Asian J. Exp. Sci., Vol. 32, No. 1, 2018; 45-50

Online available on : www.ajesjournal.com, ISSSN : 0971-5444

Insilico Identification of MicroRNA (miRNAs) and their Target Prediction from Colorado Tick Fever Virus from Complete Genome.

Vijay Laxmi Saxena*[1], Pragya Chaturvedi[2], Kulsoom Fatima[3]

[1]Bioinformatics infrastructure facility center, Department of Zoology, D.G.P.G. College, Kanpur - 208001 India

[2]School of Biotechnology, KIIT University, Bhubaneswar-751024, India & Bioinformatics infrastructure facility center, Department of Zoology, D.G.P.G. College, Kanpur-208001, India

[3]Department of Biotechnology, C.S.J.M University, Kanpur -212601, India

*Email: vijayashokkanpur@gmail.com

Received: November13, 2017; Revised: December 17, 2017; Accepted: December 22, 2017

Abstract: MicroRNAs (miRNAs) newly identified class of non-protein-coding RNA is ~20nt small RNAs which play an important roles in multiple biological process by de-gradating target mRNAs or repressing mRNA translation. It affects the gene regulation by affecting base pairing of the messenger RNA. Its misregulation may be linked with cancer. MicroRNAs (miRNAs) bind to complementary sequences on the 3' untranslated region (UTR) of mRNAs from hundreds of target genes, leading either to mRNA degradation or suppression of translation. Computational prediction is analyzed and estimation of evolutionary relationship among types of organism is done wih 14 precursors and eight potential miRNAs. miRNAs target 86 target site in 8 gene PEG3, ZIM2, MIDN1, TTN, C16orf57, LRRC27, ABCA1.

Keywords: MicroRNAs, Colorado tick fever virus, Target mRNA, Genes

Introduction

MicroRNAs (miRNAs) are non-coding RNAs that regulate gene expression of post-transcription. It is derived from longer transcripts and encodes in animal, plant and virus genomes (Cooper, *et al*., 2005). Cai. *et al*. (2009) made a brief review on the mechanisms of miRNA regulation. They also reported the intensity of constraint in mammalian genomic sequence. It is recently discovered that miRNAs is also occurring in a single-celled eukaryote. Steven *et al.* (2016) suggested that they regulate and the expression of target genes depend on the binding capacity to complementary sites in their transcripts. According to Kloosterman (2006), *miRNAs* have been implicated in processes and pathways such as development, cell proliferation, apoptosis, metabolism and morphogenesis, and in diseases including cancer. Lewis *et al.* (2003) suggested that the repression of protein synthesis is due to miRNAs in mammals. Pillai, *et al.* (2007) studied the repression of protein synthesis by miRNAs. Stefanie *et al*. (2015) described the molecular understanding of micro RNA and found that it helps in gene silencing. Zhao, *et al* (2007) studied a complex system of small RNAs in the unicellular green alga *Chlamydomon asreinhardtii*. Gregory *et al*. (2004) noticed the occurrence of microprocessor complex which mediates the genesis of microRNAs. Blackshields *et al.* (2010) described the sequence emBedding for fast construction of guide trees for multiple sequence alignment. Chen *et al.* (2008) made detailed procedure of RNA interference targeting VP1 inhibits foot-and-mouth disease by this virus replication in BHK.

Regarding Colorado tick fever virus (CTFV), it infects haemopoietic cells, particularly erythrocytes, and so its study is most important on accounts of incidence of transmission during blood transfusion. It is a double stranded linear genome. At the time the genus was created, it is transferred in tick borne and mosquito borne viruses. Presently, the genus Colti virus contains only CTFV and they are blood-feeding parasites of animals. In nature, the virus is maintained in a cycle involving larval and nymph stages of *Dermacentor andersoni tick* and in variety of rodent species and ultimately transmitted to humans by adult *D. Andersoni* (Gloria, *et al*., 2015). The distribution of CTF virus coincides with the range of the principal vector, *D. andersoni*. Isolates of CTF virus are easily obtained during the late spring and early summer from infected tick, mammals and humans.

