Online journal ISSN - 1676-5680
GMR Genetics and Molecular Research
EVOLUTION AND TECHNOLOGY

miRNApath: a database of miRNAs, target genes and metabolic pathways

A.O. Chiromatzo[1], T.Y.K. Oliveira[1], G. Pereira[1], A.Y. Costa[2], C.A.E. Montesco[3], D.E. Gras[1], F. Yosetake[4], J.B. Vilar[5], M. Cervato[4], P.R.R. Prado[1], R.G.C.C.L. Cardenas[4], R. Cerri[6], R.L. Borges[4], R.N. Lemos[4], S.M. Alvarenga[7], V.R.C. Perallis[6], D.G. Pinheiro[1], I.T. Silva[1], R.M. Brandão[1], M.A.V. Cunha[1], S. Giuliatti[1] and W.A. Silva Jr[1]

[1]Departamento de Genética, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, SP, Brasil
[2]Departamento de Bioinformática, Universidade de Ribeirão Preto, Ribeirão Preto, SP, Brasil
[3]Departamento de Física e Matemática, Universidade de São Paulo, Ribeirão Preto, SP, Brasil
[4]Departamento de Informática Biomédica, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, SP, Brasil
[5]Departamento de Biologia Geral, Instituto de Ciências Biológicas, Universidade Federal de Goiás, Goiânia, GO, Brasil
[6]Departamento de Estatística, Matemática Aplicada e Computação, Instituto de Geociências e Ciências Exatas, Universidade Estadual Paulista, Rio Claro, SP, Brasil
[7]Instituto de Biotecnologia Aplicada à Agricultura, Universidade Federal de Viçosa, Viçosa, MG, Brasil
Corresponding author: A.O. Chiromatzo
E-mail: alynne@lgmb.fmrp.usp.br

Genet. Mol. Res. 6 (4): 859-865 (2007)
Received August 03, 2007
Accepted September 25, 2007
Published October 05, 2007

ABSTRACT. MicroRNAs (miRNAs) are small non-coding RNAs that regulate target gene expression and hence play important roles in metabolic pathways. Recent studies have evidenced the interrelation of miRNAs with cell proliferation, differentiation, development, and diseases. Since they are involved in gene regulation, they are intrinsically related to metabolic pathways. This leads to questions that are particularly interesting for investigating medical and laboratorial applications. We developed an miRNApath online database that uses miRNA target genes to link miRNAs to metabolic pathways. Currently, databases about miRNA target genes (DIANA miRGen), genomic maps (miRNAMap) and sequences (miRBase) do not provide such correlations. Additionally, miRNApath offers five search services and a download area. For each search, there is a specific type of input, which can be a list of target genes, miRNAs, or metabolic pathways, which results in different views, depending upon the input data, concerning relationships between the target genes, miRNAs and metabolic pathways. There are also internal links that lead to a deeper analysis and cross-links to other databases with more detailed information. miRNApath is being continually updated and is available at http://lgmb.fmrp.usp.br/mirnapath.

Key words: MicroRNA, Metabolic pathways, Target genes, Database, Interaction, Investigation

INTRODUCTION

MicroRNAs (miRNAs) are a class of small, single-stranded non-coding RNAs about ~22 nt in length that regulate gene expression at the post-transcriptional level. Biogenesis starts in the nucleus, where the Drosha enzyme extracts from a long transcript, called primary miRNA (pri-miRNA), with one or more miRNA precursors (pre-miRNA) of ~70 nt. Each of these pre-miRNA is then exported to the cytoplasm, where they are processed by the Dicer enzyme, usually producing a mature miRNA (Bartel, 2004; Kim, 2005; Kim and Nam, 2006).

The mature miRNA enters into an RNA-induced silencing complex and then controls gene expression by cleavage, if they have sufficient complementarity with the mRNA, or by translation repression, if suitable complementary sites are present. Recent research has shown that miRNAs play key roles in many regulatory pathways, such as control of developmental timing, cell proliferation, differentiation, and apoptosis (Wu and Belasco, 2005; Hurst, 2006). Consequently, their deregulation is involved in a series of human diseases, including leukemia, cardiac hypertrophy, and Tourette's syndrome (Di Leva et al., 2006; Cheng et al., 2007).

