

Cross-kingdom sequence similarities between human micro-RNAs and plant viruses

Jovan D Rebolledo-Mendez¹, Radhika A Vaishnav^{2,3}, Nigel G Cooper,^{4,*} and Robert P Friedland^{2,*}

¹Department of Biochemistry and Molecular Biology; University of Louisville, KY USA; ²Department of Neurology; University of Louisville, KY USA; ³Department of Physiology and Biophysics; University of Louisville, KY USA; ⁴Department of Anatomical Science and Neurobiology; University of Louisville, KY USA;

Keywords: plant virus, human miRNA, molecular mimicry, sequence homology, evolutionary conservation, bioinformatics, health, *Tobacco mosaic virus*, post-transcriptional regulation, environment

Micro-RNAs regulate the expression of cellular and tissue phenotypes at a post-transcriptional level through a complex process involving complementary interactions between micro-RNAs and messenger-RNAs. Similar nucleotide interactions have been shown to occur as cross-kingdom events; for example, between plant viruses and plant micro-RNAs and also between animal viruses and animal micro-RNAs. In this study, this view is expanded to look for cross-kingdom similarities between plant virus and human micro-RNA sequences. A method to identify significant nucleotide sequence similarities between plant viruses and human micro-RNAs was created. Initial analyses demonstrate that plant viruses contain nucleotide sequences which exactly match the seed sequences of human micro-RNAs in both parallel and anti-parallel directions. For example, the bean common mosaic virus strain NL4 from Colombia contains sequences that match exactly the seed sequence for micro-RNA of the hsa-mir-1226 in the parallel direction, which suggests a cross-kingdom conservation. Similarly, the rice yellow stunt viral cRNA contains a sequence that is an exact match in the anti-parallel direction to the seed sequence of hsa-micro-RNA let-7b. The functional implications of these results need to be explored. The finding of these cross-kingdom sequence similarities is a useful starting point in support of bench level investigations.

Introduction

Micro-RNAs (miRs), first described approximately two decades ago,¹ have gained much attention in the research community for their proposed regulatory influences on biological systems.² Composed of short non-coding RNA molecules of 19–26 nucleotides in length, miRs can post-transcriptionally regulate gene expression and influence cellular phenotypes, such as: cell cycle regulation,³ oral carcinoma,⁴ heart disease⁵ and development of the central nervous system.⁶ Thus, miRs are key modulators of messenger-RNAs (mRNAs) usually operating through their complementary interactions with the 3' untranslated regions (3' UTRs) of mRNA sequences. These complementary interactions between miRs and mRNAs regulate the process of translation and thereby alter the abundance of specific proteins and modulate downstream cellular processes.⁷

Interference of the miR-associated regulatory events will likely have relevance to the etiology and expression of particular cellular and tissue phenotypes; regulatory events will also be involved in the etiology and/or expression of many pathologies. The putative target genes whose expression is regulated by particular miRs can be predicted by statistical⁸ and computational

modeling methods^{9–17} if the miR sequence is known. Similarly, bioinformatics approaches have been used for the prediction of novel miRs,¹⁸ as well as for predicting the targets of miRs.^{12,16,19}

The miR based regulatory system is also present in plants,^{20,21} and a recent provocative study described the presence of plant derived miRs in mammalian serum.²² This study demonstrates the ability of plant miRs to escape the GI tract and enter various organs and tissues of the body and influence gene expression.²² The importance of this report lies in its indication of cross-kingdom interactions. There are other recent reports of potential involvement of miRs in cross-kingdom²³ and cross-species²⁴ interactions.

Historically, plant viruses have generally been considered safe for human exposure, with little direct evidence of their ability to replicate in humans,²⁵ or cause or contribute to cellular function or disease. However, it was recently shown that *Coupea mosaic virus* (CPMV) can enter the bloodstream after oral administration²⁶ and can localize in endothelial cells of the blood brain barrier.²⁶ A commonly consumed plant virus, the *pepper mild mottle virus* (PMMV) is detectable in stool and can induce an IgM antibody response in humans.²⁷ Besides the possibility of invoking immune responses, plant viruses may be capable of influencing

*Correspondence to: Robert P. Friedland, Email: robert.friedland@louisville.edu; Nigel G. Cooper, Email: nigel.cooper@louisville.edu

Submitted: 04/25/13 Accepted: 05/06/13 Published Online: 05/21/13

Citation: Rebolledo-Mendez JD, Vaishnav RA, Cooper NG, Friedland RP. Cross-kingdom sequence similarities between human micro-RNAs and plant viruses. Commun Integr Biol 2013; 6: e24951. <http://dx.doi.org/10.4161/cib.24951>

