



## Computational Analysis Reveals Three Micro-RN as in Hepatitis A Virus Genome

Ayaz Ahmad<sup>1,#,\*</sup>, Muhammad Bilal<sup>2,#</sup>, Khalid Khan<sup>2</sup>, Waheed Ahmad<sup>3</sup>, Islam-Ud-Din<sup>4</sup>, Sajid Ali<sup>1</sup>, Amjad Iqbal<sup>5</sup>, And Muhammad Jaseem<sup>6</sup>

<sup>1</sup> Department of Biotechnology, Abdul Wali Khan University Mardan, Pakistan

<sup>2</sup> Department of Chemistry, Islamia College University Peshawar, Pakistan

<sup>3</sup> Department of Computer Science, Abdul Wali Khan University Mardan, Pakistan

<sup>4</sup> Department of Environmental Sciences, International Islamic University Islamabad, Pakistan

<sup>5</sup> Department of Agriculture, Abdul wali Khan University Mardan, Pakistan

<sup>6</sup> Institute of Basic Medical Sciences, Khyber Medical University Peshawar, Pakistan

Received: September 1, 2014

Accepted: November 13, 2014

### ABSTRACT

Micro RNAs (miRNAs) are a class of endogenous non-coding RNAs, 19–25 nt in length, that play pivotal role in the regulation of gene expression by degrading the messenger RNAs of target genes in a sequence-specific manner. Dysregulation of miRNAs results in abnormal gene expression and has been linked to the initiation, advancement and maintenance of some human diseases. Recent studies show that genomes of both virus and host have the potential to encode miRNAs, which may be beneficial either for host or for virus. Hepatitis A is one of the liver inflammations, instigated by the hepatitis A virus (HAV). In this study we, for the first time, computationally identified miRNAs in HAV genome. Initial searches through VMir software extracted 7 sequences with potential hairpin-like structures from HAV genome. MiPred program confirmed 6 candidates as real pre-miRNA hairpin structures. After measurements of free energy and applying other parameters, we confirmed three mature miRNAs in HAV genome. These findings will not only help researchers to explore the role of these miRNAs in viral pathogenesis but also in developing novel antiviral therapies.

**KEYWORDS:** MicroRNAs, miRNAs, Hepatitis, Hepatitis A Virus, HAV.

### INTRODUCTION

MiRNAs, previously known as small temporal RNAs (stRNAs), represent a class of 19–25 nt long, endogenous RNA molecules that play a vital role in post-transcriptional regulation of gene expression by guiding the RNA induced silencing complex (RISC) to bind the messenger RNAs of target genes in a sequence-specific manner thereby causing their cleavage or translational repression [1-3]. These tiny molecules have been implicated in plethora of cellular processes including developmental timing, cell fate determination, neuronal plasticity, cholesterol metabolism, immune responses, apoptosis, cell cycle and tumorigenesis [4]. Lines of evidences suggest that miRNAs are embedded not only in the intergenic regions of genomes but also in protein coding genes [2]. MiRNAs are first transcribed as long transcripts known a primary miRNA (pri-miRNAs). One to several precursors of miRNA (pre-miRNAs) may be embedded inside each pri-miRNA transcript. Consequently, the nuclear RNase III enzyme, Drosha processes each primary miRNA into its constituents of 60–70 nt long precursors of miRNA which fold into an imperfect stem-loop structure(s) and acquire characteristic hairpin shape while still in the nucleus. The resultant miRNA precursors are transported to the cellular cytoplasm by the exportin-5. Here these precursors are further sliced into ~22 nt long duplexes under the action of RNase III Dicer enzyme [2]. The mature miRNA then enters the multiprotein RNA induced silencing complex (RISC). RISC then leads to either degradation or translational silencing of the target mRNA, which in turn depends on the extent of complementarity among the RISC bound micro RNA and target messenger RNA [1, 5].

