

Secondary Structural Analyses of MicroRNA's and Their Precursors in a Group of Mammals and Insects

Rudra Prasanna Banerjee¹, Sayak Ganguli^{2*}, Moumita Adhikary² and Abhijit Datta³

¹Department of Molecular Biology and Genetics, Presidency University, Kolkata, West Bengal, India.

²DBT Centre for Bioinformatics, Presidency University, Kolkata, West Bengal, India.

³Department of Botany, Presidency University, Kolkata, West Bengal, India.

Short Communication

Received: 28/08/2012

Revised: 18/01/2013

Accepted: 09/04/2013

*For Correspondence:

DBT Centre for Bioinformatics,
Presidency University
Derozio Building (1st Floor)
86/1 College Street
Kolkata 700073, West Bengal, India.
Mobile: +91 9830200174

Keywords: microRNA, Primary RNA, hairpin precursors, stem-loop and correspondence analysis.

ABSTRACT

The recent discovery of microRNAs took many by surprise because of their unorthodox features and widespread functions. These tiny, ~22 nucleotides, RNAs control several pathways including developmental timing, hematopoiesis, organogenesis, apoptosis, cell proliferation and possibly even tumorigenesis. These RNAs are also the key players in the process of RNA interference. This work focuses on the secondary structure analyses of micro RNAs and their precursors and establishment a relation between the formation of secondary structures and their free energy values.

INTRODUCTION

MicroRNAs (mi RNAs) are a class of short (18–25 nucleotides) non-coding RNAs that are involved in the control of gene expression following transcription. Usually the action of these novel class of regulators results in translational repression and gene silencing^[1, 2, 3, 6]. In animal cells, miRNA genes are first transcribed as long pre-mi RNAs and processed to ~70 nt precursors (pre-mi RNA) with stem-loop structure by the RNase III enzyme Drosha which is localized in the nucleus. Following the nuclear processing event the precursors move on to the cytoplasm where another RNase III enzyme Dicer cuts the pre-mi RNAs to release the functional ~22 nt mature mi RNAs. Finally, RNA-induced silencing complexes are formed to regulate the expression of target genes via complementary base pair interactions.

Since, the first mi RNA was discovered in *Caenorhabditis elegans* hundreds of mi RNAs have been cloned in many organisms. However, only abundant mi RNA genes can be easily detected by PCR or northern blot due to limitations of the techniques. For finding those low-expression or tissue-specific mi RNA genes computational prediction provides an efficient strategy. Up to now, several computational approaches have been reported to identify mi RNAs. Comparative genomics was adapted to find entirely novel mi RNA families in specific animals and plants^[4, 7].

MATERIAL AND METHODS

Data mining and Clustering

All animal microRNA sequences available were downloaded from miRBASE and the stem loop (precursor) and mature sequences (products) were kept in two different clusters. The structural analysis were performed using a modified Zuker's algorithm and were then compared to the structures derived using MFOLD. The sequences were then arranged according to the ascending order of least free energies and the total numbers of stem loop structures were calculated for each sequence. The numbers of A-U, G-C &

G-U base pairs per sequences were also calculated and the cutting site of each sequences were also noted. The animal sequences that were used were: *Monodelphis domestica* [5], *Ateles geoffroyi*, *Lagothrix lagotricha*, *Saguinus labiatus*, *Nasoniagiraulti*, *Nasonia longicornis*, and *Nasoniavitripennis*.

Correspondence Analysis (CA)

This is a method which is generally used for the testing of homogeneity in a contingency table and is considered to be closely related to the chi square test. In CA the chi square statistic of the data table is always proportional to the sum of the distances of the points to their centroid (total inertia) which depends on how the points are depicted. CA decomposes the overall chi square statistic. Here a correspondence analysis was performed using the two table rows - the total number of secondary structural motifs present in the precursors and their least free energies to check whether the values were homogeneous.

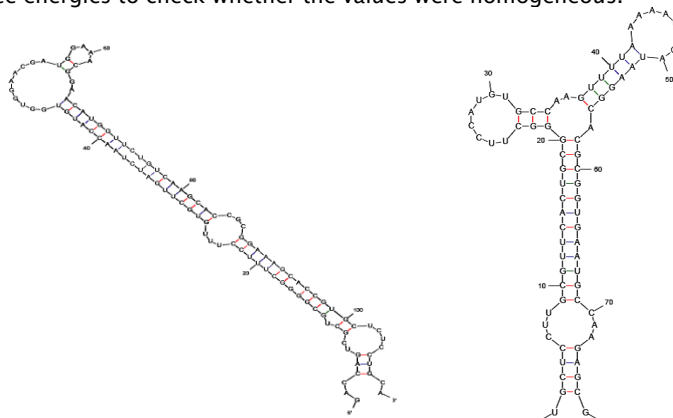


Figure 1: Stem loop structure of animal microRNAs under study.