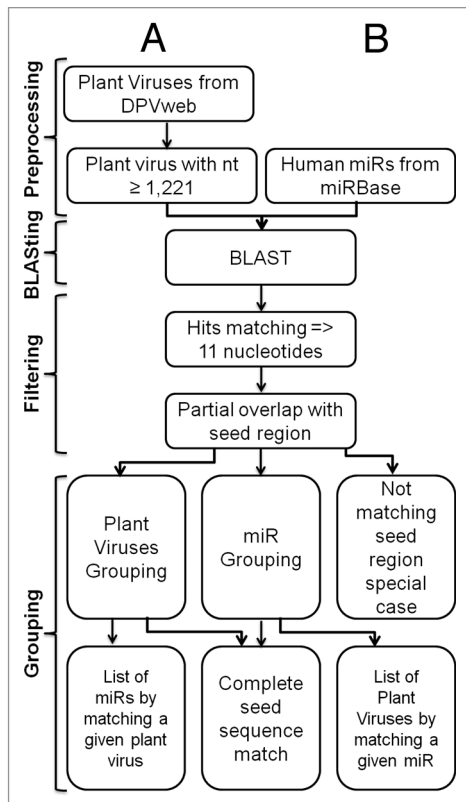


Figure 1. Pipeline method for determination of similarities between plant viruses and human miRs. Process pipeline developed for finding similarities between named plant viruses and human miRs. Step 1 includes the establishment of the two data sets including (A) plant viruses downloaded from DPVweb, and (B) hsa-miRs, downloaded from miRBase. Step 2 includes the use of NCBI BLASTn for sequence alignments to find sequence similarities between data sets A and B, where the maximum expectation value found is 9.5 and the shortest matching sequence is 6 nucleotides. Step 3 includes the use of filters based on rules: (A) plant viruses must have at least 1,221 nucleotides; (B) there must be at least 11 or more nucleotides of similarity in such match; and (C) the BLAST sequence should start in any of the nucleotides from within the miR seed region. The final database can be used for additional data integration and analyses. It was used here to perform clustering analyses and also for identification of viruses that contained sequences that showed similarities with miR seed sequence.

human gene expression via more direct interactions with genetic materials. The need to look for miR-like sequences in plant viruses is further supported by a recent bioinformatics study demonstrating that some animal viruses contain miRs.²⁸

Our goal was to demonstrate if plant virus sequences have sequence similarity with hsa-miRs, either in the parallel (plus/plus) or antiparallel (plus/minus) directions; by similarity meaning the percent of identity of nucleotide sequences. In pursuit of these objectives, we developed a sequence processing pipeline, where data has been filtered and carefully curated to allow for rational selection of candidate plant virus- human miR interactions to be used for future bench assays of function, in order to better understand cross-kingdom interactions, which may be beneficial or harmful.

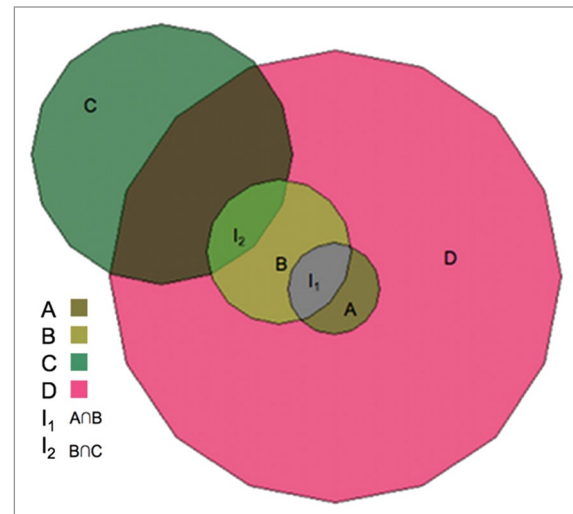


Figure 2. Overlapping clusters of plant viruses based on human miRs with which they have sequence similarities. Venn diagram showing the four subsets (A–D), based on the bit-score for all the plant viruses that had similarity sequences with human miRs. These data were obtained as the result from Step 3 (Fig. 1). The clusters represent the entire holdings in the database. The priority in this diagram was to show the BLAST bit-scores distributed among the different subsets of plant viruses. Subset (A) (color olive) contains 25 matching viruses with score of 28.4 bits and 14 nucleotides long; subset (B) (color dark khaki) contains 63 matching viruses with score of 30.2 bits and 13 nucleotides long; subset (C) (color dark green) contains 204 matching viruses with score of 24.3 bit and 12 nucleotide long; subset (D) (color pink) contains 615 matching viruses with score 22.3 bit and 11 nucleotides long; cluster I₁ (color gray) contains 14 matching viruses with score of 24.3 bits (14 nucleotides) and 26.3 (13 nucleotides); and cluster I₂ (color bright green) contains 1 matching virus, rice tungro spherical (am234049) with scores of 26.3 bits (13 nucleotides long), 24.3 bits (12 nucleotides long) and 22.3 bits (11 nucleotides long).

Results

For step 1, we used DPVweb,²⁹ which contains both complete and partial genomes for plant viruses. Therefore, for plant viruses, 37,266 genome sequences were downloaded from DPVweb.²⁹ A total of 1,424 human miRs were downloaded from miRBase³⁰ as data set-B (Fig. 1). As many as 478 of these miRs had two mature sequences. Therefore, the total number of sequences in data set-B was actually 1,902. The final data sets from these external databases included 37,266 plant viral sequences in data set-A and 1,902 miR sequences in data set-B.

In step 2 of the pipeline, the alignment of nucleotide sequences from among these two data sets was accomplished with the aid of BLASTN version 2.2.25+.^{31,32} The alignment resulted in two streams of output which included “matches” between viral and miR sequences, consisting of strings of six or more consecutive nucleotides, and “no-matches.” The data stream without matching sequences was discarded and the stream with matching sequences was saved and used for further processing.

