International Journal of Bioinformatics and Biological Science: v.1 n.1 p.9-17. March, 2013

Prediction of miRNA targets affected proteins and their homologs in *Mouse gammaherpesvirus68*

Budhayash Gautam[1]*, Pramod Katara[2], Anshu Choudhary[1], Gulshan Wadhwa[3] and Satendra Singh[1]

[1]Department of Computational Biology & Bioinformatics, Jacob School of Biotechnology and Bioengineering, Sam Higginbottom Institute of Agriculture, Technology and Sciences-Deemed to be University, Allahabad-211007, India
[2]Department of Bioscience and Biotechnology, Banasthali University, P.O. Banasthali Vidyapith, India-304022
[3]Apex Bioinformatics Centre, Department of Biotechnology, Ministry of Science and Technology, CGO Complex, Lodhi Road, New Delhi – 110 003

*Corresponding Author: Budhayash Gautam budhayash.gautam@shiats.edu.in

Received: 24 December 2012; Accepted: 03 January 2013

ABSTRACT

microRNAs are small noncoding, single-stranded RNA gene product about 20-24nt long that are processed by Dicer from precursor with a characteristic hairpin secondary structure. Precise regulation of miRNAs activity during various stages of growth and in specific cell types is of central importance for normal development because miRNAs affect morphology of plants and animals by regulation of the gene expression at the post-transcriptional level which is involved in critical developmental events. Thus, in the present study the focus is on the animal miRNAs and prediction of the miRNA target, affected proteins by miRNA and miRNA homolog of *Mousegammaherpes virus68.* Present analyses are based on sequence complimentaries between miRNA and mRNAs. As a result, we predicted 98 targets for 49 mature miRNA sequences and among these 58 mature miRNA sequences were already published in database. The study of affected proteins revealed that very less number of miRNAs, protein products are known and they mostly involved in diverse processes like elements of signal recognition. Homology analyses for miRNAs suggested that 17 miRNAs of *Mousegammaherpes virus68* show 379 miRNA homologs for different animal species.

Keywords: miRNA targets, in silico identification, Mousegammaherpes virus68, homologs, EST.

BACKGROUND

MicroRNAs (miRNAs) are small ~20-24 nucleotide endogenous and noncoding regulatory RNAs that are processed from hairpin RNA precursors encoded within the genome of metazoans. MicroRNA (miRNA) genes are a new and large class of genes that do not encode proteins. They produce transcripts that in many cases are

thought to function as antisense regulators of other mRNAs. In both animals and plants, the majority of the miRNA genes exist as independent transcriptional units, and they are transcribed by RNA polymerase II into long primary transcripts (termed pri-miRNAs)(Sunil chandra *et al* ,1994).

The first miRNA was discovered in *Caenorhabditis elegans* hundreds of miRNAs also have been cloned in various plant species(Hochreiter *et al* ,2007,Hwang *et al* ,2008,Katara *et al* ,2010)but the progress in prediction of the genes responsible for these miRNAs production, and cleavage target/sites of these miRNAs are still at very low rate which need to speed up to understand the fundamentals of gene regulation system at miRNA level, more accurately and completely(sunil chandra *et al* ,1994)

In nature, miRNA acts as post-transcriptional regulators in animals and plants. miRNAs use two distinct post-transcriptional mechanisms to regulate gene expression. They act by binding to the complementary sites on the 3' untranslated region (UTR) of the target gene to induce cleavage with near perfect complementarily(Hochreiter *et al* ,2007) or to repress productive translation; they also facilitate deadenylation, which leads to rapid mRNA decay. In most known cases in animals, miRNAs regulate the translational expression of genes. To understand the biological function of miRNAs, it is necessary to identify their targets.

Computational approaches have been successful in plants, where known target sites tend to be almost perfectly complementary to miRNAs and where miRNAs are thought to promote degradation of the target mRNA. In animals, however, the miRNA/mRNA base pairing appears to be less than perfect, which has greatly hindered computational approaches for target site identification.