Material and methods

The Databases of the present investigation were collected by the following sources:

NCBI: The National Center for Biotechnology Information (NCBI) is part of the United States National Library of Medicine (NLM), a branch of the National Institutes of Health. The NCBI houses a series of databases relevant to biotechnology and biomedicine. Major databases include GenBank for DNA sequences and PubMed, a bibliographic database for the biomedical literature.

GenBank: The NCBI has had responsibility for making available the GenBank DNA sequence database since 1992. GenBank coordinates with individual laboratories and other sequence databases such as those of the European Molecular Biology Laboratory (EMBL) and the DNA Data

Bank of Japan (DDBJ).

The microRNA database: The icroRNA database **(**miRBase) is the central online repository for microRNA (miRNA) nomenclature, sequence data, annotation and target prediction. miRBase provides a range of data to facilitate studies of miRNA genomics all miRNAs are mapped to their genomic coordinates. Clusters of miRNA sequences in the genome are highlighted and can be defined and retrieved with any inter miRNA distance. The overlap of miRNA sequences with annotated transcripts, both protein and non-coding are described.

Server:

DIANA Lab. Diana microT 3.0 is an algorithm for the prediction of miRNA target prediction which is based on several parameters genes (miTGs).

Tools:

Oligo pro calculator: OligoCalc as a web-accessible, client-based computational engine for reporting DNA and RNA single-stranded and double stranded properties, including molecular weight, solution concentration, melting temperature, estimated absorbance coefficients, inter-molecular self-complementarity estimation and intramolecular hairpin loop formation. OligoCalc has a familiar 'calculator' look and feel, making it readily understandable and usable.

Clustal omega: Multiple sequence alignments (MSAs) are essential in most bioinformatics analyses that involve comparing homologous sequence.

Mireval: Mireval is a comprehensive tool and is easy to use. It will allow users with no prior knowledge of in-silico detection of microRNAs to take advantage of the most successful approaches to investigate sequences of interest.

RNAfold: The RNAfold web server will predict secondary structures of single stranded RNA or DNA sequences. Current limits are 7,500 nt for partition function calculations and 10,000 nt for minimum free energy only prediction.

Methodology

The dataset is downloaded from NCBI and predict precursor from miRBase database of. Colorado Tick Fever Virus (CTFS) complete whole genome (12segments). Mireval were used to find out the precursor sequence. The secondary structure with optimal minimum free energy was analyzed through RNA fold web server and validated. After finding conserved region for miRNA, targets were found out through DIANA lab.

Results and discussion

Precursor analysis

The study is exclusively based on in silico firstly to find out the precursors from Colorado Tick Fever Virus having 1 to 14 precursors in the whole genome and then this sequence is submitted in MiRevel software for precursor prediction.

Potential miRNA analysis

Prediction of potential miRNA was based on (MSA) multiple sequence alignment (pair wise) and its alignment were completed with the help of EBI clustal- omega and MFOLD miRNA predicting software. The sequence was submitted in EBI which is retrieved from miRBases; and multiple sequence alignment was done with precursors of viruses and precursors of Colorado tick fever virus. Mainly, MSA focused on conserved regions for miRNA (Table 1-4).

Physical content and thermodynamic condition

Physical content and thermodynamic condition of CTFV we found with the help of Oligo pro calculator.

Target prediction analysis

Based on observation, potential miRNA predicted the different types of target sites in a human gene. Then it predicted the different type of genes and those involved in different type of target sites with conserved information

All the tables explain the different types of analyses related to Insilco prediction only, and the whole information of targeted genes in humans and targeted miRNA of Colorado Tick Fever Virus. The present results are in confirmation to the earlier investigators

Conclusion

The study of Insilco identification miRNA and their target prediction from Colorado Tick Fever Virus, 23 potential miRNAs were predicted Insilco and 86 target sites were found in 8 genes. Out of these eight genes, PEG3, ZIM2, LRRC and c16orf57 played an important role in the regulation of the central nervous system and its activity in human brain and body; it is shown that if it is damaged it could lead aseptic meningitis, encephalitis, light sensitivity and skin abnormalities like Rashes. So, if we can control the malfunction protein at the transcription level, we would be able to save patients suffering from attack of *Colorado Tick Fever Virus.* And hence these genes can be used in further studies to control Tick Fever

Acknowledgement

I like to put my sincere acknowledgements to DBT for providing us such platform. And my sincere thanks to BIF Center at D.G.P.G. College, Kanpur and all staff members there.

