**International Journal of Computational Biology**
ISSN: 2229-6700 & E-ISSN:2229-6719, Volume 4, Issue 1, 2013, pp.-52-55.
Available online at http://www.bioinfopublication.org/jouarchive.php?opt=&jouid=BPJ0000220

# IDENTIFICATION OF RNA REGULATORY MOTIFS IN mi-RNA PRECURSORS OF *Pan troglodytes*

**ADHIKARY M., DAS S.G., CHAKRABORTY H.J., MUKHOPADHYAY P., VISHAL V., GUPTA M.K., BERA A.R., BASU P., GANGULI S.* AND DATTA A.**

DBT Centre for Bioinformatics, Presidency University, Kolkata- 700073, WB, India.
*Corresponding Author: Email- sayakbif@yahoo.com

Received: March 26, 2013; Accepted: May 23, 2013

**Abstract-** The last decade has witnessed huge effort in the understanding of the various mechanisms by which small non coding RNAs regulate the various life processes associated with a cell. Such efforts have led to the identification of multiple small non coding RNAs such as miRNA, siRNA, tasiRNA, piRNA, rasiRNA and many more. Of them microRNAs with their ubiquitous presence in all domains of life has become the molecule of choice of many research efforts and mechanisms involving their biogenesis and final function in the cell has been elucidated through these works. However, still certain grey areas exist in our understanding regarding the phenomenon of regulation of these non-coding wonders. As we all are aware that most microRNAs are differentially regulated during the entire lifespan of an organism, it becomes very clear that certain regulatory proteins and their interacting partners play important role in the process. With this background this work was performed where the precursor sequences of microRNAs of *Pan troglodytes* were considered as query and through position specific weight matrix evaluation and subsequent validation a large number of RNA regulatory elements were identified in those sequences.

**Keywords-** microRNA, precursor sequence, position specific weight matrix, RNA regulatory motifs

**Citation:** Adhikary M., et al (2013) Identification of RNA Regulatory Motifs in mi-RNA Precursors of *Pan troglodytes*. International Journal of Computational Biology, ISSN: 2229-6700 & E-ISSN:2229-6719, Volume 4, Issue 1, pp.-52-55.



## Introduction

MicroRNAs are single stranded, endogenous, ranging from 19~25nt in length generated from a long precursors. The precursors itself fold into hairpin structures and repress the post-transcriptional gene expression in both plants and animals. They also induce mRNA degradation and translation repression by binding to their target mRNAs [1-6]. Short single-stranded mi-RNAs are formed in two phases. In nucleus, the miRNA genes are transcribed by RNA polymerase II to generate the primary transcripts (pri-miRNA). The pri-miRNA contains a cap, a polyA tail and the stem loop which together form the hairpin structure. Pri-miRNAs are processed which results in hairpin intermediates which are termed precursor miRNA (pre-miRNA). In the nucleus the RNase III type enzyme Drosha together with its cofactor DGCR8/Pasha form the microprocessor complex. It cleaves the pri-miRNA to form the pre-miRNA. Then the nuclear membrane protein Exportin5 exports the pre-miRNA into the cytoplasm. In the cytoplasm, the pre-miRNA binds with the cytoplasmic RNase III type protein Dicer and cleaved to form short 22nt miRNA with 3'overhangs [7-9]. Following this the microRNA associates with other partners such as Dicer and Argonaute proteins and form the RNA induced silencing complex(RISC) and initiates post transcriptional gene silencing (PTGS). But animal miRNA targets are interrupted by gaps which are nothing but mismatches between the 3'UTR of the target and the effector molecule.

Now a day's miRNAs are recognized as a major regulatory gene family which regulates at least 30% of all mammalian protein coding genes. miRNA has several regulatory motifs which regulate the formation of the secondary structure by playing an essential role in transcriptional and post-transcriptional regulation of gene expression. These regulatory motifs activate under proper condition when the appropriate regulatory factor binds with it [10-14]. So, here we identify the regulatory motifs by using Position Specific Weight Matrix and the result was validated by RegRNA Server. RegRNA is an integrated web server for identifying the homologs of regulatory RNA motifs and elements from the precursor miRNA sequences.

