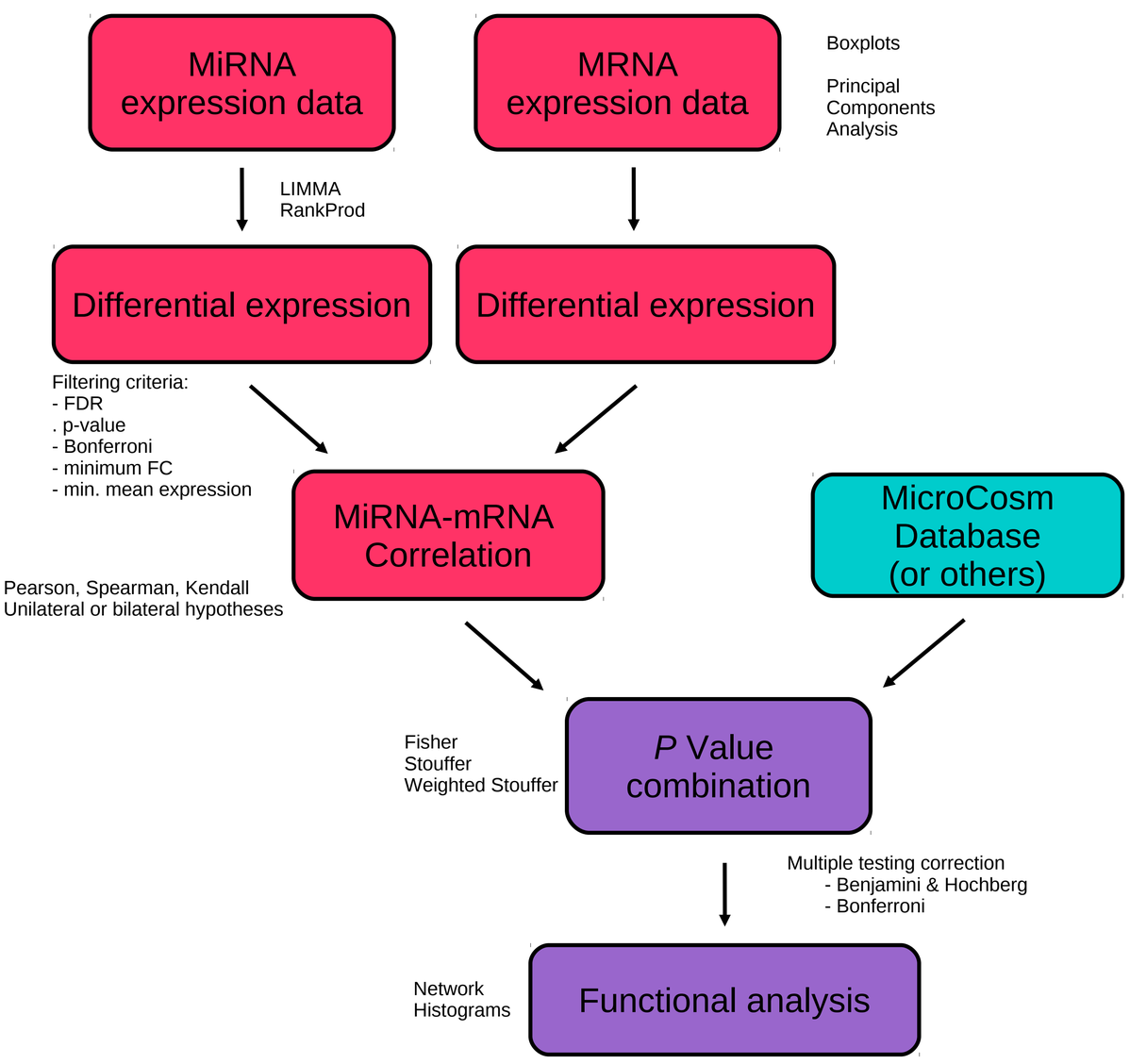

Figure 1. Outline of the pipeline. Basic steps performed in an analysis with the miRComb package.

http://bioinfo.ciberehd.org/mircomb/

Package output example

We tested miRComb in a prostate cancer dataset (GSE21032) containing miRNA and mRNA microarray information from 139 tissue samples (111 Tumors, 28 Controls).