Table-1. Predicted precursor from miRNAmiRBase DATABASE

S. no.	Annotated pre. Name	Pre-miRNA	Genome in position	Length (nt)
1.	>Pre_1_miR_CTFV (segment 1)	No conserved region		
2.	>pre_2_miR_CTFV (segment 2)	CUUCUGGUCUCAUCGUCUCGGAC AGGGAGUAUGGAGGCAAUUUCCA GCUGUGCGAGAGGAUCGUGAGCG GCGGUUUGAUAACCAC	2331-2415	86
3.	>Pre_3_miR_CTFV (segment 3)	No conserved region		
4.	>Pre_4_miR_CTFV (segment 4)	No conserved region		
5. 6.	>pre_5_miR_CTFV (segment 5)$_A$ >pre_5_miR_CTFV (segment 5)$_B$	GCCGUUAUCUUGGAGAAGAUGGC AACAUCUCUUGCAUUGGCGGCGG CGAAGGAGGGUAACCCGGCAGAG GCCUUCUAUCGUGACA ACGACUUUGGAUCUGAUGUUUCA CCAGCUGUUCCUAUUAUCAACGG GUGAUGUUUGGCUUGGAGCGUU GGUGAUUCAUGAUCCCC	1791-1875 2121-2205	86 86
7.	>pre_6_miR_CTFV (segment 6)	No conserved region		
8.	>pre_7_miR_CTFV (segment 7	GUGCAUUGGCAUCCACUCAUGUU GCCUCUCUAGACAACGUUGGAUG AAAAUGGUGCUGAUUUGGCAAA AUUCGAGCAGCGCGAGC	541-625	86
9. 10. 11.	>pre_8_miR_CTFV (segment 8)$_A$ >pre_8_miR_CTFV (segment 8)$_B$ >pre_8_miR_CTFV (segment 8)$_C$	UUGAUCAAGCGGCUUUAGCGGCU CAAGCUGUUGCUGAUGCAGGAGG UGGGGUUGCACAACAACAGGCGG CUGCAGCGGCGGCAGG GGCUUUAGCGGCUCAAGCUGUUG CUGAUGCAGGAGGUGGGGUUGCA CAACAACAGGCGGCUGCAGCGGC GGCAGGUGUGCAGGCU GCUCAAGCUGUUGCUGAUGCAGG AGGUGGGGUUGCACAACAACAGG CGGCUGCAGCGGCGGCAGGUGUG CAGGCUCAAGAGGAUU	531-615 541-625 551-635	86 86 86
12. 13.	>pre_9_miR_CTFV (segment 9)$_A$ >pre_9_miR_CTFV (segment 9)$_B$	UGGCCUGGUGAGCUGUACCUCAU UAACAUCGAGAGUGCAGGCUUUC GUGAAGGAGCUUCAGCGGCUGCG GCUUACCAGGCGGGCU GCGUGCGCGACGCAUCUCCGUUA ACUGAGAGCGGGGUUGUUUAUCA GGUGGAGGCUGCUCAGUACGGAA UGCGAACGACGCGAGU	1061-1145 -1665	86 86
14.	>pre_{10}_miR_CTFV (segment 10)	ACGUCCCUGACGGACCGCAUCGU CAAGGAGGCAUUGGCUACAGUGA AAGUGUUCAGUUUCGUCGUUGGG CGAGAUCACUUGCCUG	1121-1205	86
15. 16.	>pre_{11}_miR_CTFV (segment 11)$_A$ >pre_{11}_miR_CTFV (segment 11)$_B$	GGAGCCGUGCCGCAUCUUCAAUU UGUUAACAAUGAGUUGAAGGUC GGGUGGAAGGUCGCGCCGAUACG UCGUGAUGGGCGAAACU	801-885	86