Beside animals, plants and insects, many virus genomes have been reported to contain miRNAs [4]. These virus encoded miRNAs have been shown to play key roles in virus-host interactions by targeting both host and virus mRNAs of various important genes [4]. The Epstein–Barr virus (EBV) genome was shown to encode five miRNAs, each of which has not only the capability to regulate the expression of virus gene

\* **Corresponding Author:** Ayaz Ahmad, Department of Biotechnology, Abdul Wali Khan University Mardan, Pakistan. Phone# 0092-937-542046, E-mail: ahdazyb5@awkum.edu.pk

involved in latency but also modulates host cell gene expression [6]. The identification of miRNAs in some double stranded DNA viruses revealed their evolution to use RNA silencing potential for regulation of the expression of viral genes, host cellular genes, or both, which needs further investigation to assess whether net benefit is to virus or host [7]. Some of these miRNAs have been identified through experimental strategies like cDNA cloning and confirmed by Northern blotting while the rest are identified computationally [4]. Experimental techniques for identifying viral miRNAs are technically challenging, laborious and time consuming. Computational prediction methods serve as fast, better and more affordable for exploring novel miRNAs and range from custom-made programs used to search for hairpin loops and other features like thermodynamics stability to advanced algorithms using machine learning approaches [8].

Hepatitis A is one of the liver infections, caused by hepatitis A virus (HAV) which has a single chain RNA genome of 7478 nt size. Hepatitis A happens sporadically and in epidemics all over the world. Every year, approximately 14 lac people around the world suffer from this disease [9]. To query whether the strategy of transcribing miRNAs is employed by HAV also, we computationally analyzed HAV genome for miRNA-encoding potential.

## MATERIALS AND METHODS

### Source of Genome

In silico prediction of miRNAs in hepatitis A virus (HAV) was performed by downloading the complete genome sequence of hepatitis A virus. The retrieval of the genome sequence of strain K02990 was carried out by the genome data bank (NCBI). The GenBank entry is (<http://www.ncbi.nlm.nih.gov/nucleotide/329596?report=genbank>). The genome size of this strain is 7478 nucleotides.

### Pre-miRNA Extraction

HAV genome was scanned through Vmir software (program version 2.3, scoring algorithm version 1.4) for hairpin-structure miRNA precursors (pre-miRNAs) [10]. Initially, sequences which acquired fold-back, hairpin shape were considered as potential pre-miRNAs candidates.

### Extraction of Potential Pre-miRNA Candidates

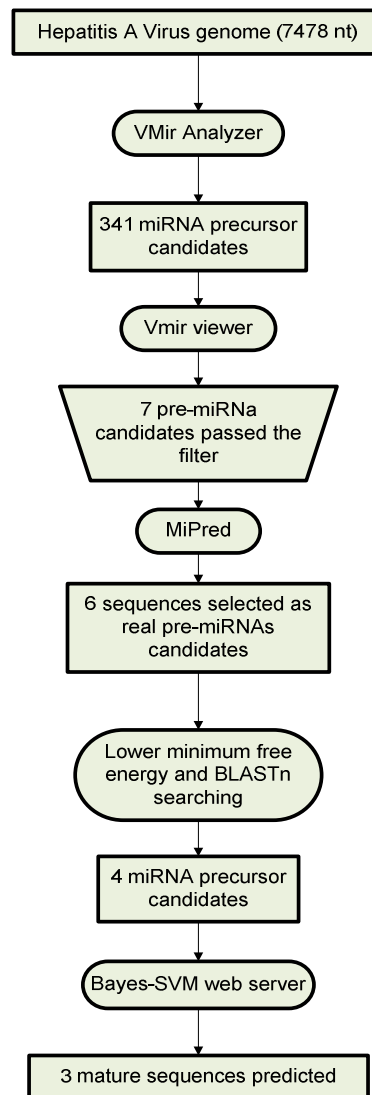
Pre-miRNA candidates were investigated for secondary structure prediction and minimum free energy (MFE). Sequences with a hairpin-like secondary structure, having lower MFE (equal or less than  $-25$  kcal/mol) were selected as potential miRNA precursors.

### Confirmation of Real Pre-miRNAs

In the next step, real and pseudo miRNA precursors were distinguished using MiPred program [11] with RF algorithm (<http://www.bioinf.seu.edu.cn/miRNA/>) [12]. BLASTn tool on the NCBI database was used to keep only unique sequences and remove any repeated sequences.

### Prediction of Mature miRNA

Finally, mature miRNA sequences were predicted by Bayes-SVM-MiRNA online web server v1.0. The web tool can be accessed at (<http://wotan.wistar.upenn.edu/BayesSVMmiRNAfind/>). The overall computational prediction procedure is represented in the form of a flowchart in Figure 1.