The saved data stream contained over twenty-thousand matching sequences, which were of two categories. One category included 11,653 matching pairs of parallel sequences (plus/plus)

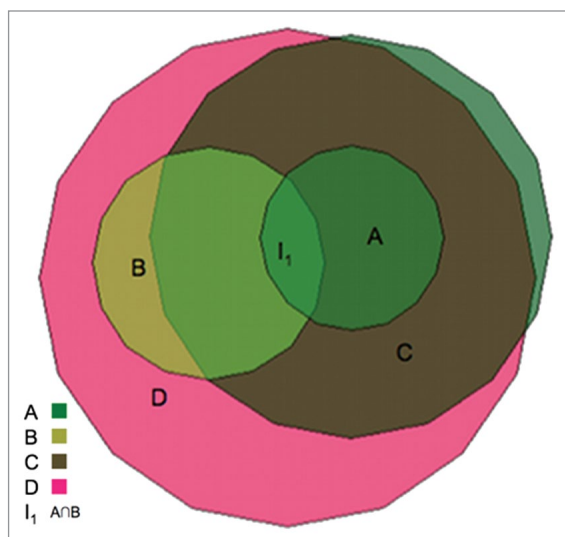


Figure 3. Overlapping clusters of hsa-miRs based on the number of plant viruses which they have sequence similarities. Diagram showing the clustering of human miRs based on their sequence similarities with plant viruses. These data were obtained as an outcome from Step 3 (Fig. 1). This figure represents the entire holdings in the database. The priority for this particular result was set to show the subsets of hsa-miRs based on their BLAST bit-scores. Subset (A) (color bright green) contains 10 matching miRs with score of 28.2 and 14 nucleotides long; subset (B) (color dark khaki) contains 16 matching miRs with score of 26.3 bits and 13 nucleotides long; subset (C) (color dark green) contains 48 matching miRs with score of 24.3 bits and 12 nucleotides long; subset (D) (color pink) contains 73 matching miRs with 11 nucleotides long; and subset (I_1) (color yellow green) contains 2 matching miRs that are contained in each other score subset.

and the other category included 11,595 matched pairs of anti-parallel (plus/minus) or complementary sequences. The output of step 2 was a list of statistically significant matching sequences that showed some level of sequence matching (plus/plus and plus/minus matching) between plant viral sequences and human miR sequences.

The data stream was further processed in step 3 with the aid of selection filters to insure that only high quality matches were retained. For example, one of the filters selected only those matching sequences whose plant virus length part would be at least 1,221 nucleotides to avoid possible duplications from partial viral sequences. In addition, the sequence match had to be 11 or more nucleotides in length, and this could include the first nucleotide of the mature miR. This filter made it less likely that sequence matches were the result of chance. Finally, we included the matches whose sequence region included the complete seed region (nucleotides 2–7) of the miR.³³ The total filtering process led to reduction of the data to include a total of 2,479 matching sequences, including parallel and anti-parallel sequence sub-categories. From the 2,479 matched sequence pairs there were interactions between 78 miRs (4.1% of the 1,902 miRs) and 1,320 different plant viruses (3.54% of the 37,266 plant viruses).

Data analyses

A clustering analysis of the sequences in the database was adopted, as described below, to get a better understanding of the overall content of the database.

Plant viruses

This analysis showed the presence of four major clusters (A–D), in which plant viruses were matched through various levels of sequence identity with one or more hsa-miRs. The identification of clusters was based on the BLAST bit-scores and the length of their shared sequence similarities (Fig. 2). The four clusters were labeled according to size. For example, the smallest cluster A contained 25 viruses, each of which had sequences of 14 nucleotides in length that shared identities with hsa-miRs. The cluster B contained 63 viruses with slightly shorter sequence identities and cluster C contained 204 matching viruses with still shorter sequence identities. Cluster D, the largest cluster, contains 615 viruses each with sequences of 11 nucleotides in length sharing identities with hsa-miRs. Each of the major clusters intersected with one or more of the other clusters. For example, the intersection between clusters A and B consists of 14 viruses, with either 14 or 13 nucleotide sequence lengths, sharing identities. The

Table 1. Plant virus ranking

Rank	Virus Name	DPVweb plant virus index	Virus strain	hsa miRs
1	Rice yellow stunt virus viral cRNA	ab516283		mir-1253, mir-1248, let-7b, let-7a-1, let-7a-3, mir-107, let-7a-2, mir-103a-1, mir-1-1, let-7c, mir-103b-1, mir-103b-2
2	Citrus tristeza virus	eu937520	Strain T30	*mir-122, mir-1244-1, mir-1244-2, mir-1244-3, let-7f-1, let-7f-2, mir-1184-1, mir-1184-2, mir-1184-3, mir-1228
		fj525431, fj525433, fj525434, fj525435,	Isolate NZRB-M12, isolate NZRB-TH28, isolate NZRB-TH30, isolate NZRB-M17	mir-1244-1, mir-1244-2, mir-1244-3, let-7f-1, mir-1184-1, mir-1184-2, mir-1184-3, let-7f-2, mir-1180, mir-1193
		fj525432	Isolate NZRB-G90	mir-1244-1, mir-1244-2, mir-1244-3, let-7f-1, mir-1184-1, mir-1184-2, mir-1184-3, let-7f-2, mir-1193
		af260651	T30	*mir-122, mir-1244-1, mir-1244-2, mir-1244-3, let-7f-1, mir-1184-1, mir-1184-2, mir-1184-3, let-7f-2

Ranking of the two plant viruses that have the most number of sequences matching across the human miRs. *Rice yellow stunt virus viral cRNA* followed by *Citrus tristeza virus* were the plant viruses that ranked the best matching similarities. The latest plant virus has seven

