

The miRò project: towards a unified resource for miRNA research

Laganà A^{1,2}, Giudice A², Forte S³, Veneziano D², Russo F¹, Catania V², Giugno R², Pulvirenti A², Shasha D⁴, Ferro A^{1,2}

Motivation

microRNAs (miRNAs) are small non coding RNAs responsible of post-transcriptional gene regulation. Their crucial role in several physiological and pathological processes has been demonstrated, but their mechanisms of action and functions still remain unclear. Many data banks and bioinformatics tools for miRNA research are currently available on the web but it is fundamental to have systems that integrate data from heterogeneous sources into a unified database equipped with an intuitive and flexible query interface and data mining facilities. miRò is a new system which provides users with miRNA-phenotype associations in humans. miRò is a web-based environment that allows users to perform simple searches and sophisticated data mining queries. The main goal of miRò is to provide users with powerful query tools for finding non-trivial associations among heterogeneous data and thereby to allow the identification of relationships among genes, processes, functions and diseases at the miRNA level. miRò is an ongoing project. Here we describe the new features which are currently being developed.

Methods

The miRò web-site integrates data from different sources. The figure shows the logical architecture of the system. A continuous line indicates an already available feature, a dotted line indicates a module being developed. miRNAs are annotated with information about their precursor and mature sequences coming from miRBase, and with expression profiles obtained from the Mammalian microRNA Atlas. miRNAs are also associated to Gene Ontology terms and diseases through their targets: each miRNA inherits all the annotations of its target genes. Experimentally supported miRNA/target pairs come from miRecords. The predicted targets are taken from the TargetScan, PicTar and miRanda web-sites. The ontological terms with which the target genes are annotated (processes and functions) are obtained from the Gene Ontology Database. The gene-disease relations come from the Genetic Association Database (GAD). Several sources are currently being added to the system. They include other target prediction tools such as MirTarget2, PITA and RNA22, pathway data coming from Pathway Commons

¹ Dept. of Biomedical Sciences, University of Catania, Italy ² Dept. of Mathematics and Computer Sciences, University of Catania, Italy ³ IOM Ricerca, Viagrande, Italy ⁴ Courant Institute of Mathematical Sciences, New York University, U.S.A.

and KEGG, targets polymorphisms from DBsnp, targets expression profiles from Protein Atlas and Human Transcriptome Map. Moreover information about chromosomal fragile sites, translocation breakpoints, CpG islands and Repetitive Elements affecting both miRNAs and their targets are being integrated. We are also developing a new scoring function for miRNA's annotations based on the semantic associations between terms, mined from the biomedical literature. All the data are collected and maintained up-to-date in a MySQL database. The back-end and front-end are being developed in Ruby on Rails.

Results

miRò has been tested on some known cases coming from the literature. It has been able to identify miRNA-disease and miRNA-process associations previously reported. The new data will allow the users to perform more complex and sophisticated queries as well as the automatic extraction of significant knowledge about miRNA functions. The new interface will also provide visualization tools, downloadable report facilities and web services. The main goal of the project is the establishment of a unified resource for miRNA research which allows non trivial analysis through sophisticated mining techniques.

Availability

<http://ferrolab.dmi.unict.it/miro/>

Contact e-mail

lagana@dmi.unict.it

Image