Several computational methods are developed and have been widely used by scientist for the prediction of miRNAs(Katara *et al* ,2010 and Lok *et al* ,2002) and their target genes using different approaches(Nash *et al* ,2001) Several approaches have been used for miRNA target predictions in plants and animals. For plant as well as animal sequences, similarity based simple pattern matching approaches or BLAST similarity searches have shown high performance, because complementarity in miRNA and their targets is nearly perfect(Lok *et al* ,2002 and Hwang *et al* ,2008)

Mouse gammaherpes virus (MHV-68) is genetically related to the *Human gammaherpes viruses, Epstein–Barr virus* (EBV) and Kaposi's sarcoma-associated herpesvirus (KSHV)**.** Its long term infection causes the development of lymphomas and several immunological changes occur. How this virus gains access to the CNS and how frequently this occurs remain unknown(Nguyen *et al* ,2008,Sarid *et al* , 1999,Stewart *et al* 1995, Svobodova *et al* ,1982 and Virgin *et al* ,1997).

Due to all above reasons it is now very much required that some kind of treatment should be available for this problem. For this, miRNA prediction is one of the important solutions(Katara *et al* 2010 and Lok *et al* ,2002). Thus in the present

work, identification of miRNA targets of *Mouse gammaherpes virus 68* was done. Along with that identification of affected proteins and their functions was also done. miRNA homologs which are present in other species were also identified.

MATERIALS AND METHODS

Retrieval of miRNA sequences from miRBase database

MiRBase database was used for the sequence retrieval (www.miRBase.org). It contains 8619 miRNA loci from 87 species including *Homo sapiens*, flies, plants, animals and viruses. Among 87 species 21 are plant species,and rest are animal species including flies, viruses etc. Present work focuses on *Mouse gammaherpesvirus 68*. All available sequences of mature sequences and stem loop sequences miRNAs of Mouse gammaherpesvirus 68 were downloaded from miRBase database.

Identification of miRNA targets

To find the targets for miRNAs, all the sequences were compared with their EST database dbEST (http://bioedit.ncbi.nim.nih.gov/dbEST) available at NCBI

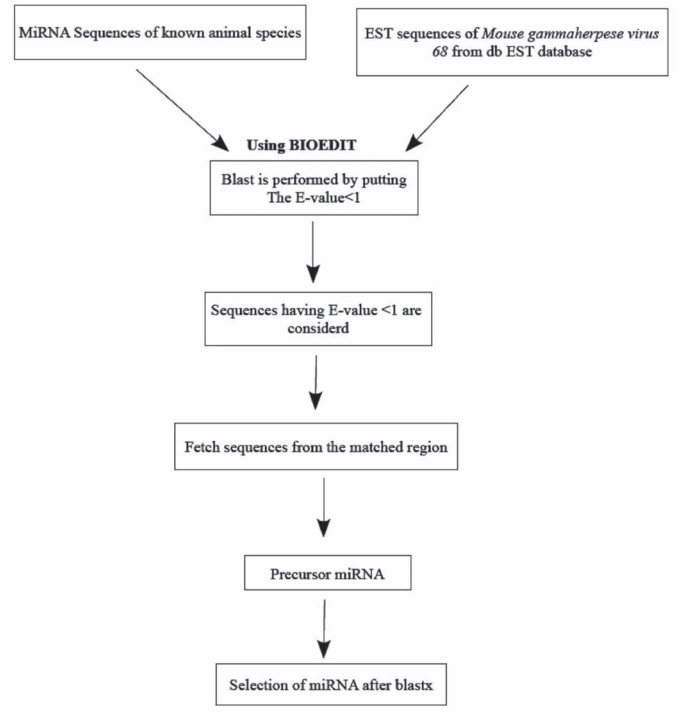


Figure 1: Flow chart of miRNA prediction in *Mousegammaherpes virus 68*

(National Centre for Biotechnology Information) with the help of BIOEDIT tool(Wheeler *et al,*2004) Corresponding mRNA of EST, with high similarities (E-value less than 1) had been considered as potential miRNA targets(Katara *et al* ,2010) A complete flow chart of the whole process is given in Figure-1.