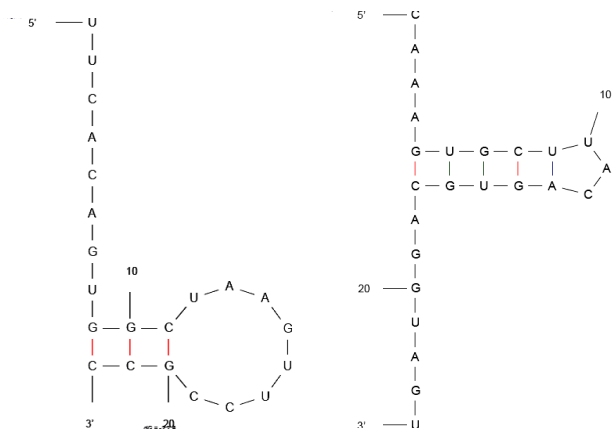
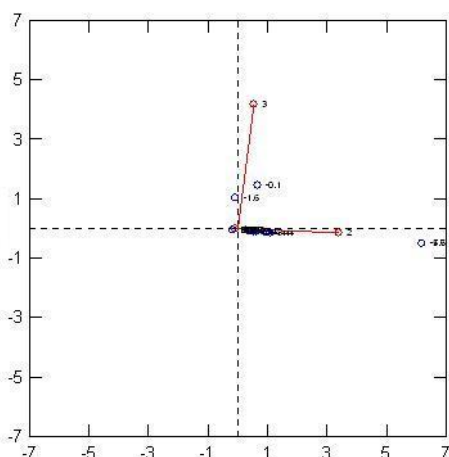


Figure 2: Mature structure of animal microRNAs under study.

Correspondence Plot



Correspondence Plot

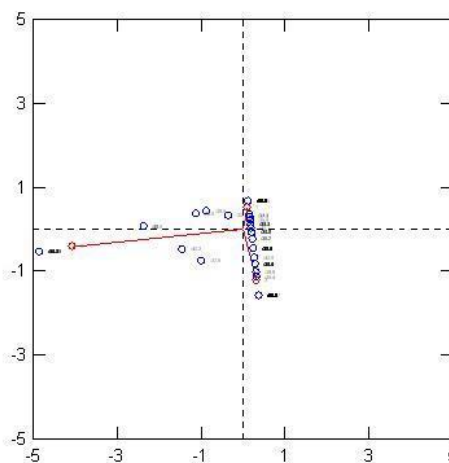


Figure 3: Correspondence plot of the mature and stem loop sequences of animal microRNAs under study.

RESULT AND DISCUSSION

Analysis of RNA secondary structures has shown that most of the animals had one or more stem loop structures in precursor sequences. In case of mature miRNAs almost all studied sequences had single loop and indicated lack of variability and also increased structural similarity due to the near perfect similarity.

The range of free energy of the mature (-11.8 to +2.2) and stem loop sequences (-67.60 to -9.90) were found to match the variously reported secondary structure energies of the miRNA. Sundaralingam (1969) and Leontis and Westhof (2003), have defined RNA motifs as directed and ordered array of non-Watson and Crick base pairs which form clear folding of phosphor-di-ester backbone of interacting strands^[9]. In case of miRNAs secondary structural motifs may contribute to the overall stability of the molecule in the cell as well as determine the nature of interacting partners that the molecules may interact with. As we are still in a nascent stage to completely fathom the the process of RNA folding and its equilibrium parameters, secondary structural analyses provides an alternative strategy to gain novel insights to the overall landscape. In this analysis the relationship of free energies and motif formation were tested using a correspondence analysis based on one of the most common secondary structural motif, the stem loop^[8], and variables are found to be closely homogeneous. In both the plots obtained - one for the hairpin loop structure and the other for mature miRNA sequences, a clustering of the sequences around the centroid is observed which indicates their strong association.

ACKNOWLEDGEMENTS

The authors acknowledge the financial contributions of the Department of Biotechnology Govt. of India under its BTBI scheme.

REFERENCES

1. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE and Mello CC. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*. 1998;391 (6669): 806-811.
2. Ganguli S, Datta A. 2011. RNA Interference—Interactomics and Therapeutics in RNAi Technology, edited by V. K. Gupta, CRC Press, 347-356.
3. Ganguli S, Dhar P, Basu P, Roy P, Datta A. Catalytic RNA world relics in Dicer RNAs. *Intl J Genetics*. 2010; 2(1): 8-17.
4. Hammond SM, Bernstein E, Beach D, Hannon GJ. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature*. 2000;404(6775):293-296.
5. Mikkelsen TS, Wakefield MJ, Aken B, Amemiya CT, Chang JL, et al. Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. *Nature*. 2007; 447(7141):167-177.
6. Kim VN. Small RNAs: classification, biogenesis, and function. *Molecules and Cells*. 2005;19 (1):1-15.
7. Leontis BN, Westhof E. Analysis of RNA motifs. *Curr Opin Srtuct Biol*. 2003;13(1): 300 - 308.
8. Moore PB. Structural motifs in RNA. *Annu Rev Biochem*. 1999;68:287-300.
9. Sundaralingam M. Stereochemistry of nucleic acids and their constituents. IV. Allowed and preferred conformations of nucleosides, nucleoside mono-, di-, tri-, tetraphosphates, nucleic acids and polynucleotides. 1969. *Biopolymers*. 7(6):821 - 860.