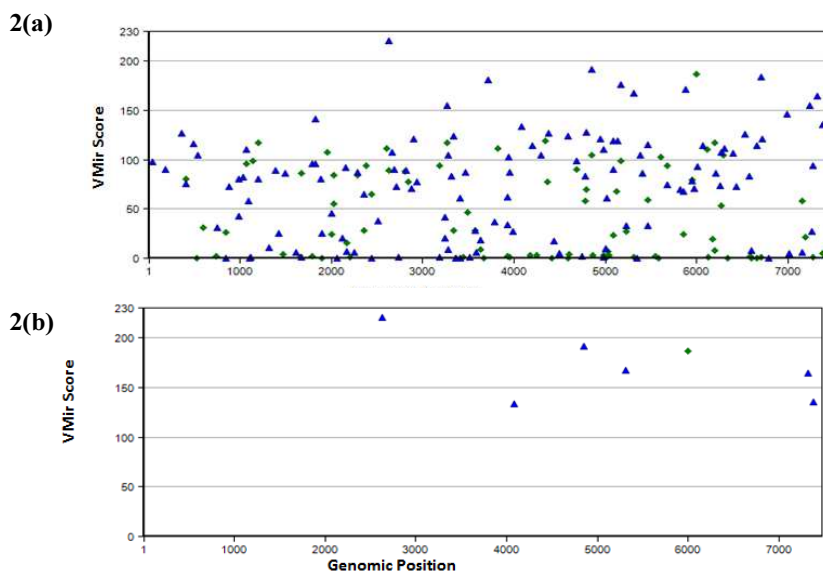


**Figure 1. Flowchart of the Computational Prediction Process**

## RESULTS AND DISCUSSION

VMir analysis of HAV genome reveals that miRNA precursors are extensively distributed across the viral genome and seven high scoring filtered hairpins (with scores between 133 and 220) are located between nucleotides 2500 and 7500. VMir software provides an updated scoring algorithm, particularly designed to predict viral miRNAs [13]. This *ab initio* algorithm has been effectively practiced by several groups earlier for the identification of miRNAs in a number of viral genomes particularly related to herpesvirus and polyomavirus families. It contains numerous quality filters, which make the prediction easier. In the present work, we fed fasta format of HAV genome into VMir analyzer. Scan was performed in both orientations across the viral genome with a window of 500 nt and step size of 10 nt. VMir Analyzer primarily identified a total of 341 sequences as candidate miRNA precursors (Figure 2.a). By applying filter values, set in the earlier studies [13, 14], i.e. 115 for minimal scores and 35 for window counts, only 7 hairpins passed the window filter (Figure 2.b). These candidate miRNA precursors were assigned a VMir score. These candidates are widely dispersed across the viral genome and seven high scoring filtered hairpins (with scores between 133 and 220) are located between nucleotides 2500 and 7500. Figures 2.a and 2.b show the locations and VMir

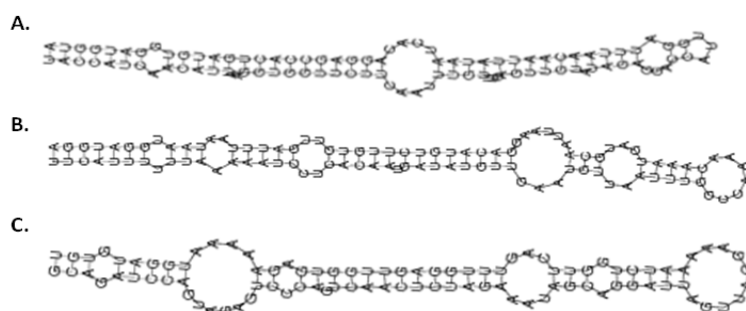
scores for unfiltered and filtered hairpins. These seven sequences with potential hairpin-like structures were analyzed for the secondary structure validation through RNAfold web server (Figure 3). The web tool is available at (<http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi>).



**Figure 2. VMir analysis of the HAV genome;** (a) . represent the unfiltered output from the VMir prediction. All hairpins that fold in 35 or more windows and achieved a VMir score of 115 or above are shown. (b). only those hairpins are shown which passed the filter and achieved a VMir score of 130 and above. Hairpins are plotted according to genomic location and VMir score.