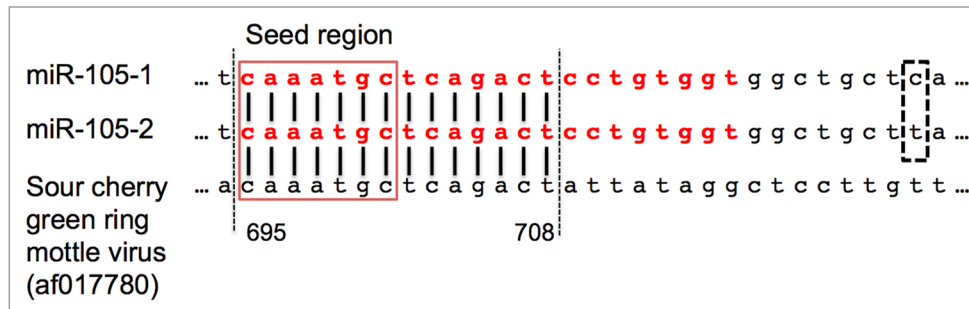


Figure 4. Nucleotide sequences showing similarity between miR-105-1, miR-105-2 and *sour cherry green ring mottle virus* (af017780). The partial sequences for miR-105-1 and miR-105-2 are aligned with a section of af017780. Note that miR105-1 and miR105-2 are both quite similar, with the exception of a single nucleotide (dotted-lined box). Both miR-105-1 and miR-105-2 contain 81 nucleotides. The length of the plant virus is 8,372 nucleotides, but the significant subset, nucleotides 695–708, that accompanies the similarities with miR-105-1 and miR-105-2 is shown here. The solid-lined box shows the seed region matching completely in all three sequences.

Table 2. miR ranking

Rank	miRs	Number of viruses/strains
1	hsa-mir-1253	136
2	hsa-mir-1243	110
3	hsa-mir-105-1	83
3	hsa-mir-105-2	83
4	hsa-mir-1207	81

Top rank ordered hsa-miRs that share identity with one or more plant viruses, with 11 or matching nucleotides including the complete seed region

intersection between B and C contained just one virus, rice tungro spherical virus (am234049) that matched sequences of either 11 or 12 nucleotides in length depending on the major cluster. While clusters A and B are subsets of cluster D, cluster C was not. It overlapped with clusters D and B, but not with A.

Rank-ordered plant viruses

The viruses that shared some sequence identity (plus/plus or plus/minus) with one or more miRs were rank-ordered according to the number of miRs with which they showed some level of similarity. The two top ranked viruses are listed (Table 1) with their corresponding miR partners. The top ranked plant virus was the rice yellow stunt virus (ab516283), which has sequence similarities with twelve different human miRs. The second most frequent was *Citrus tristeza virus*, for which there are multiple strains. Strain T30 (eu937520) matched with ten miRs. Strains [*Citrus tristeza virus*-isolate *NZRB-M12* (fj525431); *Citrus tristeza virus* isolate *NZRB-TH28* (fj525433); *Citrus tristeza virus* isolate *NZRB-TH30* (fj525434); and, *Citrus tristeza virus* isolate *NZRB-M17* (fj525435)] all matched the same ten miRs. Finally, *Citrus tristeza virus* isolate *NZRB-G90* (fj525432) shared some sequence identity with nine miRs. *Citrus tristeza virus* T30 (af260651) also shares sequence identity with nine different miRs.

Human micro-RNAs

This analysis showed the presence of four major miR-clusters (A–D), in which hsa-miRs shared some level of identity with one or more viruses based on the quality of the BLAST bit-scores

length of shared sequence similarities (Fig. 3). For example, the smallest cluster A had the best bit-scores, with ten miRs, each of which had a 14 nucleotide length of sequence similarities with plant viruses. The cluster B had 16 miRs each with a 13-nucleotide length which had sequence identities with viruses. Cluster C contained 48 miRs with lengths of 12 nucleotides sharing identities with viruses. Cluster D contained 73 miRs, each of which had sequence similarities of 11 nucleotides in length with viruses. Some of the clusters had overlapping subsets. For example, the intersection of clusters A and B contained two miRs shared by each cluster. In addition, several miRs were found to have more than one sequence similarity with a virus sequence. While there are 314 such matches, there was only one instance of a miR having three sequence similarities with a virus sequence (data not shown).

Rank-ordered micro-RNAs

The top four rank-ordered human miRs are listed (Table 2) with their corresponding numbers of plant virus partners. These sequences contained the complete seed sequence of the identified miR and sequence similarities of 11 or more nucleotides in length. Some of the plant viruses have a number of different strain identities (for example, see Table 1). The miR with the highest number of shared sequence identities was hsa-mir-1253 with some sequence identity to 136 plant viruses/strains (Table 2). The second highest was hsa-mir-1243 with some level of similarity to 110 plant viruses. The third ranked included hsa-mir-105-1 and hsa-mir-105-2 with similarities to the same 83 plant viruses. The fourth ranked miR was hsa-mir-1207, which had similarities to 81 plant viruses. A top four miR listing, rank ordered by the number of shared identities with named plant viruses, is available as a Table S1.