Identification of affected proteins

These targets encode specific protein whose translation is being suppressed due to binding of miRNA affected proteins, with their functions. Target sequence (nucleotide sequence) were converted into protein query sequence. For this purpose blastx was used, which uses nucleotide sequences as queries and translate them in all six reading frames to produce translated protein sequences. These translated protein sequences were further used as protein query for protein sequence Database. Further, to get more specified results, target sequences with significant statistics values had been selected, and blastx was performed for these selected target sequences against Swiss Prot database (highly annotated protein database). Hits with highest bit score and lowest E-value had been considered as effected proteins by miRNAs of *Mouse gammaherpes virus 68*. Below, a flowchart about methodology is given in Figure1.

Identification of miRNA homologs

To identify miRNA homologs, sequences of mature miRNAs (which are conserved since their sequences are crucial for target-interaction) were taken as input for sequence similarity search against mature miRNA sequences of all other viruses available in miRBase Database. MiRNAs of all other viruses species with bit score more than 100 against query sequence had been selected as homolog miRNAs(Kataara *et al* ,2010).

RESULTS AND DISCUSSION

Identification Of miRNA Targets

Previous studies suggest that miRNAs bind to their mRNA target with nearly perfect sequence complementary, and degrade the target mRNA. On the basis of this concept sequence similarity search between query sequence of miRNA and EST database were used and got 98 targets for 49 mature miRNA sequences among 58 mature miRNA sequence published in database (Table 1). These targets were also showing significant sequence similarities with E value ranging from (1e-49) to 0.99Among 98 miRNA targets of *Mouse gammaherpesvirus 68,* 61% of targets were showing E-value around 0.99 and 17% were showing E-value around 0.71, thus most of the miRNA with their respective target sequence shared very high similarities which are significant and provide accuracy. High number of predicted targets (mRNAs) for the miRNA candidates of *Mouse gammaherpesvirus 68* were observed, as compared to other well studied animal species miRNA in *Mus musculus*, revealed that a broad range of genes were regulated by miRNA in *Mouse gammaherpesvirus 68*(Katara,*et al*,2010).

Table 1: Resultant mature and stem loop sequences

Mature Sequence (query numbers)	Stem Loop Sequence (query numbers)
i. Query= mghv-miR-M1-1 MIMAT0001564	i. Query= mghv-mir-M1-1 MI0001669
ii. Query= mghv-miR-M1-11-3p MIMAT0018169	ii. Query= mghv-mir-M1-10 MI0016250
iii. Query= mghv-miR-M1-12 MIMAT0018171	iii. Query= mghv-mir-M1-11 MI0016251
iv. Query= mghv-miR-M1-8 MIMAT0001572	iv. Query= mghv-mir-M1-13 MI0016253
v. Query= mghv-miR-M1-10* MIMAT0018166	v. Query= mghv-mir-M1-14 MI0016254
vi. Query= mghv-miR-M1-4* MIMAT0017188	vi. Query= mghv-mir-M1-15 MI0016255
vii. Query= mghv-miR-M1-13-3p MIMAT0018173	vii. Query= mghv-mir-M1-2 MI0001670
viii. Query= mghv-miR-M1-9 MIMAT0001573	viii. Query= mghv-mir-M1-3 MI0001671
ix. Query= mghv-miR-M1-8* MIMAT0017191	ix. Query= mghv-mir-M1-5 MI0001673
x. Query= mghv-miR-M1-7-5p MIMAT0001570	x. Query= mghv-mir-M1-6 MI0001674
xi. Query= mghv-miR-M1-6 MIMAT0001569	xi. Query= mghv-mir-M1-7 MI0001675
xii. Query= mghv-miR-M1-14-5p MIMAT0018174	xii. Query= mghv-mir-M1-8 MI0001676
xiii. Query= mghv-miR-M1-11-5p MIMAT0018168	xiii. Query= mghv-mir-M1-9 MI0001677
xiv. Query= mghv-miR-M1-7-3p MIMAT0001571	
xv. Query= mghv-miR-M1-5* MIMAT0017189	
xvi. Query= mghv-miR-M1-3 MIMAT0001566	
xvii. Query= mghv-miR-M1-2-3p MIMAT0001565	
xviii. Query= mghv-miR-M1-13-5p MIMAT0018172	