## Material and Methods

### Data Mining and Clustering

The work has been done on the miRNA precursor sequences of *Pan troglodytes*, commonly known as Robust Chimpanzee. They are great apes that are most closely related to human (they shared 98% similarity with human). There are several types of regulatory RNA motifs such as: (a) motifs in mRNA 5' -UTR and 3' -UTR; (b) motifs involved in mRNA splicing; (c) motifs involved in transcriptional regulation; (d) riboswitches; (e) splicing donor/acceptor sites; (f) inverted repeat etc. Here we have identified different types of regulatory motifs which are involved in splicing.

The micro RNA precursor sequenceswere retrieved from the miR-

Base database. Then with the help of position specific weight matrix the regulatory splicing motifswere identified which were placed within the sequence which represents the variation found in the aligned positions. The score values of the matrix gave us a weighted match to any given substring of fixed length. The results were then validated using the RegRNA web server.

Among the exonic & Intronic splicing regulatory motifs there are four types of exonic splicing regulatory motifs which are exonic splicing enhancer motifs, exonic splicing silencer motifs, exon enhancer motifs and exon silencer motifs; the three types of intronic splicing regulatory motifs such as intronic splicing enhancer motifs, intron enhancer motifs and intron silencer motifs. The exonic motifs and intronic motifs were then aligned by R-COFFEE Multiple Sequence Alignment Server and finally the logos were designed using the WEBLOGO3 Server.

The phylogenetic trees were constructed using maximum parsimony of PHYLIP's dnapars/protpars algorithm and the results were validated by using the SeaView-4.3.4 software.

**Result**

Among the exon specific motifs the exon enhancer motifs were present in greater amount then the exon silencer motifs & exon splicing silencer motifs, and it was observed that the exon splicing enhancer motifs were the least abundant. Among the intron specific motifs the intron enhancers were the most frequent motif pattern while the other two i.e. the intron splicing enhancer and the intron silencer were the least abundant [Fig-1] and [Fig-2].

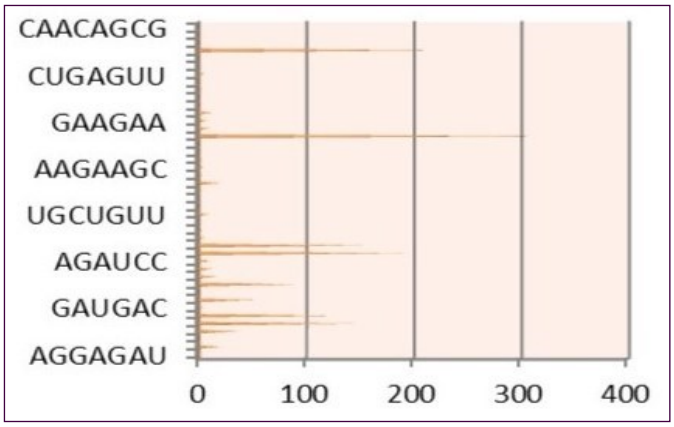


**Fig. 1-** Exonic Motifs Bar Graph

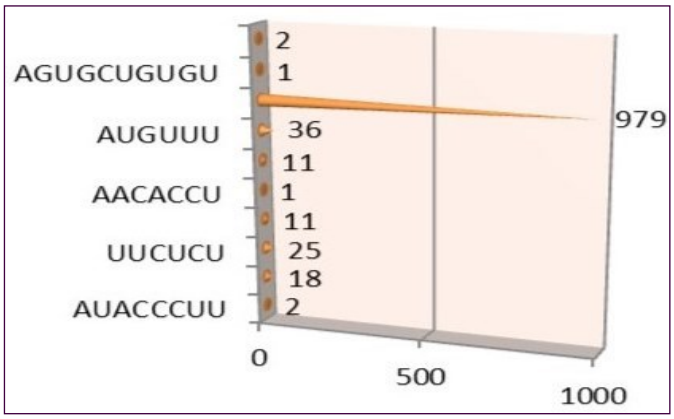


**Fig. 2-** Intronic Motifs Bar Graph

The R-COFFEE results indicate that the alignment between the intronic motifs range from average to good & the alignment score was found to be 57 while the alignment between the exonic motifs were better with a score of 81 [Fig-3a], [Fig-3b], [Fig-4a], [Fig-4b].