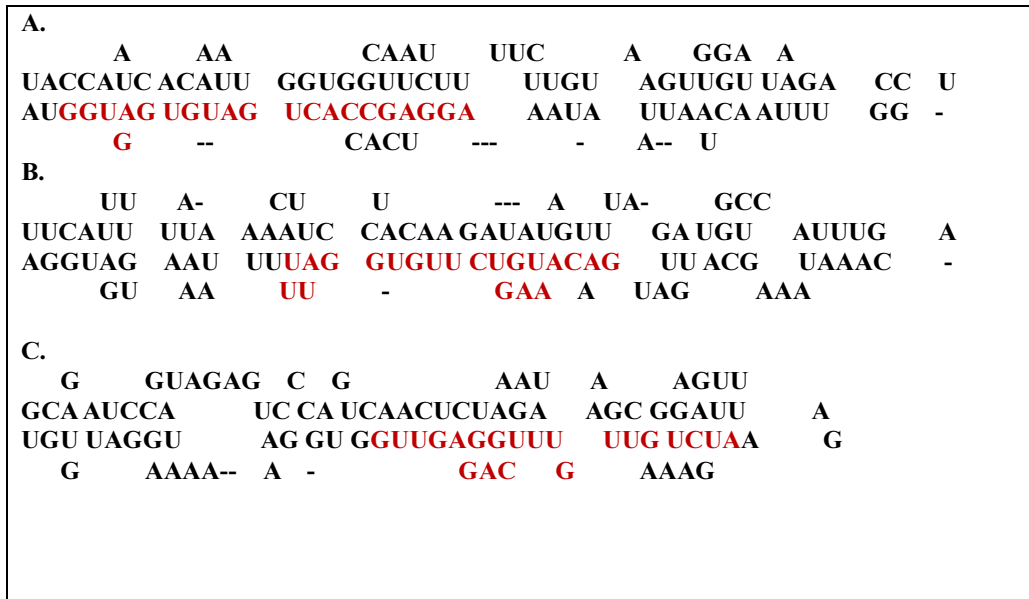
Almost all miRNA precursors attain the characteristic stem-loop hairpin shapes. Therefore, numerous pseudo pre-miRNAs (sequences with analogous stem-loops) can be found in many genomes. In order to distinguish the real pre-miRNAs from pseudo ones, we used MiPred program, a hybrid tool with combined features like local contiguous structure-sequence composition, MFE and a Monte Carlo randomization test [11]. MiPred makes prediction at 98.21% specificity and 95.09% sensitivity. MiPred was used with default parameters to analyze these 7 sequences. Out of the total 7 sequences, MiPred confirmed 6 candidates as real pre-miRNAs like hairpin sequences. After performing BLASTn searches, and analyzing these sequences for MFE, a total of 4 sequences were screened as potential miRNA candidates.

Bayes-SVM-miRNA web server v1.0 offers two classifiers i.e. SVM and Naïve Bayes for the identification of mature miRNA candidates. This web server predicted the formation of mature miRNAs in only three sequences. The positions of these 3 mature miRNAs inside the stem-loop hairpin structures are shown in Figures 3 and 4.



**Figure 3: Secondary Structures of the three pre-miRNAs using RNAfold program**

Computational methods are widely used for the identification of miRNAs most of which mainly rely on hairpin structures of pre-miRNAs as well as other features like evolutionarily conserved nature of sequence. Similarly, other approaches like phylogenetic shadowing strategy remained effective for identification of novel miRNAs. Recently proposed algorithms are independent of microRNA sequence conservation and have facilitated in detection of mouse, human and viruses microRNAs. However, it is quite interesting to know that virus encoded miRNAs have undergone rapid evolution i.e. their homologs are lacking in other viruses [15]. This demands the development of novel and improved algorithms for *ab initio* prediction of microRNAs.



**Figure 4. Secondary Structure Predictions of HAV Pre-miRNAs:** *The putative mature miRNAs sequences are shown in red*

Previous studies targeted on viral miRNAs, mainly in herpesvirus family, shed light on the role of these miRNAs on virus-host interactions during viral infections and pathogenesis [16, 17]. It has been reported that viral miRNAs down-regulate host's immune defense genes and hence take part in immune evasion [18]. Viruses increase the chances of their survival by resisting the host defense system through an intricate strategy which is comprised of protein-facilitated as well as microRNA-facilitated regulations. Viruses, for instance herpesviruses, which have extended dormancy stages, required to retain infected cells of the host alive for a long time period. This session is significantly prolonged by those viruses which cause latent infection. Thus, viral miRNAs can support virus replication by at least two ways i.e. extending cell existence and eluding immune recognition. It has been suggested that those microRNAs which are transcribed by dormant or tumor developing viruses may interact with the host cellular factors which are assumed to be responsible for antiviral processes and hence creating the cellular environment favorable to viral latency and oncogenesis [19].