Identification of affected proteins

In *Mouse gammaherpes68,* 80 targets sequences are available using suitable statistical parameter i.e concerning bit score (40 and above) an E-value9Katara, *et al* ,2010).All of the targets encodes for signal recognition particle, kynurenine, perilipin, ribosomebiogenesis, thioredoxin, matrix metalloproteinases, homeobox protein, chordin, cadherins,biotin synthase, trna-pseudouridine synthase proteins (Table 2).

Identification of miRNA homologs

In Mouse gammaherpesvirus68 17 miRNA targetswere having homologs within species like *Homo sapiens, Mus musculus, Bos Taurus, Gallus gallus, Pongo*

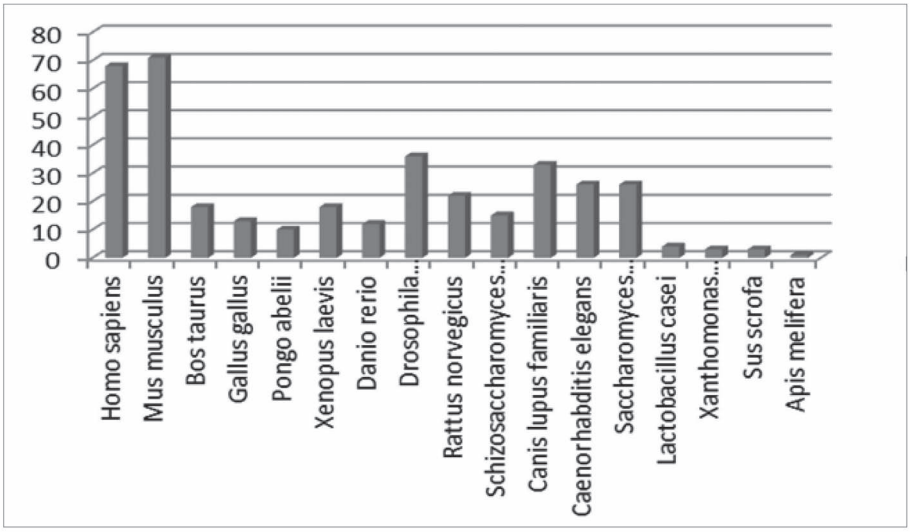
Table 2: Proteins with their specific functions

Proteins	Functions
The Signal Recognition Particle	The signal recognition particle (SRP) is an abundant, cytosolic, universally conserved ribonucleoprotein (protein-RNA complex) that recognizes and targets specific proteins to the endoplasmic reticulum in eukaryotes and the plasma membrane in prokaryotes.
Kynurenine	Kynureninase catabolizes the conversion of kynurenine into anthranilic acid while kynurenine-oxoglutarate transaminase catabolizes its conversion into kynurenic acid. Kynurenine 3-hydroxylase converts kynurenine to 3-hydroxykynurenine.
Ribosome Biogenesis	Ribosome biogenesis is the process of making ribosomes. In prokaryotic cells, it takes place in the cytoplasm with the transcription of many ribosome gene operons. In eukaryotes, it takes place both in the cell cytoplasm and in the nucleolus of eukaryotic cells. It involves the coordinated function of over 200 proteins in the synthesis and processing of the four rRNAs, as well as assembly of those rRNAs with the ribosomal proteins.This organelle also involve in the protein synthesis during translation process within the cytoplasm of a cell.
Perilipin	Perilipin is a protein that coats lipid droplets in adipocytes,the fat-storing cells in adipose tissue. Perilipin acts as a protective coating from the body's natural lipases, such as hormone-sensitive lipase, which break triglycerides into glycerol and free fatty acids for use in metabolism, a process called lipolysis. In humans, perilipin is expressed in three different isoforms, A, B, and C, and perilipin A is the most abundant protein associated with the adipocyte lipid droplets.
Thioredoxin	Thioredoxins are proteins that act as antioxidants by facilitating the reduction of other proteins by cysteine thiol-disulfide exchange. Thioredoxins are found in nearly all known organisms and are essential for life in mammals. Thioredoxins are characterized at the level of their amino acid sequence by the presence of two vicinal cysteines in a CXXC motif. These two cysteines are the key to the ability of thioredoxin to reduce other proteins. Thioredoxin proteins also have a characteristic tertiary structure termed the thioredoxin fold.
Matrix Metalloproteinases	Matrix metalloproteinases (MMPs) are zinc-dependent endopeptidases; other family members are adamalysins, serralysins, and astacins. The MMPs belong to a larger family of proteases known as the metzincin superfamily. They were first described in vertebrates, including humans, but have since been found in invertebrates and plants. They