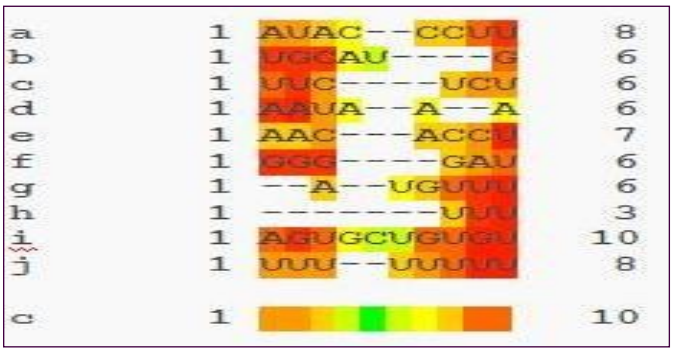


**Fig. 3a-** Multiple Sequence Alignment of Intronic Motifs

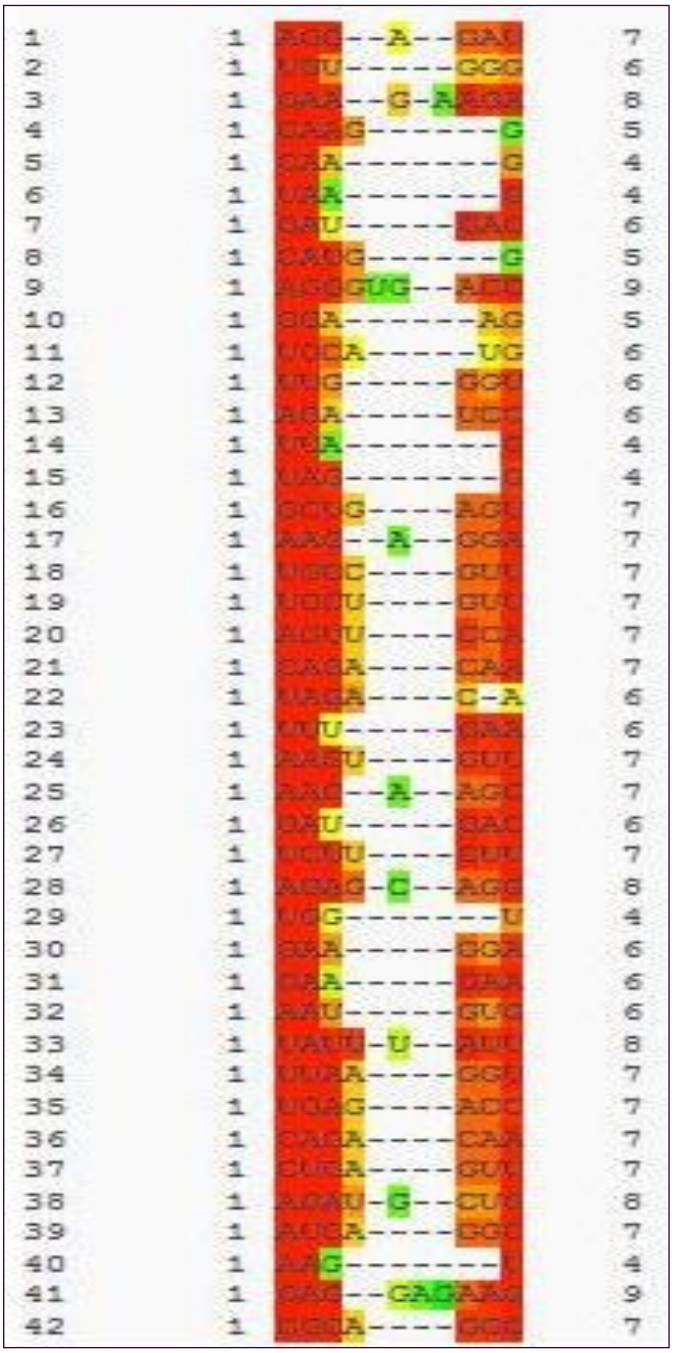

**Fig. 3b-** Multiple Sequence Alignment of Exonic Motifs

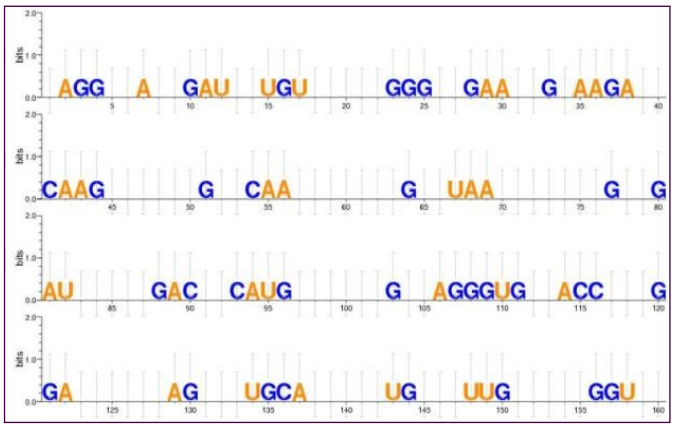

**Fig. 4a-** Sequence Logo of Intronic Motif

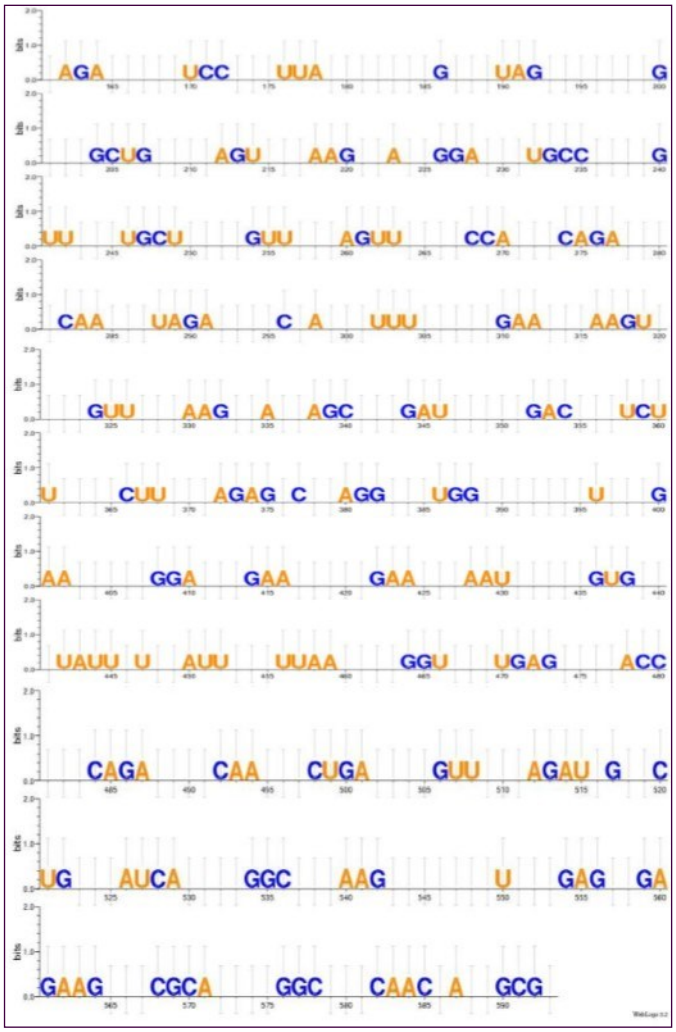

**Fig. 4b-** Sequence Logo of Exonic Motifs

The sequence logos of both exonic & intronic motifs were generated by using the basic Bayesian calculation starting with explicit Dirichlet priors (The data was added and then the posterior mean relative entropy (the stack height) and Bayesian 95% confidence intervals for error bars were calculated [Fig-5a], [Fig-5b].

## Conclusion

A large number of conserved RNA regulatory motifs were identified in the *Pan troglodytes* precursor sequences. This vindicates our hypotheses that the possible modes of microRNA regulation lie in their sequence specific patterns which serve as probable binding sites of activator or repressor proteins involved in the process of cellular regulation. The identification of RNA regulatory motifs is further evidence to the fact that there is a possibility that other regulatory elements are also functional in the process of control of microRNA biogenesis. Future endeavors should focus on the identification of these myriad mechanisms of control of the biogenesis pathway of microRNAs and identify regulators of the regulatory RNAs.