Due to the complexity of genomic and genetic information, bioinformatics has been widely used for the identification of possible drug targets (Jiang and Zhou, 2005; Ricke et al., 2006). A common approach used to determine the pathway signaling components is to interfere in gene expression. Since miRNAs are important regulators, they can be used for the investigation of potential drugs.

Although it is important to know the influence of miRNA on metabolic pathways, there is no information about their interrelation (Hsu et al., 2006; Shahi et al., 2006; Sethupathy et al., 2006). The Sanger Institute elaborated the miRBase database, which provides information about the genomic coordinates, predicted precursor sequence, mature sequence, and target gene prediction of the mature miRNAs (John et al., 2004; Griffiths-Jones et al., 2006). Other databases, such as miRNAMap, offer new predicted targets (Hsu et al., 2006) and miRGen centralizes genomic and target gene information (Megraw et al., 2007).

The main purpose of the miRNApath online database is to provide information about the influence of miRNA on metabolic pathways. This kind of study is made possible by using one of the six web services made available. The search result includes cross-links to other databases (miRBase, KEGG and Entrez) for detailed information about the gene, metabolic pathway and miRNA. The links within miRNApath include further information about the search.

MATERIAL AND METHODS

The miRNApath database is built based on well-known databases. Its main stored information is the miRNA data from the miRBase ftp site (http://microrna.sanger.ac.uk/sequences/), the target genes from the miRGen assembly (http://www.diana.pcbi.upenn.edu/miRGen.html) and the metabolic pathways network obtained from the KEGG database, using the SOAP/WSDL interface (http://www.genome.jp/kegg/soap/).

Nowadays, the most important institutions involved are the National Center for Biological Information (NCBI) and the University of California at Santa Cruz. Usually, each institution establishes its own code(s) to identify and retrieve information from their databases. Since miRNApath works with other databases, there is a problem with data denomination heterogeneity. Hence, miRNApath adopted the gene symbol, the miRNA name, the metabolic pathway name, and the NCBI access number (GenBank database), as well as geneid (Refseq database) and unigene (Unigene database) to represent the data. miRNApath handles only those organisms that are in all databases. The selected organisms are *Homo sapiens*, *Mus musculus*, *Rattus norvegicus*, and *Canis familiaris*.

Integration of all database information is made using the Application Programming Interfaces supplied by the institutions in Perl language to parse the downloaded data. This is a meticulous task that demands caution in order to maintain miRNApath consistency. In miRNApath, all the target genes from the miRGen and KEGG metabolic network genes must have equivalent code identifications from NCBI databases. Also, all the miRNAs of miRGen must be registered in the miRBase. The miRNApath is stored in a Mysql database and the web interface was developed using PHP language.

RESULTS

The miRNApath offers an online interface with six web services separated into the categories: QUICK SEARCH, which returns superficial information and ADVANCED ANALYSIS, which presents a detailed result. Figure 1A illustrates the main page. QUICK SEARCH has three options: Search by Gene, Search by miRNA and Search by Pathways. The difference between them is the object of study; for example, if the user wants to analyze a target gene list,

he must click the link “Search by Gene”, and so on. Figure 1B shows the Search by Gene site. To make a search, the user has to write a target gene list or upload a file, specify which organism and format what the input refers to and then press the search button to submit the data. The figures below show an example of how to make a search and how the result will appear. The other searches are analogous.