Contd.

Proteins	Functions
	are distinguished from other endopeptidases by their dependence on metal ions as cofactors, their ability to degrade extracellular matrix, and their specific evolutionary DNA sequence
Homeobox Proteins	In molecular biology, the engrailed homeobox proteins are a family of homeobox proteins which are characterised by the presence of a region of some 20 amino-acid residues located at the C-terminal of the 'homeobox' domain. This region forms a signature pattern for this subfamily of proteins.
Chordin	Chordin is a polypeptide that dorsalizes the developing embryo by binding ventralizing TGFâ proteins such as bone morphogenetic proteins. It may also play a role in organogenesis. There are five named isoforms of this protein that are produced by alternative splicing.In humans, the chordin peptide is encoded by the CHRD gene
Cadherins	Cadherins (named for "calcium-dependent adhesion") are a class of type-1 transmembrane proteins. They play important roles in cell adhesion, ensuring that cells within tissues are bound together. They are dependent on calcium (Ca^{2+}) ions to function, hence their name.
Biotin Synthase	This enzyme belongs to the family of transferases, specifically the sulfurtransferases, which transfer sulfur-containing groups. The systematic name of this enzyme class is dethiobiotin:sulfur sulfurtransferase. This enzyme participates in biotin metabolism. It employs one cofactor, iron-sulphur
trna-pseudouridine Synthase	This enzyme belongs to the family of isomerases, specifically those intramolecular transferases transferring other groups. The systematic name of this enzyme class is tRNA-uridine uracilmutase. Other names in common use include tRNA-uridine isomerase, tRNA pseudouridylate synthase I, transfer ribonucleate pseudouridine synthetase, pseudouridine synthase, and transfer RNA pseudouridine synthetase

abelii, Xenopus laevis, Danio rerio, Drosophila melanogaster, Rattus norvegicus, Schizosaccharomyces pombe, Canis lupus familiaris, Caenorhabditis elegans, S. cerevisiae, Lactobacillus casei, Xanthomona campestris, Sus scrofa, Apis



melifera. Total number of miRNA homologs for *Mouse gammaherpesvirus68* was 98 and 71 miRNA sequences were showing homology with miRNA of species *Mus musculus.* Therefore miRNAs of *Mouse gammaherpesvirus68* were more conserved with in *Mus musculus.* miRNAs of *Homo sapiens, Drosophila melanogaster, Canis lupus familiaris, Caenorhabditis elegans* were also showing homology with *Mouse gammaherpesvirus68* (Figure 2). Analysis suggests that miRNA are evolutionary conserved and shared similarities with different animal species, over wide evolutionary distances(Katara,*et al* ,2010) Here it is notable that *Mouse gammaherpesvirus68* share maximum similarities with species *Mus musculus, Drosophila melanogaster.*

