

# **BABELOMICS: a systems biology perspective in the functional annotation of transcriptomics experiments**

Fátima Al-Shahrour, Pablo Minguez, Joaquín Tárrega, Eva Alloza, Ignacio Medina, José Carbonell, David Montaner, and Joaquín Dopazo

Bioinformatics Department, Centro de Investigación Príncipe Felipe (CIPF), Autopista del Saler 16, E-46013, Valencia, Spain.

## **ABSTRACT**

We present Babelomics, a complete suite of web tools for functional analysis of groups of genes in high-throughput experiments [1-3]. Babelomics includes conventional functional enrichment analysis, where two sets of genes (typically the genes selected in the experiment and the background of genes) are compared for significant functional enrichment and the most sophisticated Gene Set enrichment Methods, where the collective behavior of sets of functionally-related genes is studied in order to find coordinate, significant differential expression [4, 5].

Babelomics uses functional information on gene ontology terms, interpro motifs, KEGG pathways, Biocarta pathways, Swissprot keywords and bioentities defined using textmining methods [6]; regulatory information such as predicted transcription factor binding sites, CisRed motifs, miRNA targets from miRBase; protein-protein interaction analysis, chromosomal positions and presence in tissues with determined histological characteristics.

The analysis is carried out by means of series of integrated modules, some of which were previously published, such as FatiGO [7, 8] or FatiScan [1, 9]. Different visualization options are available, including the possibility of representing KEEG pathways, or the schematic representation of gene-set enrichment experiments in the FatiScan .

The strong multiple-testing nature of the contrasts made by the tools has carefully been taken into account. All the tools are integrated in the gene expression analysis package GEPAS [10-13]. This allows to quickly producing a functional interpretation of different transcriptomics analyses (e.g. genes differentially expressed, genes co-expressing in a cluster, etc.)

Babelomics can be find at: <http://www.babelomics.org>

Usage ratios can be viewed at: <http://bioinfo.cipf.es/webalizer/babelomics/index.html>.

Currently Babelomics has an average of 250 experiments analysed per day.

## References

1. Al-Shahrour F, Arbiza L, Dopazo H, Huerta-Cepas J, Minguez P, Montaner D, Dopazo J: **From genes to functional classes in the study of biological systems.** *BMC Bioinformatics* 2007, **8**:114.
2. Al-Shahrour F, Minguez P, Tarraga J, Montaner D, Alloza E, Vaquerizas JM, Conde L, Blaschke C, Vera J, Dopazo J: **BABELOMICS: a systems biology perspective in the functional annotation of genome-scale experiments.** *Nucleic Acids Res* 2006, **34**(Web Server issue):W472-476.
3. Al-Shahrour F, Minguez P, Vaquerizas JM, Conde L, Dopazo J: **BABELOMICS: a suite of web tools for functional annotation and analysis of groups of genes in high-throughput experiments.** *Nucleic Acids Res* 2005, **33**(Web Server issue):W460-464.
4. Dopazo J: **Functional interpretation of microarray experiments.** *Omic* 2006, **10**(3):398-410.
5. Al-Shahrour F, Dopazo J: **Ontologies and functional genomics.** In: *Data analysis and visualization in genomics and proteomics*. Edited by Azuaje F, Dopazo J: Wiley; 2005: 99-112.
6. Minguez P, Al-Shahrour F, Dopazo J: **A function-centric approach to the biological interpretation of microarray time-series.** *Genome Inform* 2006, **17**(2):57-66.
7. Al-Shahrour F, Diaz-Uriarte R, Dopazo J: **FatiGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes.** *Bioinformatics* 2004, **20**(4):578-580.
8. Al-Shahrour F, Minguez P, Tarraga J, Medina I, Alloza E, Montaner D, Dopazo J: **FatiGO +: a functional profiling tool for genomic data. Integration of functional annotation, regulatory motifs and interaction data with microarray experiments.** *Nucleic Acids Res* 2007, **35**(Web Server issue):W91-96.
9. Al-Shahrour F, Diaz-Uriarte R, Dopazo J: **Discovering molecular functions significantly related to phenotypes by combining gene expression data and biological information.** *Bioinformatics* 2005, **21**(13):2988-2993.
10. Herrero J, Al-Shahrour F, Diaz-Uriarte R, Mateos A, Vaquerizas JM, Santoyo J, Dopazo J: **GEPAS: A web-based resource for microarray gene expression data analysis.** *Nucleic Acids Res* 2003, **31**(13):3461-3467.
11. Herrero J, Vaquerizas JM, Al-Shahrour F, Conde L, Mateos A, Diaz-Uriarte JS, Dopazo J: **New challenges in gene expression data analysis and the extended GEPAS.** *Nucleic Acids Res* 2004, **32**(Web Server issue):W485-491.
12. Montaner D, Tarraga J, Huerta-Cepas J, Burguet J, Vaquerizas JM, Conde L, Minguez P, Vera J, Mukherjee S, Valls J *et al*: **Next station in microarray data analysis: GEPAS.** *Nucleic Acids Res* 2006, **34**(Web Server issue):W486-491.
13. Vaquerizas JM, Conde L, Yankilevich P, Cabezon A, Minguez P, Diaz-Uriarte R, Al-Shahrour F, Herrero J, Dopazo J: **GEPAS, an experiment-oriented pipeline for the analysis of microarray gene expression data.** *Nucleic Acids Res* 2005, **33**(Web Server issue):W616-620.