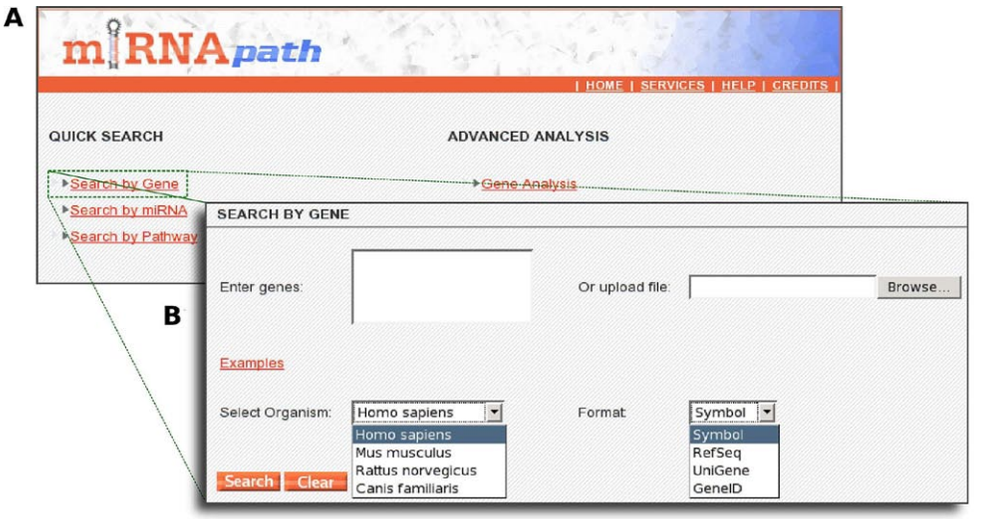

Figure 1. A. Main page. The green box links to the Search by Gene site. **B.** Search by Gene input page. The user can write a list of target genes or upload a file and then select the corresponding organism and input data type.

The Search by Gene result is given in Figure 2A. It shows, for each target gene entered, the quantity of miRNAs that target the gene and the number of pathways in which the gene is involved. The orange links lead the user to internal sites with detailed information, and the blue links direct the user to outside databases. If the user clicks on miRNAs amount (blue box), it will open another site (Figure 2B) with target gene AXIN2 information, including all the 17 miRNAs that affect this gene and the number of pathways for which each miRNA has a target gene. If the user clicks on the pathways amount (green box), he will indicate a site (Figure 2C) with information about the AXIN2 target gene, the hsa-miR-449 miRNA and the name of all the 94 pathways in which the miRNA has a target gene. The links in dark red mean that the gene AXIN2 is regulated by the hsa-miR-449 miRNA and it participates in the Wnt signaling pathway.

Finally, we present an example of the ADVANCED ANALYSIS. The gene analysis search and the gene search have the same input interface, but with different results. Figure 3A shows the search result site, which has three ways of displaying data that are ordered by pathway, gene and miRNA. Figure 3B shows the order by pathway view. It displays the pathways in which the target genes of the input list are involved, the number of target genes from the input list that are in the pathway and the number of miRNAs that affect these target genes. In the