CONCLUSION

Present study was focused on the prediction of the miRNA targets, affected proteins and miRNA homolog of *Mouse gammaherpesvirus68*; animal sequences are taken from mirbase. 98 mRNA targets were found for 58 mature miRNA sequences. Most of the targets encode for signal recognition protein, kynurenine, perilipin and thioredoxin that are very important elements. Homology analysis for miRNA suggest that 17 miRNA of *Mouse gammaherpesvirus68* show 379 miRNA homologs from different animal species, these homologies indicates that miRNAs of most of the animal belongs to same family. *Mouse gammaherpesvirus68* share maximum similarities with *Mus musculus, Homo sapiens, Drosophila melanogaster, Canis lupus familiaris*. Mined potential miRNA targets and affected functions further can be validated by using available approach. Availability of microarray data can be used to assess the relative expression levels of potential mRNA targets in tissues in which their miRNAs were expressed, which will provide the means to understand and validate miRNA targets.

REFERENCES

Hochreiter, R., Ptaschinski, C., Kunkel, S.L., Rochford, R. 2007. Murine gammaherpesvirus-68 productively infects immature dendritic cells and blocks maturation. *Journal of General Virology*, **88:**90-96.

Hwang, S., Wu, T.T., Tong, L.M., Kim, K.S., Guzman, D.M., Colantonio, A.D., Uittenbogaart, C.H. Sun, R. 2008. Persistent Gammaherpesvirus Replication And Dynamic Interaction With The Host in vivo. *J Virology*. 82:12498-12509.

Katara, P., Gautam, B., Kuntal, H., Sharma, V. 2010. Prediction of miRNA Targets, Affected Proteins And Their Homologs in *Glycine max. Bioinformation.* **5:**162-165.

Lok, S.S., Haider, Y., Howell, D., Stewart, J.P., Hasleton, P.S., Egan, J.J. 2002. Murinegammaherpes Virus As A Cofactor In The Development Of Pulmonary Fibrosis In Bleomycin Resistant Mice. *Eur Respir J.* **20:**1228-1232.

Nash, A.A., Dutia, B.M., Stewart, J.P., Davison, A.J. 2001. Natural History of Murine gamma-herpesvirus Infection. *Philos Trans R Soc Lond B Biol Sci*. **356:**569-579.

Nguyen, Y.N., McGuffie, B.A., Anderson, V.E., Weinberga, J.B. 2008. Gammaherpes virus odulation of Mouse Adenovirus Type 1 Pathogenesis. *Virology*. **380:**182-190.

Sunil, N.P., Chandra, J. Arno, Fazakerley, J., Nash, A.A. 1994. Lymphoproliferative Disease In Mice Infected With Murine gammaherpesvirus 68. *Journal of Pathology*., **145**(4) 818–826.

Sarid, R., Olsen, S.J., Moore, P.S. 1999. Kaposi's sarcoma-associated herpesvirus: epidemiology, virology, and molecular biology, *Advances in Virus Research*, **52:**139-232.

Stewart, J.P., Usherwood, E.J., Ross, A., Dyson, H., Nash, T. 1998. Lung Epithelial Cells Are A Major Site Of Murine Gammaherpesvirus Persistence. *Journal of Experimental Medicine*. **187:**1941-1995.

Svobodová, J., Blaskovic, D., Mistríková, J. 1982. Growth Characteristics Of Herpesviruses Isolated From Free Living Small Rodents. *Acta virologica*. **26:**256-263.

Virgin, H.W., Latreille, P., Wamsley, P., Hallsworth, K., Weck, K.E., Canto, A.J., Speck, S.H. 1997. Complete Sequence And Genomic Analysis of Murine Gammaherpesvirus 68. *Journal of Virology.* **71:**5894-5904.

Wheeler, D.L., Church, D.M., Edgar, R., Federhen, S., Helmberg, W., Madden, T.L., Pontius, J.U., Schuler, G.D., Schriml, L.M., Sequeira, E., Suzek, T.O., Tatusova, T.A., Wagner, L. 2004. Database Resources Of The National Center For Biotechnology Information. *Nucleic Acids Res*., **32:**D35-D40.