Table-2. Structure and MFE with the help of RNA fold web server

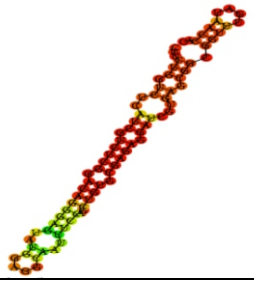
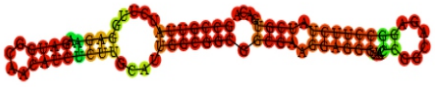
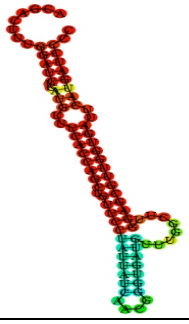
Annotated precursor name	Structure	MFE (kcal/mol)
>Pre_1_miR_CTFV (segment 1)	No conserved region	
>pre_2_miR_CTFV (segment 2)		-29.00
>Pre_3_miR_CTFV (segment 3)	No conserved region	
>Pre_4_miR_CTFV (segment 4)	No conserved region	
>pre_5_miR_CTFV (segment 5)A		-31.00
>pre_5_miR_CTFV (segment 5)B		-29.30
>pre_6_miR_CTFV (segment 6)	No conserved region	

Table-3 For physical constant and thermodynamics constant condition with the help of OLIGO PRO CALCULATOR for CTFV

S.NO	Annotated Name	Physical Constants			Thermodynamic constants conditions		
		Length	% GC	M.W.	Delta G	Delta H	Delta S
1	>miR$_{2}$_CTFV(A)	24	54	7158.5	33	199.8	521.5
2	>miR$_{2}$_CTFV(B)	21	67	6550.2	31.9	185.5	478.9
3	>miR$_{5A}$_CTFV	24	50	7348.7	32.1	196.3	513.2
4	>miR$_{5B}$_CTFV(A)	20	40	6025.8	23.9	159.1	419.8
5	>miR$_{5B}$_CTFV(B)	20	40	6025.8	23.9	159.1	419.8
6	>miR$_{7}$_CTFV(A)	23	48	6893.4	29.8	185.4	485.9
7	>miR$_{7}$_CTFV(B)	24	46	7332.7	32.9	197.6	514.7
8	>miR$_{8A}$_CTFV(A)	23	52	6972.4	33	192	496.2
9	>miR$_{8A}$_CTFV(B)	24	75	7439.8	41.1	218.6	556
10	>miR$_{8B}$_CTFV(A)	19	58	5749.7	26.4	156.2	402
11	>miR$_{8B}$_CTFV(B)	24	75	7433.8	40.3	216.4	551.6
12	>miR$_{8C}$_CTFV	24	54	7324.7	33.3	192.5	497
13	>miR$_{9A}$_CTFV(A)	17	41	5059.3	18.3	131.9	350.3
14	>miR$_{9A}$_CTFV(B)	23	70	6985.5	36.7	203.5	521.5
15	>miR$_{9B}$_CTFV(A)	24	63	7259.6	38.4	217.7	561.8
16	>miR$_{9B}$_CTFV(B)	24	63	7259.6	38.4	217.7	561.8
17	>miR$_{10}$_CTFV(A)	23	70	6905.4	36.2	211.5	549.1
18	>miR$_{10}$_CTFV(B)	20	60	6078.9	28.9	176.5	459.7
19	>miR$_{11A}$_CTFV(A)	24	54	7221.6	34.5	204.6	532.2
20	>miR$_{11A}$_CTFV(B)	22	68	6736.3	35	201.5	520.8
21	>miR$_{11B}$_CTFV(A)	20	35	5929.8	22.6	155.7	413
22	>miR$_{11B}$_CTFV(B)	25	60	7675.9	38.9	226.3	588.1
23	>miR$_{11C}$_CTFV	22	50	6689.3	28.6	181.2	476
24	>miR$_{12C}$_CTFV(A)	17	59	5090.3	22.3	143.4	374.4
25	>miR$_{12C}$_CTFV(B)	25	60	7549.8	36.3	215.5	561.7