The package first allows to explore the data by Boxplots or Principal Components Analysis (Figure 2).

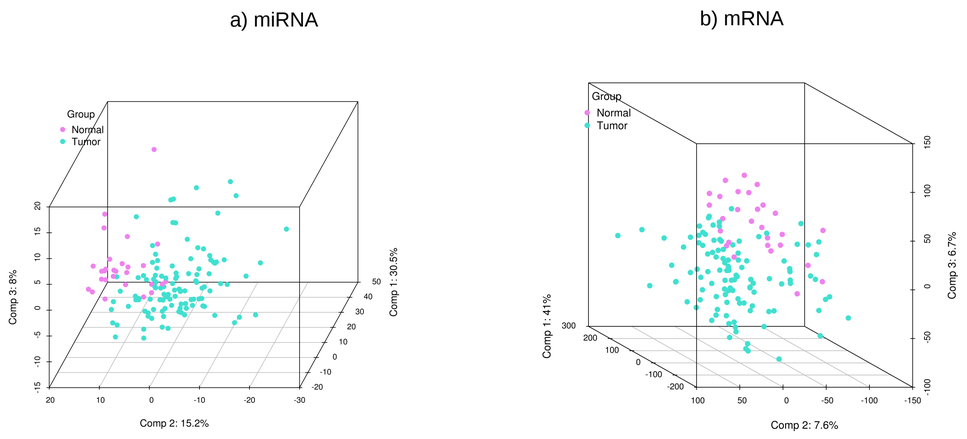

Figure 2. Principal Component Analysis for the miRNA (a) and mRNA (b) expression data showing samples distribution.

The next step is to compute the differentially expressed miRNAs and mRNAs. It is also possible to plot Volcano Plots or Heatmaps of the differentially expressed miRNAs and mRNAs (data not shown). In this dataset there were a total of 185 deregulated miRNAs and 6512 deregulated mRNAs.

Package output example

Then the package carries out correlation measures of the expression of deregulated mRNAs versus the expression of deregulated miRNAs (Figure 3).

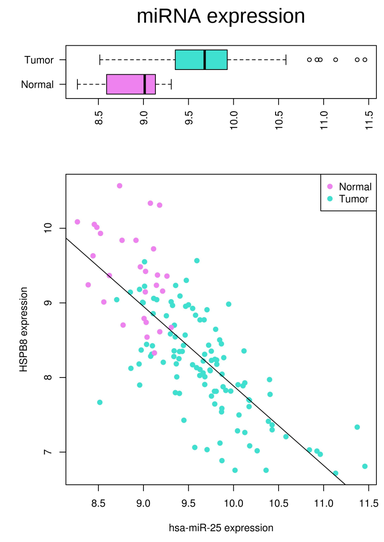

Figure 3. Example of one correlation. Boxplots show that both miRNA and mRNA are deregulated in the disease.

The following step is to select negative miRNA-mRNA correlations and combine the corresponding p-values with predicted p-values from MicroCosm Database. Furthermore, a multiple testing correction method (Benjamini & Hochberg or Bonferroni) is applied in order to obtain a list of all the potentially occurring miRNA-mRNA interactions in a specific disease context. Optionally, a network of the most confident ones can be drawn (Figure 4).