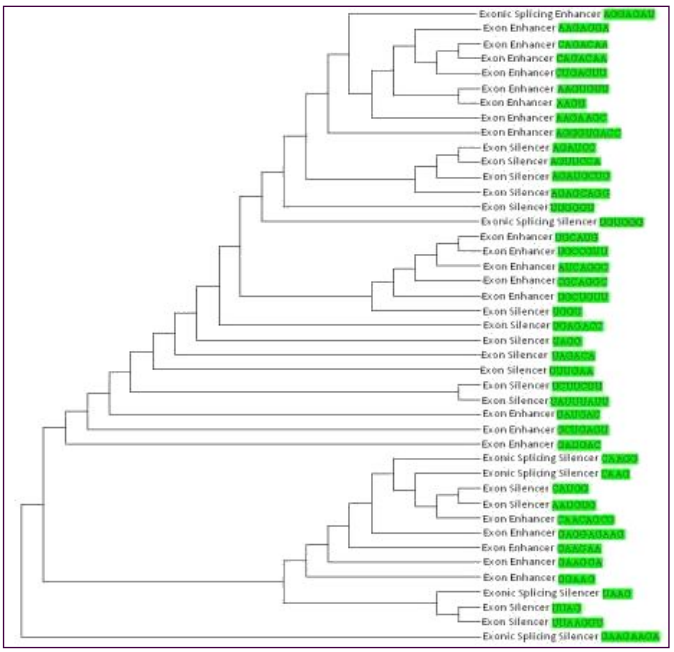

**Fig. 5a-** Phylogenetic Tree of Exonic Motifs

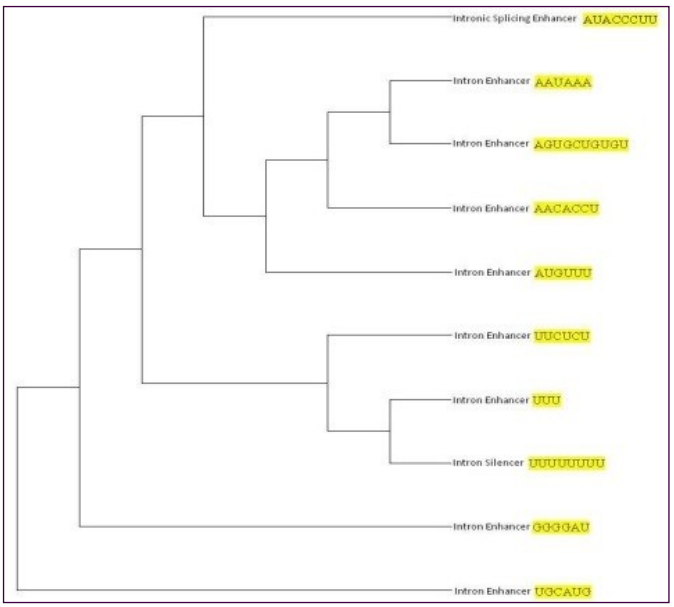


**Fig. 5b-** Phylogenetic Tree of Exonic Motifs

## Acknowledgement

The authors acknowledge the contribution made by the Department of Biotechnology, Ministry of Science of Technology and its "Development of RNAiIDB scheme".

## References

[1] Kim Y. and Kim V.N. (2012) *Molecular Cell*, 46, 384-386.

[2] Adhikary M., Ganguli S., Das G.S., Datta A. (2011) *Int. J. of Comp. Biol.*, 2, 35-37.

[3] Manolo Gouy, Stéphane Guindon and Olivier Gascuel (2010) *Molecular Biology and Evolution,* 27(2), 221-224.

[4] Stephane Guindon, Jean-Francois Dufayard, Vincent Lefort, Maria Anisimova, Wim Hordijk and Olivier Gascuel (2010) *Systematic Biology*, 59(3), 307-321.

[5] Moretti S., Wilm A., Higgins D.G., Xenarios I., Notredame C. (2008) *Nucleic Acids Res*., 36, W10-3.

[6] Wilm A., Higgins D.G., Notredame C. (2008) *Nucleic Acids Res*., 36(9), e52.

[7] Hsi-Yuan Huang, Chia-Hung Chien, Kuan-Hua Jen and Hsien-Da Huang (2006) *Nucleic Acids Res*., 34, 429-434.

[8] Humphreys D.T., Westman B.J., Martin D.I., Preiss T. (2005) *Proc. Natl. Acad. Sci. USA*, 102(47), 16961-16966.

[9] Xiaoman Li and Wing H. Wong (2005) *Proc. Natl. Acad. Sci. USA*, 102, 9481-9486.

[10] Crooks G.E., Hon G., Chandonia J.M., Brenner S.E. (2004) *Genome Research*, 14, 1188-1190.

[11] Eric C. Lai (2002) *Nature Genetics,* 30, 363-364.

[12] Wightman B., Ha I. and Ruvkun G. (1993) *Cell*, 75, 855-862.

[13] Schneider T.D. and Stephens R.M. (1990) *Nucleic Acids Res.*, 18(20), 6097-6100.

[14] Sundaralingam M. (1969) *Biopolymers*, 7, 821-838.