Table-4 Annotated name of miRNA and sequence of potential miRNA and their genes and number of target. (19 Threshold energy on Diana target)

S.No	Annotated miRNA name	miRNA nucleotide sequence	Name and no of Gene	Target site
1	>miR$_{2}$_CTFV1	CUUCUGGUCUCAUCGUCUCGGACA	5 genes (PEG_3,ZIM_2) ($MIDN_1$) (TTN)	44
2	>miR$_{2}$_CTFV2	GAGAGGAUCGUGAGCGGCGGU	2 gene (C_16orf_{57}) ($LRRC_{27}$)	37
3	>miR$_{7}$_CTFV1	GUGCAUUGGCAUCCACUCAUGUU	1 gene ($ABCA_1$)	5

References

Amarzguioui M., Rossi J. J., Kim D. (2005): Approaches for chemically synthesized siRNA and vector-mediated RNAi. FEBS Lett. 579:5974-5981.

Blackshields G., Sievers F., Shi W., Wilm A., Higgins D.G. (2010): Sequence emBedding for fast construction of guide trees for multiple sequence alignment. Algorithms MolBiol 5: 21.

Cai Y., Yu X., Hu S., Yu J. (2009): A brief review on the mechanisms of miRNA regulation. Genomics Proteomics Bioinformatics; 7(4):147-54..

Chen W., Yan W., Du Q., Fei L., Liu M. (2008): RNA interference targeting VP1 inhibits foot-and-mouth disease virus replication in BHK-21 cells.

Cooper, G.M. et al. (2005): Distribution and intensity of constraint in mammalian genomic sequence. Genome Res., 15, 901–913.

Elbashir S.M., Harborth J., Lendeckel W., Yalcin A., Weber K. (2001): Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nat. 411:494-498.

Emmons R.W. (1966): Colorado tick fever: Prolonged virememia in hibernating Citelluslateralis. American journal of Tropical Medicine and Hygiene; 15, 428-433.

Garzon R., Fabbri M., Cimmino A., Calin G.A. and Croce C.M. (2006): MicroRNA expression and function in cancer. Trends Mol. Med.; 12, 580–587.

Gloria Bertoli, Claudia Cava, Isabella Castiglioni (2015): MicroRNAs: New Biomarkers for Diagnosis, Prognosis, Therapy Prediction and Therapeutic Tools for Breast Cancer *Theranostics*; 5(10):1122-1143. doi:10.7150/thno.1154.

Gregory R.I., Yan K.P., Amuthan G., Chendrimada T., Doratotaj B., Cooch N., Shiekhattar R., (2004): The Microprocessor complex mediates the genesis of microRNAs. Nat. 432 (7014):235-240.

Kloosterman,W.P. and Plasterk, R.H. (2006): The diverse functions of microRNAs in animal development and disease. Dev. Cell, 11, 441–450.

Kirby C. Stafford. (2007): Tick Management Handbook, The Connecticut General Assembly, revised edition.

Lewis B.P., Shih I.H., Jones-Rhoades M.W., Bartel D.P., Burge C.B. (2013): Prediction of mammalian microRNA targets. Cell, 115(7):787-798.

Pillai, R.S., Bhattacharya, S.N. and Filipowicz, W. (2007): Protein Synthesis by Mirans: How many mechanisms? Trends Cell Biol 17:118-126.

Stefanie Jonas and Elisa Izaurralde. (2015): Towards a molecular understanding of microRNA-mediated gene silencing, Nature Reviews Genetics 16, 421–433

Steven Maltby, Maximilian Plank, Hock L. Tay, Adam Collison and Paul S. Foster. (2016): Targeting MicroRNA Function in Respiratory Diseases: Mini-Review, Front. Physiol.

Zhao, T., Li, G., Mi, S., Li, S., Hannon, G.J., Wang, X.J. and Qi,Y. (2007): A complex system of small RNAs in the unicellular green alga *Chlamydomonas reinhardti*i. Genes Dev., 21, 1190–1203.