

Computation Approach on Mirna Research: Plethora of Opportunities and Adding New Dimension in Biology

Sourav Pakrashi¹, Jyotirmoy Rakshit¹ and Sandipan Chakraborty^{2*}

¹School of Biotechnology & Biological Sciences, West Bengal University of Technology, BF-142, Salt Lake, Sector-1, Kolkata 700064, India

²Department of Microbiology, University of Calcutta, 35 Ballygunge Circular Road, Kolkata 700 019, India

***Corresponding Author:** Sandipan Chakraborty, Department of Microbiology, University of Calcutta, 35 Ballygunge Circular Road, Kolkata 700 019, India; E-mail: sandipanchakraborty.13@gmail.com

Introduction

Early 20th century, zoology, botany, physiology and ecology were the foremost cult-classic biological areas where observations are made by morphological, histological and anatomical based studies and cross species interactions observed in nature. Observations made either by naked eye and with the aid of microscope were key tools used in scientific research. Data originated from diverse fields of science was unable to blend in a proper manner to achieve a new perception to explore complex biological system. But it is needless to say that the intrinsic biological system is highly complex by nature. Cell, tissue, organ, organisms are interconnected and controlling functions of each other, thus a system based approach should be the prime goal of modern biology [1].

1970s witnessed a tremendous growth both in computational hardware as well as software applications. Applications of computation and High-Throughput Technology (HTS) in research arena have altered the scenario immensely. Recent development in software applications as well as algorithm and with the aid of database technology, it is now

possible to manage gigantic amount of data generated through HTS technology. Modern biology is now a combinatorial approach where application of *in silico* prediction tools on experimental data provides details insight on complex biological systems. Primary and secondary databases are created to accumulate wet lab experiments derived data in an ordered manner and to facilitate its easy retrieval. Extensive *in silico* tools are used in genomic constructs, prediction of active and regulatory parts of gene like exons, introns and [2] and inference of gene regulatory from expression profile [3-5]. On the other hand, computer derived models, can be used to create prediction to be tested by experiments, as well as, to explore questions that are not amenable to experimental inquiry [1].

Here, we will focus in great detail on the applications of bioinformatics based approach on micro RNA research.

The Micro RNA Revolution

According to the central dogma of molecular biology, RNA was initially considered as passive intermediate between DNA and protein. But discovery of miRNA, small quantity yet versatile RNA molecule, open up new dimensions in post-transcription and post translational gene expression. miRNAs are small, noncoding RNA molecules of ~22 nucleotides, which typically repress messenger RNAs (mRNA) by binding to their 3' UTR [6 - 9]. The specificity of interaction is determined by a stretch of six to eight nucleotides in the 5' end of miRNA, it is

referred as seed region, which has perfect complementarity to the target mRNA [10]. This recognition activates the RISK complex which is followed by adenylation, decapping and/or transitionally repressed by inhibition of initiation and elongation [11, 12] leading to the degradation of the targeted mRNA. Many plant miRNAs are mostly involved in transcriptional regulation, whereas the regulatory impact of animal miRNAs are more pervasive, with roles identified during developmental stage, cell death, cell proliferation, hematopoiesis, and differentiation of the nervous system. Currently more than 1500 human miRNAs as well as hundreds of viral miRNAs have been identified. They are believed to play crucial role in regulating expressions of dozens of genes [10].

Application of Computational Approach and its Limitation in miRNA Research

miRNAs are found across different organisms including plant, animal and even in viruses. Its small size makes it difficult to clone and study its expression. Application of bioinformatics based approach has been widely used in miRNA research from early days. Homology analysis revealed critical importance of stem-loop site adjacent to miRNA gene to attain proper secondary structure of premature miRNA. Generally, *in silico* tools are used to search potential miRNA gene and identification of probable miRNA target across the genome. Basic idea behind this searching is to apply the knowledge of already experiment derived miRNA-mRNA duplex, especially its sequence and structural homology to annotate novel miRNA. But this approach suffers due to complexity of biological system. Increase specificity and decrease false positive target site across genome are two biggest challenges which need to be addressed. Since length of miRNA is very small and quantity of existing hairpin loop creating DNA sequence is too large. Thus more stringent algorithm development along with additional filtering is needed to enhance the specificity [13].

Another computational challenge is the lack of information of miRNA expression *in vivo* [14]. Expression level of mRNA as well as miRNA tremendously varied with different tissues, physical, physiological condition and developmental

stage of the organism. Epigenetic modification is difficult to predict with the aid of computer. Post transcriptional and translational modification of mRNA may only be identified using a cloning approach [15, 16]. So experimental validations are needed to confirm existence of *in silico* predicted miRNA. MirBase and Targetscan are two most widely used *in silico* tools applied in miRNA research.

MirBase:- MirBase is a dedicated interface for miRNA research and also perform as a centralized nomenclature system of new miRNA gene. This database contains all experimentally published mature miRNA sequence together with their predicted sources, hairpin loop precursors, annotation relating to their discovery and function [15]. The number of submitted miRNA hairpin loci in MirBase is growing rapidly, from 2,906 hairpin loci (June, 2005, release 7) to nearly 24,521 (June, 2013, release 20) (Figure 1) [7, 15].