**Table 1. Sequences and Genomic Positions of Predicted miRNAs in HAV Genome**

S.No	Predicted mature miRNA sequence (5' to 3')	Position, orientation
1	AGGAGCCACUGAUGUGGAUGG	2653-2673, +
2	AAGGACAUGUCUUGUGUUGAU	4868-4888, +
3	AUCUGGUUCAGUUUGGAGUUG	5325-5345, +

Future studies combining bioinformatics with microarray analysis would be helpful to clear the image of host-pathogen interactions modulated by viral microRNAs. Moreover, the functional analysis of the identified HAV microRNAs in pathological processes is mandatory which would be helpful in designing new preventive and antiviral therapeutic strategies.

## ACKNOWLEDGEMENTS

The authors would like to thank Professor Zuhong Lu of Southeast University, Nanjing, China for his help in using MiPred program. Dr. Ayaz Ahmad's Lab is funded by Abdul Wali Khan University Mardan (AWKUM), Higher Education Commission (HEC) of Pakistan, Directorate of Science and Technology (DoST) KPK and Relief International (RI).

## REFERENCES

1. Moreno-Moya, J.M., F. Vilella, and C. Simón, *MicroRNA: key gene expression regulators*. Fertility and sterility, 2013.
2. Dong, H., et al., *MicroRNA: function, detection, and bioanalysis*. Chemical reviews, 2013. **113**(8): p. 6207-6233.
3. He, L. and G.J. Hannon, *MicroRNAs: small RNAs with a big role in gene regulation*. Nature Reviews Genetics, 2004. **5**(7): p. 522-531.
4. Grundhoff, A. and C.S. Sullivan, *Virus-encoded microRNAs*. Virology, 2011. **411**(2): p. 325-343.
5. Okamura, K., et al., *Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways*. Genes & development, 2004. **18**(14): p. 1655-1666.
6. Pfeffer, S., et al., *Identification of virus-encoded microRNAs*. Science, 2004. **304**(5671): p. 734-736.
7. Sullivan, C.S. and D. Ganem, *MicroRNAs and viral infection*. Molecular cell, 2005. **20**(1): p. 3-7.
8. Gomes, C.P., et al., *A review of computational tools in microRNA discovery*. Frontiers in genetics, 2013. **4**.
9. Matheny, S.C. and J.E. Kingery, *Hepatitis A*. Am Fam Physician, 2012. **86**(11): p. 1027-34; quiz 1010-2.
10. Sullivan, C.S. and A. Grundhoff, *Identification of viral microRNAs*. Methods in enzymology, 2007. **427**: p. 1-23.
11. Jiang, P., et al., *MiPred: classification of real and pseudo microRNA precursors using random forest prediction model with combined features*. Nucleic acids research, 2007. **35**(suppl 2): p. W339-W344.
12. Xue, C., et al., *Classification of real and pseudo microRNA precursors using local structure-sequence features and support vector machine*. BMC bioinformatics, 2005. **6**(1): p. 310.
13. Grundhoff, A., C.S. Sullivan, and D. Ganem, *A combined computational and microarray-based approach identifies novel microRNAs encoded by human gamma-herpesviruses*. Rna, 2006. **12**(5): p. 733-750.
14. Xiang, J., et al., *Computational identification of microRNAs in Anatid herpesvirus 1 genome*. Virology journal, 2012. **9**(1): p. 1-4.
15. Cai, X., et al., *Kaposi's sarcoma-associated herpesvirus expresses an array of viral microRNAs in latently infected cells*. Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(15): p. 5570-5575.
16. Pfeffer, S., et al., *Identification of microRNAs of the herpesvirus family*. Nature methods, 2005. **2**(4): p. 269-276.
17. Veksler-Lublinsky, I., et al., *Gene bi-targeting by viral and human miRNAs*. BMC bioinformatics, 2010. **11**(1): p. 249.
18. Stern-Ginossar, N., et al., *Host immune system gene targeting by a viral miRNA*. Science, 2007. **317**(5836): p. 376-381.
19. Yao, Y. and V. Nair, *Role of Virus-Encoded microRNAs in Avian Viral Diseases*. Viruses, 2014. **6**(3): p. 1379-1394.