Analysis of the seed sequence

The region of the miRs considered to be the most important in miR-mRNA interactions is the miR-seed region (nucleotides 2–7). For our analysis, the total seed sequence of the miRs had to be evident in the viral sequences in a parallel or complementary fashion (plus/plus or plus/minus). In addition, for this particular analysis, the total sequence similarity had to be at least 11 nucleotides long, which is about 50% the length of a mature

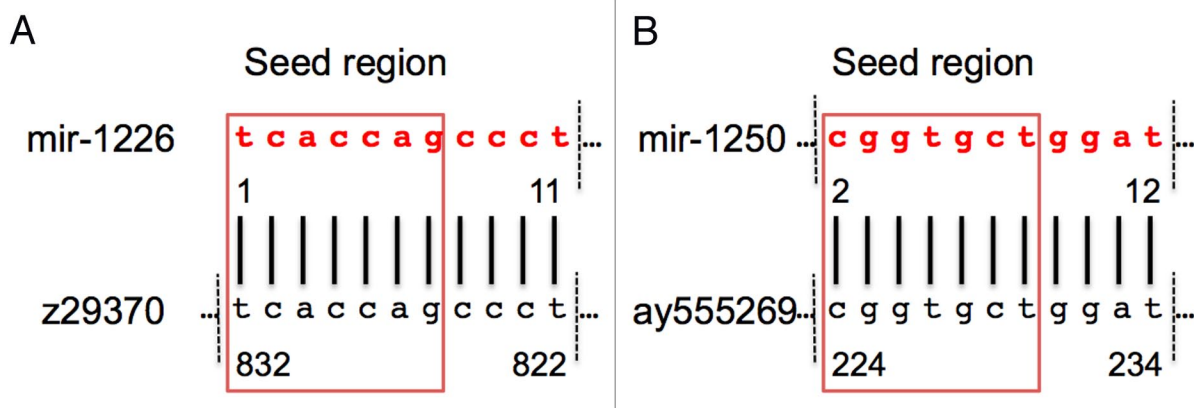


Figure 5. Two strains of *Tobacco mosaic virus*, z29370 and ay555269, show sequences similarity to human miRs. The mir-1226 and virus z29370 have an antisense similarity (plus/minus), while mir-1250 and virus ay555269 have a sense (plus/plus) similarity. The seed region of the miRs is shown in the solid-line box. In (A) the similarity starts one nucleotide before the seed region, and it is followed by three nucleotides after such similarity. (B) shows the similarity is extended four nucleotides after the seed region.

miR. Together, these requirements led to the output of a total of 1,073 sequence matches.

This list was examined for sequence matches that appeared to be the most representative of genuine miRs. For example, it contained sequence matches between viruses and miRs of between 11 and 15 nucleotides in length. Two subgroups were also identified: (1) those in which the seed sequence between nucleotides 2–7 were identical between plant virus and miR; and (2) those in which the seed sequences were identical and where the first nucleotide prior to the seed sequence was also a match. These requirements led to a list of 555 sequences with the plus/plus or parallel configuration (whose selected top 20 candidates are found in the **Table S2**). A few of these are listed below (**Table 3**). While there were a number of sequence similarities between plant viruses and miRs which had lengths exceeding 14 nucleotides, their bit-scores were either not as good as those described above or they did not match the complete seed sequence (data not shown). Perhaps it is not surprising that strains of a given virus have the same sequence similarities even when it contains a miR seed sequence (**Table 3**, **Fig. 4**).

Statistical analysis of similarities

Although we used BLAST's bit-score to define statistical significance, we decided to compare these results with those obtained from randomly generated sequences. The correspondent p values were equal to $8.94E-3$ for the length of the sequence similarity, and equal to $4.23745E-10$ for the relative location of the similarity inside the plant virus genome. These results demonstrate that two data sets, namely the similarities from the database obtained from the randomly generated plant viruses and the one using the real human miRs database, are significantly different from each other and that their statistical distribution is also different. These results indicate the presence of a much higher degree of specificity in the nature.

Specific cases

While the presence of viruses in plants has obvious relevance to the agricultural economy, it has been suggested that there may be consequences for human health.^{26,27,34} The presence of miR-like sequences in plant viruses associated with four particular plants where some health relevance is clearly plausible was investigated. In this case, the viral sequence had to match half the length of

Table 3. Best plus/plus matching ranking based on seed region

DPVweb plant virus index	Virus name	Virus sequence	miR sequence	miR name
dq666332	<i>Bean common mosaic virus strain NL4 from Colombia</i>	gtgagggcatgcag	gtgagggcatgcag	1226
eu761198	<i>Bean common mosaic virus isolate MS1</i>	gtgagggcatgcag	gtgagggcatgcag	1226
af017780	<i>Sour cherry green ring mottle virus</i>	caaatgctcagact	caaatgctcagact	105-1
aj312438	<i>Bean common mosaic virus cowpea isolate Y</i>	gtgagggcatgcag	gtgagggcatgcag	1226
af017780	<i>Sour cherry green ring mottle virus</i>	caaatgctcagact	caaatgctcagact	105-2
ay575773	<i>Blackeye cowpea mosaic virus</i>	gtgagggcatgcag	gtgagggcatgcag	1226
aj312437	<i>Bean common mosaic virus cowpea isolate R</i>	gtgagggcatgcag	gtgagggcatgcag	1226
l32603	<i>Sonchus yellow net virus</i>	cagttatcacagt	cagttatcacagt	101-1

Highest ranking viruses which contain the complete seed sequence of hsa-miRs with plus/plus strand and a similarity of 14 nts. Among these, *Bean common mosaic virus* has different strains that have a high matching with the seed region of miRs. miR 1226 shares complete seed region matching with two different plant viruses.