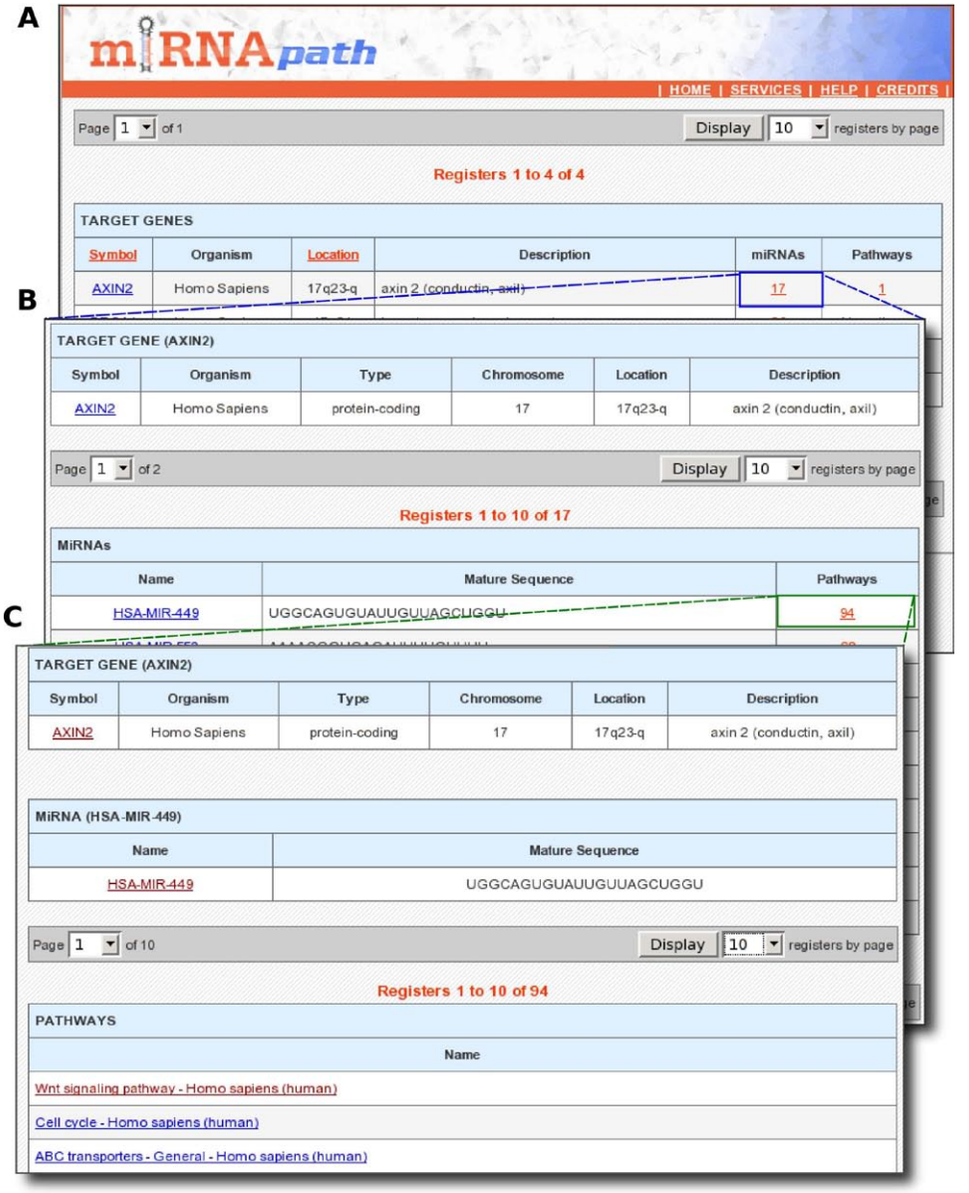

Figure 2. A. The result of the Search by Gene. **B.** Detailed site with information about the AXIN2 gene and all the miRNAs that target this gene. **C.** Information about the target gene AXIN2, the hsa-miR-449 miRNA and all the metabolic pathways in which this miRNA targets a gene. The links in dark red mean that the gene AXIN2 is regulated by the hsa-miR-449 miRNA and it participates in the Wnt signaling pathway.

example, the MAPK signaling pathway has two genes from the input target list and there are 44 miRNAs that regulate those genes. If the user clicks on amount of genes, a site like Figure 3C will appear. This site has information about the pathway, the genes from the input list that belong to the pathway and the respective miRNAs that repress the genes.

