A dry lab for exploring miRNA functions and applications in cancer subtype discovery

Thuc Duy Le [1,*], Lin Liu [1], Taosheng Xu [2], Junpeng Zhang [3], Jiuyong Li [1]

[*1]School of Information Technology and Mathematical Sciences,
University of South Australia, Mawson Lakes, SA 5095, AU
[2]Institute of Intelligent Machines,
Hefei Institutes of Physical Science,
Chinese Academy of Sciences, Hefei 230031, China
[3]School of Engineering, Dali University, China.

November 17, 2017

microRNAs (miRNAs) are important gene regulators controlling a wide range of biological processes and are involved in several types of cancers. Computational methods are proved to be an effective approach to exploring miRNA functions by predicting miRNA-mRNA regulatory relationships.

In the last 2 years, we have developed 4 Bioconductor packages, miRLAB[3], miRSponge, miRBaseConverter, and CancerSubtypes[4], for exploring miRNA functions and for cancer subtype discovery based on miRNA and mRNA expression data. We briefly describe the packages in the following.

1. miRLAB provides a set 12 computational methods, including our causal inference method[1], for predicting miRNA targets from expression data. The package also provides tools for validating the predictions, comparing different methods, and incorportating target binding information to the expression-based methods.

2. miRsponge provides 7 computational methods for identifying miRNA sponge interactions and 4 methods for finding miRNA sponge modules (see [2] for the review). Tools for validating competing endogenous interactions and functional enrichment analyses are also provided in the package.

3. miRBasedConverter provides tools for converting and retrieving miRNA names, ID, and other information in different versions of miRBase. The package will be useful for working with multiple miRNA datasets where different miRBase versions were used for the miRNA naming convention.

*to whom correspondence should be addressed

4. CancerSubtypes provides 6 machine learning methods for identifying cancer subtypes from data, including our own method of utilising miRNA-Transcription Factor-mRNA network information in stratifying cancer subtypes[5]. The package also contains 4 pre-selecting gene methods and 4 stastistical methods for validating the performance of the subtyping methods.

In this talk, I will present the motivations and utilities of the 4 Bioconductor packages. I will demonstrate the ease of use in function calls of those packages and some typical workflows and scenarios of using the packages.

References

[1] Thuc Duy Le, Lin Liu, Anna Tsykin, Gregory J Goodall, Bing Liu, Bing-Yu Sun, and Jiuyong Li. Inferring microRNA–mRNA causal regulatory relationships from expression data. *Bioinformatics*, 29(6):765–771, 2013.

[2] Thuc Duy Le, Junpeng Zhang, Lin Liu, and Jiuyong Li. Computational methods for identifying miRNA sponge interactions. *Briefings in bioinformatics*, p. bbw042, 2016.

[3] Thuc Duy Le, Junpeng Zhang, Lin Liu, Huawen Liu, and Jiuyong Li. mirlab: An r based dry lab for exploring mirna-mrna regulatory relationships. *PloS one*, 10(12):e0145386, 2015.

[4] Taosheng Xu, Thuc Duy Le, Lin Liu, Ning Su, Rujing Wang, Bingyu Sun, Antonio Colaprico, Gianluca Bontempi, and Jiuyong Li. Cancersubtypes: an r/bioconductor package for molecular cancer subtype identification, validation, and visualization. *Bioinformatics*, 2017.

[5] Taosheng Xu, Thuc Duy Le, Lin Liu, Rujing Wang, Bingyu Sun, and Jiuyong Li. Identifying cancer subtypes from mirna-tf-mrna regulatory networks and expression data. *PloS one*, 11(4):e0152792, 2016.