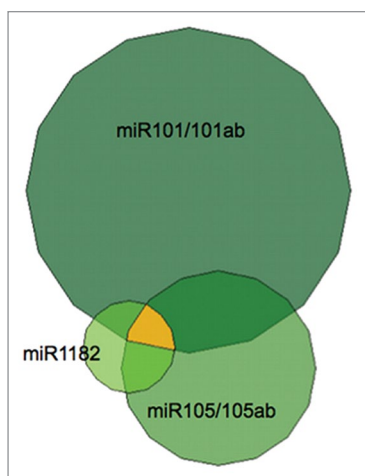


Figure 6. Gene sets similarities across miRs relevant to miR-101/101ab, miR-105/105ab, and miR-1182. Clusters from the sets that contain the target genes, taken from miRBase, for the three miRs: miR-101/101ab (with 803 genes), miR-105/105ab (with 479 genes) and miR-1182 (with 110 genes). The cluster that belongs to all the three groups is formed by 6 genes/loci: LOC100507421, KLF12, C17orf39, RUNX1, ANKRD52, CAMK2G.

a miR (11 nucleotides) or more. Those data that included complete seed sequences are summarized below and in Table 4. For these following analyses the matching sequences that contained the virus in both forms: complete and incomplete genomes were selected.

Tobacco mosaic virus (TMV) is a positive-sense, single-stranded RNA virus that infects more than 150 plant species including tobacco, tomatoes and peppers.³⁵ TMV is found in cigarettes and chewing (smokeless) tobacco and in the saliva of smokers,³⁶ and we have detected antibodies to TMV in human serum.^{34,37} Two TMV strains shared some sequence identities with hsa-mir-1226 and hsa-mir-1250, with a plus/minus and a plus/plus strand respectively (Fig. 5 and Table 4). Both of these shared sequences included coverage of the complete miR seed sequences (Fig. 5).

Potato virus Y (PVY) is an RNA virus in the family Potyviridae that infects potatoes, peppers, tomatoes and tobacco. Infection with PVY limits crop yield but does not destroy all growth. PVY is found worldwide, and it is estimated that 15% of potato crops are infected. It is likely that some potatoes consumed by humans are infected with PVY.³⁴ Previous studies have shown similarities between this plant virus protein and amyloid A β 1–42 peptide, a protein found in plaques in Alzheimer-diseased brains. Specifically, mouse antibodies to PVY bind to amyloid A β .³⁴ An analysis showed sequence similarity of PVY with five human miRs including hsa-mir-1243, hsa-mir-101-1, hsa-mir-105-1, hsa-mir-105-2 and hsa-mir-1182 (Table 4).

The miR-gene targets for each of these miRs were identified in miRbase (data not shown). The gene targets for these miRs were examined for any potential overlap (Figs. 6 and 7). Overlapping gene identities included KLF12, C17orf39, RUNX1, ANKRD52, CAMK2G and the locus ID, LOC100507421 (Fig. 6). Similarly, the clusters of the targeted genes for miR-1243, miR-101 and miR-105 contained three genes in common, including SMAD2, SLMO2, NAP1L1 (Fig. 7).

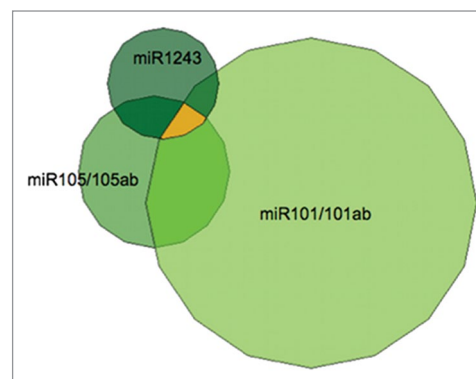


Figure 7. Gene sets similarities across miRs relevant to miR-101/101ab, miR-105/105ab, and miR-1243. Groups from the sets that contain the target genes, taken from miRBase, for the three miRs: miR-101/101ab (with 803 genes), miR-105/105ab (with 479 genes) and miR-1243 (with 139 genes). The cluster that belongs to all the three groups is formed by three genes: SMAD2, SLMO2, NAP1L1.

Cowpea mosaic virus (CPMV) is a plant comovirus in the superfamily of picornaviruses, and which has recently been explored as a prospective nanoparticle delivery system for use via oral and systemic routes. It has been shown to enter the bloodstream and cells of mammals following oral intake.^{26,38} Plant virus strains af395678, aj515903, s66280 and x00206 share sequence similarity with hsa-mir-1226. Strains af395678 and s66253 have a match with hsa-mir-1245b. Strain dq198144 has a match with hsa-mir-1228, and strain eu170481 has matches with three miRs including: hsa-mir-124-1, hsa-mir-124-2 and hsa-mir-124-3 (Table 4).

Pepper mild mottle virus (PMMoV) is a non-enveloped, rod-shaped, single-stranded, positive-sense RNA virus classified in the genus Tobamovirus. It is extremely resistant to physical and chemical agents and is one of the major pathogens of *Capsicum* (chili peppers). The *pepper mild mottle virus* is highly abundant in human stool²⁷ and is easily transmitted by the oral route, by handling during cultivation, and is also transmitted through contaminated seeds. Seven strains of PMMoV (aj308228, fn594778, fn594779, fn594780, fn594781, fn594782, fn594783, fn594784, fn594785, fr671374) contain sequences that shared sequence identity including the complete seed sequence with the miR hsa-let-7i (Fig. 8 and Table 4).

