BIOINFORMATICS APPLICATIONS NOTE Vol. 27 no. 23 2011, pages 3329–3330
doi:10.1093/bioinformatics/btr556

Databases and ontologies

Advance Access publication October 7, 2011

miREnvironment Database: providing a bridge for microRNAs, environmental factors and phenotypes

Qingqing Yang[1,2,†], Chengxiang Qiu[1,†], Jie Yang[1,2,†], Qing Wu[2] and Qinghua Cui[1,3,4,*]

[1]Department of Biomedical Informatics, Peking University Health Science Center, Beijing 100191, [2]School of Computer Science, Hebei University of Technology, Tianjin 300401, [3]Institute of Systems Biomedicine and [4]MOE Key Lab of Molecular Cardiovascular Science, Peking University, Beijing 100191, China

Associate Editor: Ivo Hofacker

ABSTRACT

The interaction between genetic factors and environmental factors has critical roles in determining the phenotype of an organism. In recent years, a number of studies have reported that the dysfunctions on microRNA (miRNAs), environmental factors and their interactions have strong effects on phenotypes and even may result in abnormal phenotypes and diseases, whereas there has been no a database linking miRNAs, environmental factors and phenotypes. Such a resource platform is believed to be of great value in the understanding of miRNAs, environmental factors, especially drugs and diseases. In this study, we constructed the miREnvironment database, which contains a comprehensive collection and curation of experimentally supported interactions among miRNAs, environmental factors and phenotypes. The names of miRNAs, phenotypes, environmental factors, conditions of environmental factors, samples, species, evidence and references were further annotated. miREnvironment represents a biomedical resource for researches on miRNAs, environmental factors and diseases.

Availability: http://cmbi.bjmu.edu.cn/miren.

Contact: cuiqinghua@hsc.pku.edu.cn

Received on July 7, 2011; revised on September 14, 2011; accepted on October 2, 2011

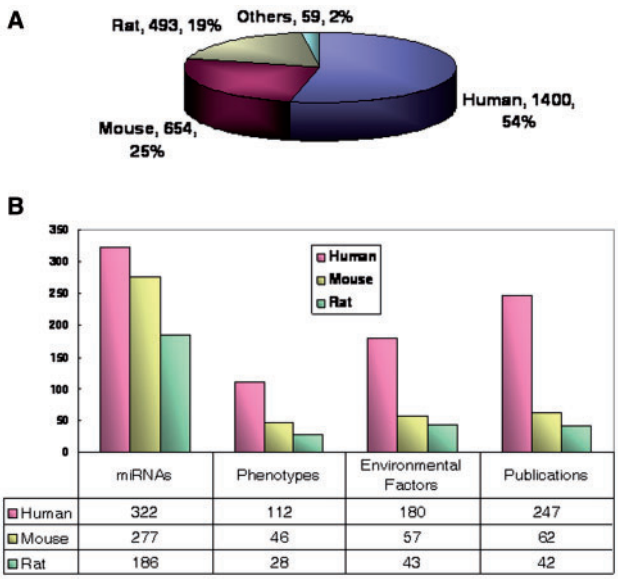
Phenotypes of an organism are determined by the complex interactions between genetic factors and environmental factors. Environmental factors contribute a lot to the formation and development of many diseases, especially complex diseases including cancer and cardiovascular diseases (Catania *et al.*, 2009; Chow *et al.*, 2010; Das, 2010; Soto and Sonnenschein, 2010). A number of databases for coding-gene and environment associations have been developed (Kitsios and Zintzaras, 2010; Mattingly *et al.*, 2006; Turner *et al.*, 2010). Furthermore, they have provided great helps in biomedical researches. MicroRNAs (miRNAs) are one class of newly identified genetic factors, which mainly repress the expression of genes at the post-transcriptional level (Bartel, 2004). miRNAs play critical roles in various biological processes, including cell growth, proliferation, differentiation, development and apoptosis (Esquela-Kerscher and Slack, 2006). Dysfunctions of miRNAs are thus associated with various diseases (Lu, 2008). Like other genetic factors, it has been reported in recent years that miRNAs also have complex interactions with a wide spectrum of experimental factors including stress (Gidron *et al.*, 2010), drugs (Lima *et al.*, 2011), virus (Lin and Flemington, 2011), alcohol (Ladeiro *et al.*, 2008), cigarette (Izzotti *et al.*, 2010), air pollution (Jardim, 2011), radiation (Niemoeller *et al.*, 2011), diet (Alisi *et al.*, 2011) etc. These interactions have crucial roles in many phenotypes including diseases. The studies on experimental factor and miRNA associations are becoming increasingly important in biomedical sciences. Therefore, a database linking miRNAs, experimental factors and phenotypes becomes emergently needed but is still not available.

To meet such need, we searched the PubMed database for literature that matched this study by the union of two keyword sets. One keyword set is 'miR or miRNA or microRNA', which ensures that literature about miRNA study is retrieved. The other keyword set contains a list of experimental factors (Supplementary File S1). We further manually curated miRNAs and environmental factors that have associations. The researchers read the original references and manually summarized the data. The condition of environmental factors, treated samples, species, the details of associations and the reference PubMed ID are also manually retrieved. The data are further manually standardized and annotated. Items (i.e. miRNA genes) that cannot be standardized or annotated are represented by 'n/a'. Furthermore, different researchers double checked the data including miRNAs environmental factors, conditions of environmental factors, phenotypes, evidences describing the relationships among miRNAs, environmental factors and phenotypes, and references. Based on above data, we constructed the miREnvironment database. As a result, miREnvironment integrated >2500 entries including ~800 miRNAs, ~260 experimental factors, ~180 phenotypes, 17 species from ~370 publications. Human, mouse and rat are the top three species that have the greatest numbers of entries. They represent 98% of the total entries (Fig. 1A). We also list the statistical details for data of human, mouse and rat (Fig. 1B). The database also provides hyperlinks to the original references in NCBI (http://www.ncbi.nlm.nih.gov/) for each entry.