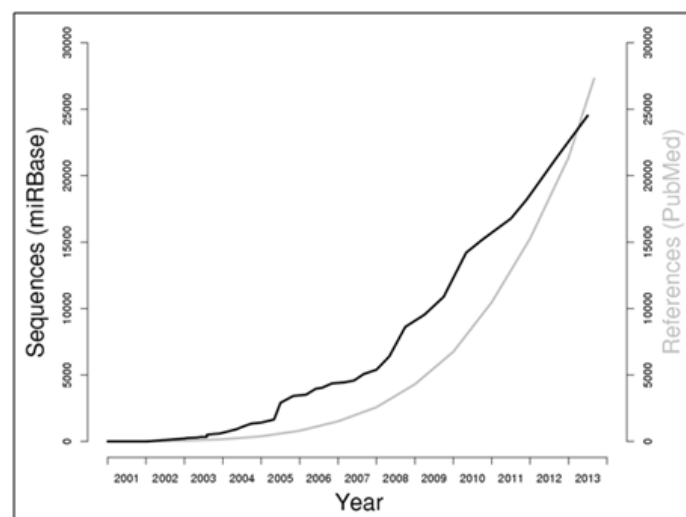


Figure 1: Extensive research on miRNA causes rapid submission of unique stem-loop miR gene sequence in mirBase. (<http://www.mirbase.org/blog>)

Targetscan: - Animal miRNA and its target 3' UTR mRNA can be determined by applying Targetscan (Release 6.2, June 2012), a computation based approach. Binding free energy between seed region of miRNA and 3' UTR of target mRNA is the fundamental criteria to achieve conserveness across vertebrate species. miRNA of 10 different species such as human, mouse,

rat, dog, chicken, chimpanzee, rhesus, cow, opossum and frog can be identified through Targetscan.

Conclusion

Computation is now playing a pivotal role in biological research. Applications of bioinformatics tools scale up many time consuming processes such as nucleic acid and protein sequencing, phylogenetic analysis, annotation of genomic constructs, identification of miRNA sequence with its target, structure prediction of macromolecules, probable interacting site(s) of enzymes and ligands, to name a few. Tools such as miRBase and Targetscan have given an impetus to miRNA research. Algorithms have been developed which aids in resolving the complexity of biological systems. Now systems biological approaches are coming up where rather looking at a particular process, we can simulate complex biological networks, metabolic pathways to provide complex biological processes in greater details. In this aspects development of computational powers as well as algorithms helps biological research to take big leap in solving complex biological phenomenon.

References

- Hiroaki K (2002). Computational systems biology. *Nature* 420: 206-210.
- Baldi P, Brunak S. *Bioinformatics: The Machine Learning Approach*. 2nd ed. Cambridge, MIT Press; 299-320.
- Onami S, Kyoda KM, Morohashi M, Kitano H. 2001. The DBRF Method for Inferring a Gene Network from Large-Scale Steady-State Gene Expression Data. In: Kitano, H. (ed), *Foundations of Systems Biology*, pp.59-75.
- Ideker TE, Thorsson V, Karp RM. 2000. World Scientific In: Altman, R B, Dunker A K, Hunter L, Lauderdale K, Klein T E (eds) , *Pac. Symp. Biocomput*, pp 305–316.
- Ideker T, Ozier O, Schwikowski B, Siegel AF (2002). Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics* 18 Suppl 1:S233-240.
- Ambross V (2004). The function of animal micro RNAs. *Nature* 431:350-355.
- Bartel DP (2004). MicroRNAs: genomics, biogenesis, mechanism and function .*Cell* 116: 281-297.
- He L, Hannon GJ (2004). MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet* 5: 522-531.
- Filipowitch W, Bhattacharyya SN, Sonenberg N (2008). Mechanism s of post- transcriptional regulations by microRNAs: are the answer insights? *Nat Rev Genet* 9: 102 -114.
- Haneklaus M, Garlic M, O'Neill LA, Masters SL (2013). miR-223: infection, Inflammation and cancer. *Journal of International Medicine* 274:215-226.
- Eulalio A, Huntzinger E, Izaurralde E (2008). Getting to the root of miRNA- mediated gene silencing. *Cell* 132:9-14.
- Meijer HA, Kong YW, Lu WT, Wilczynska A, Spriggs RV, Robinson SW, Godfrey JD, Willis AE, Bushell M (2013). Translational repression and eIF4A2 activity are critical for microRNA- mediated gene regulation. *Science* 340:82-85.
- Yong K, Jin-Hua H (2005). MicroRNA: Biological and Computational Perspective. *Gen Prot Bioinfo* 3:62-72.
- Aravin A, Tuschl T (2005). Identification and characterization of small RNAs involved in RNA silencing. *FEBS Lett* 579:5830-840.
- Pfeffer S, Zavolan M, Grässer FA, Chien M, Russo JJ, Ju J, John B, Enright AJ, Marks D, Sander C, Tuschl T (2004). Identification of virus-encoded microRNAs. *Science*. 304:734-6.
- Stark A, Brennecke J, Bushati N, Russell RB, Cohen SM (2005). Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution. *Cell* 123:1133-46.