Discussion

We report here the novel finding of sequence similarities between plant viruses and human cellular miRs. To our knowledge, such a large-scale comparison of plant virus and human miR sequences has not previously been reported. While miRs have been identified in DNA-viruses,³⁹ there has been less evidence for the presence of miRs in RNA viruses and it is currently held that those viruses most likely to encode miRs would be DNA viruses.⁴⁰ The novel finding of cross-kingdom similarities, in particular between the seed sequences of plant viral sequences, most of which are RNA viruses, and hsa-miRs, may provide an indication of the evolutionary conservation of this particular structural-functional

feature of miRs. As viruses are known to undergo rapid molecular evolution it may be useful to explore further the relative conservation between the virus seed-like and non-seed sequences. Computational tools are helping to open up the field of host-pathogen interactions, in part through the ability to assess similarities and/or complementarities between nucleotide sequences in large data sets.

The results presented here indicate the presence of sequence similarities both for the parallel and anti-parallel directions. One example of similarity in the parallel (plus/plus) direction between a plant virus and a hsa-miR is the *bean common mosaic virus* cowpea isolate Y and hsa-mir-1226, where 14 nucleotides are shared and which includes the miR's seed region. It remains to be seen if a full length viral miR would be comprised of this 14 nucleotides sequence plus other contiguous nucleotides, but certainly, virus encoded miRs have been reported.⁴¹ The existence of viral miR-like sequences opens the intriguing possibility that they may be able to interact with cellular mRNAs, and essentially compete with cellular miRs. In the human, there are 312 conserved targets for hsa-mir-1226, including NRP1, involved in control of cell migration, and NPAS3, involved in neurogenesis. Thus, viral miR-like sequences could have far reaching consequences if they gain access to the cellular machinery. However, it is not clear that the real targets of plant virus miRs will be found in the cells of metazoans, whether plant or animal, and it is possible that the natural targets of viral miR-like sequences will be found within the RNA genome of the virus itself. These are questions that await further bench level investigations.

The principal action of cellular miRs is to downregulate the activity of cellular mRNAs through complementary base-pairing interactions. Perez-Quintero and colleagues (2010) propose a broader view in which plant miRs provide antiviral activity for plants.²³ Similarly, animal miRs targets appear to have antiviral properties against animal viruses.^{42,43} Plant cellular miRs do not appear to have more putative base-pairing targets with plant infecting viruses than with animal infecting viruses.²³ If so, it would seem reasonable to suggest that the defensive nature of the miR activities is a conserved property which to some extent, at least, employs a mode of action that crosses kingdoms. This suggests that base-pairing between miR sequences with other RNA nucleotide sequences may have implications beyond the cellular mRNA regulatory machinery with which it has been characteristically associated. For example, hsa-miR-122 has been shown to interact with complementary sites on the Hepatitis-C RNA genome⁴⁰ and may be regarded as a regulator of viral replication.²³ Another example is the inhibitory effect exerted by *respiratory syncytial virus* (RSV) on hsa-mir-221.⁴⁴ Uncovering the general principle of such activities could lead to a more complete understanding of host defense mechanisms. The presence of miR-like sequences in the viral RNA genome will compete with and impair appropriate cellular miR function⁴⁵ if accessibility is not an issue.^{22,38}

We also found that the sets of real plant viruses and randomly created sequences have similarities with human miRs, based on the statistical significance of BLAST's bit-score.

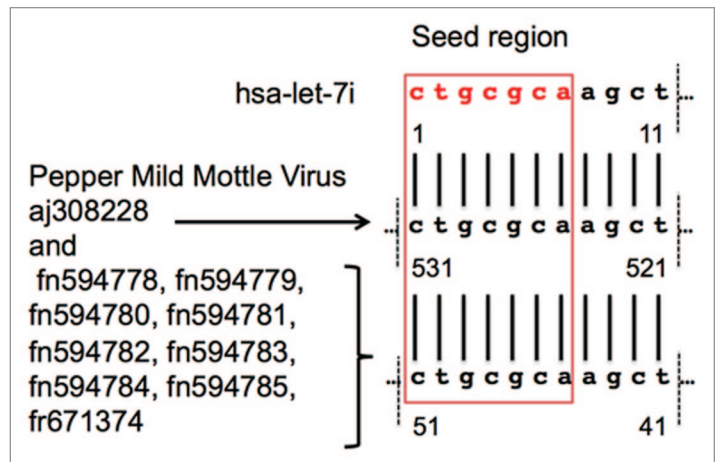


Figure 8. Sequence similarities between *pepper mild mottle virus* and hsa-let-7i. The seed region is inside the solid-line box. The complete sequence similarity between the different strains of the virus and the hsa-mir-let-7i consists of 11 nucleotides. Every single PMMoV is complementary (Plus/Minus) to the miR. The virus strain aj308228 shares the similarity from nucleotide 531, while all the other strains do so from nucleotide 51 (antisense).

Nevertheless and based on p values, we obtained a strong evidence that randomness does not match statistically significant different examples of plant virus sequences that were anti-parallel to hsa-miR seed sequences were found. For example the *soybean dwarf virus* (ab038147) contains a plus/minus sequence and has a 100% match with the seed regions of hsa-mir-103a-1 and hsa-mir-103a-2, while also matching a total of 14 nucleotides with them. For those viruses that contain such anti-parallel or plus/minus sequences, it could be that they too are players in the viral regulatory machinery. Alternatively, if cellular systems are considered, these viral sequences would most likely base pair with the cellular miRs to reduce their activity. The family of hsa-mir-103a has 650 predicted target genes associated with it, including SCN1A, a sodium channel which is responsible for action potential initiation and propagation in neurons, GABRG2, a GABA-A receptor with associated chloride channels and CACNA2D1, a calcium channel gene, also present in neurons. Again, it is not clear that such viral complementary sequences would have access to hsa-miRs or vice versa, but if they did they might have rather widespread effects. It has been shown that noncoding RNA in *herpes virus saimiri* binds to and causes the degradation of human miR-27 resulting in enhanced infection.^{46,47} Alternatively, cellular miRs may bind to complementary viral sequence and thereby regulate its ability to transfect cells. Other studies have suggested interactions between viral sequences and cellular miRs.^{18,28,49,50}

To our knowledge this is the first large-scale demonstration of sequence similarities between plant viruses and human miRs. We found a large number of such similarities, many of which demonstrate that plant virus genomes contain sequences that are identical to human miRs. In addition, some plant viruses contain sequences that were complementary to human miRs. These results may be important from an evolutionary point of view. We propose to use the data reported in our study to support future bench level investigations of functional significance.