All data were organized in the 'miREnvironment' database using SQLite, a lightweight database management system. The website is presented using Django, a Python web framework and is available at http://cmbi.bjmu.edu.cn/miren. miREnvironment contains pages for browsing, searching, submitting and downloading.

*To whom correspondence should be addressed.

†The authors wish it to be known that, in their opinion, the first three authors should be regarded as joint First Authors.



	miRNAs	Phenotypes	Environmental Factors	Publications
Human	322	112	180	247
Mouse	277	46	57	62
Rat	186	28	43	42

Fig. 1. Statistics and distribution of data in miREnvironment. (**A**) Entry distribution in different species. (**B**) Number of miRNAs, phenotypes, environmental factors, and publications for human, mouse and rat.

miREnvironment provides several search options such as names of the miRNAs, environmental factors and phenotypes.

The miREnvironment database houses a manually curated collection of experimentally supported interactions of miRNAs, experimental factors and phenotypes in >15 species. We believe that miREnvironment will not only be useful for bioinformatics analysis but also for the scientists who are interested in miRNA function and disease, medicine and pharmacology. However, miREnvironment represents the first step in this project and further extensions will be developed. We will try to annotate the environmental factors in more details. For example, the drugs will be linked to some drug resources, i.e. DrugBank (http://www.drugbank.ca/). By linking to DrugBank, the users can easily know the features of interested drugs, such as chemical formula and structure, etc. It is also useful to incorporate the dysfunctional protein-coding genes for entries at the same experimental conditions as the miRNAs have. Furthermore, some analysis tools are being developed and will be integrated into miREnvironment in the future. For example, we are developing tools to identify new disease indications for FDA-approved drugs, the so-called drug repositioning (Sirota *et al.*, 2011). The drug–miRNA interaction represents a novel dimension of information to drugs, and is expected to be useful in drug repositioning. We plan to continuously update the miREnvironment.

ACKNOWLEDGEMENTS

We thank Prof. Edwin Wang at the Biotechnology Research Institute, National Research Council Canada for language corrections.

Funding: Natural Science Foundation of China (Grant No. 30900829).

Conflict of Interest: none declared.

REFERENCES

Alisi,A. *et al.* (2011) Mirnome analysis reveals novel molecular determinants in the pathogenesis of diet-induced nonalcoholic fatty liver disease. *Lab. Invest.*, **91**, 283–293.

Bartel,D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 281–297.

Catania,A.S. *et al.* (2009) Vitamins and minerals with antioxidant properties and cardiometabolic risk: controversies and perspectives. *Arq. Bras. Endocrinol. Metabol.*, **53**, 550–559.

Chow,W.H. *et al.* (2010) Epidemiology and risk factors for kidney cancer. *Nat. Rev. Urol.*, **7**, 245–257.

Das,U.N. (2010) Obesity: genes, brain, gut, and environment. *Nutrition*, **26**, 459–473.

Esquela-Kerscher,A. and Slack,F.J. (2006) Oncomirs - microRNAs with a role in cancer. *Nat. Rev. Cancer*, **6**, 259–269.

Gidron,Y. *et al.* (2010) Influence of stress and health-behaviour on miRNA expression. *Mol. Med. Report*, **3**, 455–457.

Izzotti,A. *et al.* (2010) Modulation of microRNA expression by budesonide, phenethyl isothiocyanate and cigarette smoke in mouse liver and lung. *Carcinogenesis*, **31**, 894–901.

Jardim,M.J. (2011) microRNAs: implications for air pollution research. *Mutat Res.*

Kitsios,G.D. and Zintzaras,E. (2010) Synopsis and data synthesis of genetic association studies in hypertension for the adrenergic receptor family genes: the CUMAGAS-HYPERT database. *Am. J. Hypertens.*, **23**, 305–313.

Ladeiro,Y. *et al.* (2008) MicroRNA profiling in hepatocellular tumors is associated with clinical features and oncogene/tumor suppressor gene mutations. *Hepatology*, **47**, 1955–1963.

Lima,R.T. *et al.* (2011) MicroRNA regulation of core apoptosis pathways in cancer. *Eur. J. Cancer*, **47**, 163–174.

Lin,Z. and Flemington,E.K. (2011) miRNAs in the pathogenesis of oncogenic human viruses. *Cancer Lett.*, **305**, 186–199.

Lu,M. *et al.* (2008) An analysis of human microRNA and disease associations. *PLoS One*, **3**, e3420.

Mattingly,C.J. *et al.* (2006) The comparative toxicogenomics database: a cross-species resource for building chemical-gene interaction networks. *Toxicol. Sci.*, **92**, 587–595.

Niemoeller,O.M. *et al.* (2011) MicroRNA expression profiles in human cancer cells after ionizing radiation. *Radiat. Oncol.*, **6**, 29.

Sirota,M. *et al.* (2011) Discovery and preclinical validation of drug indications using compendia of public gene expression data. *Sci. Transl. Med.*, **3**, 96ra77.

Soto,A.M. and Sonnenschein,C. (2010) Environmental causes of cancer: endocrine disruptors as carcinogens. *Nat. Rev. Endocrinol.*, **6**, 363–370.

Turner,S.W. *et al.* (2010) A methodology to establish a database to study gene environment interactions for childhood asthma. *BMC Med. Res. Methodol.*, **10**, 107.