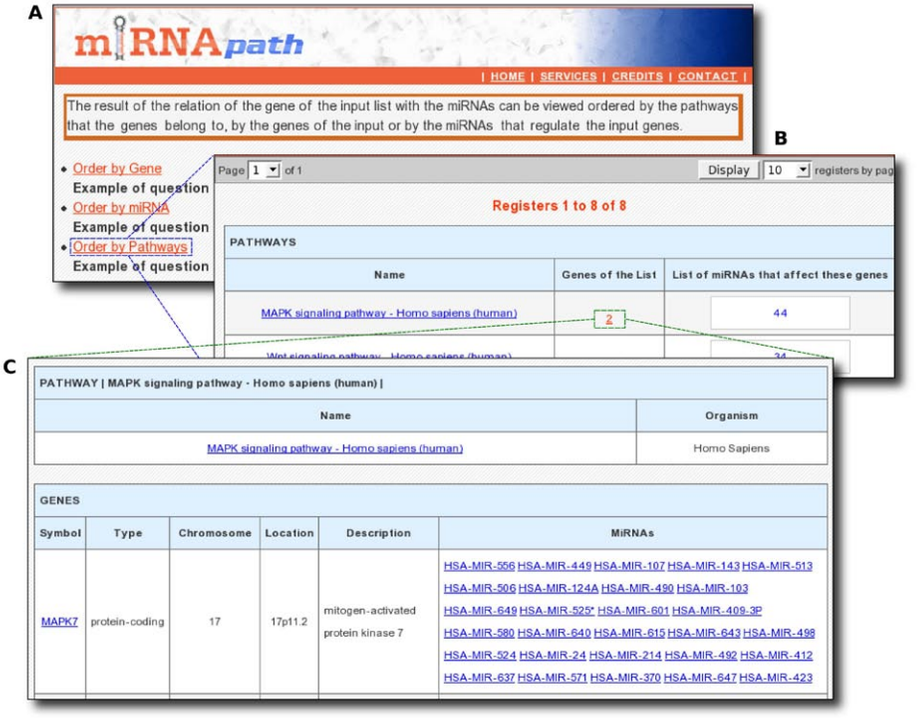

Figure 3. A. Query result with the three viewing options. **B.** Order by pathway view shows the quantity of target genes from the input list and the quantity of miRNAs that affect these genes. **C.** Information about the genes of the list that belong to the selected pathway.

DISCUSSION

The miRNApath handles many questions about the relationship between miRNAs, target genes and metabolic pathways. Investigating miRNAs, one can discover the pathways affected by these miRNAs and possibly determine a disease caused by them. Or, by analyzing a list of target genes, it is possible to find an miRNA (or a set of them), which when deregulated could be the cause of a certain symptom. In conclusion, miRNApath provides an easy and friendly means of investigating a large number of questions about this relationship. Soon, we intend to allow the user to download the results from the searches, to implement various searches, to add new input formats, and to provide other useful information.

ACKNOWLEDGMENTS

Research supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento Pessoal de Nível Superior (CAPES), Financiadora de Estudos e Projetos (FINEP), and Fundação de Apoio ao Ensino, Pesquisa e Assistência do Hospital das Clínicas FMRP-USP (FAEPA).

REFERENCES

Bartel DP (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116: 281-297.

Cheng Y, Ji R, Yue J, Yang J, et al. (2007). MicroRNAs are aberrantly expressed in hypertrophic heart: do they play a role in cardiac hypertrophy? *Am. J. Pathol.* 170: 1831-1840.

Di Leva G, Calin GA and Croce CM (2006). MicroRNAs: fundamental facts and involvement in human diseases. *Birth Defects Res. C. Embryo Today* 78: 180-189.

Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, et al. (2006). miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* 34: D140-D144.

Hsu PW, Huang HD, Hsu SD, Lin LZ, et al. (2006). miRNAMap: genomic maps of microRNA genes and their target genes in mammalian genomes. *Nucleic Acids Res.* 34: D135-D139.

Hurst LD (2006). Preliminary assessment of the impact of microRNA-mediated regulation on coding sequence evolution in mammals. *J. Mol. Evol.* 63: 174-182.

Jiang Z and Zhou Y (2005). Using bioinformatics for drug target identification from the genome. *Am. J. Pharmacogenomics* 5: 387-396.

John B, Enright AJ, Aravin A, Tuschl T, et al. (2004). Human MicroRNA targets. *PLoS. Biol.* 2: e363.

Kim VN (2005). MicroRNA biogenesis: coordinated cropping and dicing. *Nat. Rev. Mol. Cell Biol.* 6: 376-385.

Kim VN and Nam JW (2006). Genomics of microRNA. *Trends Genet.* 22: 165-173.

Megraw M, Sethupathy P, Corda B and Hatzigeorgiou AG (2007). miRGen: a database for the study of animal microRNA genomic organization and function. *Nucleic Acids Res.* 35: D149-D155.

Ricke DO, Wang S, Cai R and Cohen D (2006). Genomic approaches to drug discovery. *Curr. Opin. Chem. Biol.* 10: 303-308.

Sethupathy P, Corda B and Hatzigeorgiou AG (2006). TarBase: a comprehensive database of experimentally supported animal microRNA targets. *RNA* 12: 192-197.

Shahi P, Loukianiouk S, Bohne-Lang A, Kenzelmann M, et al. (2006). Argonaute - a database for gene regulation by mammalian microRNAs. *Nucleic Acids Res.* 34: D115-D118.

Wu L and Belasco JG (2005). Micro-RNA regulation of the mammalian lin-28 gene during neuronal differentiation of embryonal carcinoma cells. *Mol. Cell Biol.* 25: 9198-9208.