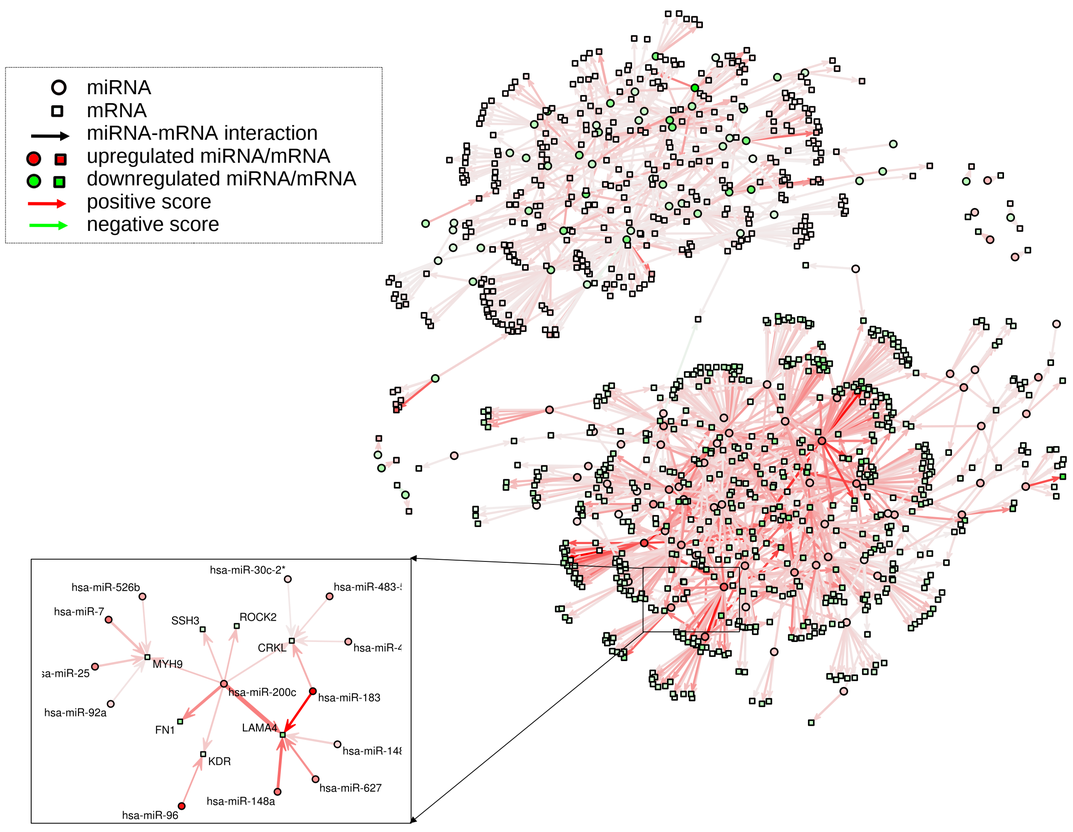

Figure 4. Network of miRNA-mRNA interactions with corrected p-value (Bonferroni) < 0,05.

Finally, the package also allows to explore which are the miRNA hubs (Figure 5) and the overrepresented functions associated with the set of mRNA targets of a specific miRNA (Figure 6). In this dataset each deregulated miRNA targets a mean of 46 mRNAs, while a mRNA is targeted by a mean of 1.3 deregulated miRNAs.

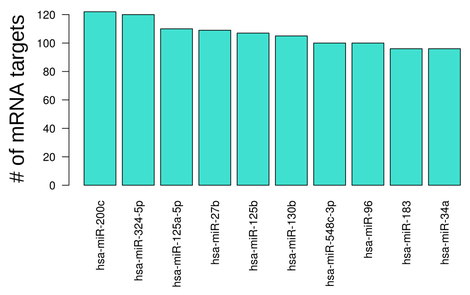

Figure 5. Top 10 miRNAs sorted by number of targets.

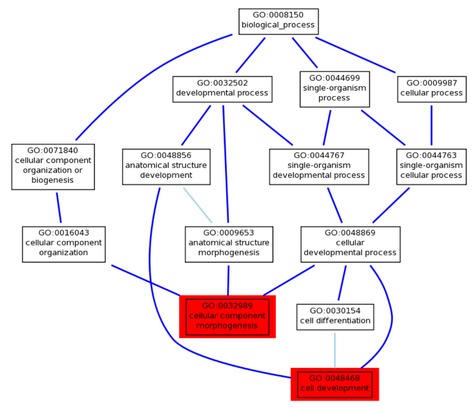

Figure 6. Overrepresented Gene Ontology Biological Functions (red) in the subset of the 122 hsa-miR-200c targets.

Conclusions

- ✓ The miRComb package is a useful tool for finding potential biologically relevant miRNA-mRNA interactions.
- ✓ This package allows to carry out all the required analysis in a single software.
- ✓ The pipeline proposed here filters the high amount of information obtained from the pre-existing miRNA target prediction databases according to a specific disease context.