Table 4. Selected plant viruses matched sequences

Plant virus index	Plant virus name	Plant virus sequence	miR sequence	hsa miR
ay55269	Tobacco mosaic virus coat protein mRNA, complete cds.	224 cggctgctggat 234	2 cggctgctggat 12	hsa-mir-1250
z29370	Tobacco mosaic virus (Crucifer) genomic RNA for RNA-dependent RNA polymerase; 122K protein; transport protein and coat protein	832 tcaccagccct 822	1 tcaccagccct 11	hsa-mir-1226
am236832	Potato virus Y partial gene for polyprotein, helper component-proteinase, HC-Pro region, genomic RNA, isolate Anhuis	683 actgatacaatt 672	2 actgatacaatt 13	hsa-mir-1243
aj439544	Potato virus Y gene for polyprotein, genomic RNA, isolate SON41	6646 cagttatcacag 6657	1 cagttatcacag 12	hsa-mir-101-1
aj890348	Potato virus Y strain C gene for polyprotein, genomic RNA, isolate Adgen	6757 acggatgtttg 6767	1 acggatgtttg 11	hsa-mir-105-1
aj890348	Potato virus Y strain C gene for polyprotein, genomic RNA, isolate Adgen	6757 acggatgtttg 6767	1 acggatgtttg 11	hsa-mir-105-2
u06789	Potato virus Y-VN genome, 3' untranslated region.	850 gagggctctgg 860	1 gagggctctgg 11	hsa-mir-1182
af395678	Blackeye Cowpea mosaic virus polyprotein gene, partial cds.	1054 gtgagggcatgcag 1067	1 gtgagggcatgcag 14	hsa-mir-1226
aj15903	Blackeye Cowpea mosaic virus partial gene for polyprotein, genomic RNA	629 gtgagggcatgcag 642	1 gtgagggcatgcag 14	hsa-mir-1226
s66280	Polymerase, coat protein [blackeye Cowpea mosaic virus BICMV, W, Genomic, 254 nt, segment 2 of 2].	193 gtgagggcatgcag 206	1 gtgagggcatgcag 14	hsa-mir-1226
x00206	Cowpea mosaic virus bottom component RNA (B RNA)	4830 tcaccagccct 4820	1 tcaccagccct 11	hsa-mir-1226
af395678	Blackeye Cowpea mosaic virus polyprotein gene, partial cds.	267 cagatgacta 257	2 cagatgacta 12	hsa-mir-1245b
s66253	Polymerase, coat protein [blackeye Cowpea mosaic virus BICMV, W, Genomic, 945 nt, segment 1 of 2].	348 cagatgacta 338	2 cagatgacta 12	hsa-mir-1245b
dq198144	Blackeye Cowpea mosaic virus strain Coimbatore coat protein gene, partial cds.	417 tcacactgcc 427	1 tcacactgcc 11	hsa-mir-1228
eu170481	Southern Cowpea mosaic virus coat protein gene, complete cds.	773 gtgttcacagc 763	2 gtgttcacagc 12	hsa-mir-124-1
eu170481	Southern Cowpea mosaic virus coat protein gene, complete cds.	773 gtgttcacagc 763	2 gtgttcacagc 12	hsa-mir-124-2
eu170481	Southern Cowpea mosaic virus coat protein gene, complete cds.	773 gtgttcacagc 763	2 gtgttcacagc 12	hsa-mir-124-3
aj308228	Pepper mild mottle virus proviral gene for 183 kDa protein, gene for 126 kDa protein, gene for 28 kDa protein and cp gene, genomic RNA	531 ctgccaagct 521	1 ctgccaagct 11	hsa-let-7i
fn594778	Pepper mild mottle virus partial 126K gene for replicase small subunit, isolate P96/41, genomic RNA	51 ctgccaagct 41	1 ctgccaagct 11	hsa-let-7i
fn594779	Pepper mild mottle virus partial 126K gene for replicase small subunit, isolate P03/17, genomic RNA	51 ctgccaagct 41	1 ctgccaagct 11	hsa-let-7i
fn594780	Pepper mild mottle virus partial 126K gene for replicase small subunit, isolate P95/25, genomic RNA	51 ctgccaagct 41	1 ctgccaagct 11	hsa-let-7i
fn594781	Pepper mild mottle virus partial 126K gene for replicase small subunit, isolate P95/22, genomic RNA	51 ctgccaagct 41	1 ctgccaagct 11	hsa-let-7i
fn594782	Pepper mild mottle virus partial 126K gene for replicase small subunit, isolate P02/4, genomic RNA	51 ctgccaagct 41	1 ctgccaagct 11	hsa-let-7i
fn594783	Pepper mild mottle virus partial 126K gene for replicase small subunit, isolate P02/5, genomic RNA	51 ctgccaagct 41	1 ctgccaagct 11	hsa-let-7i
fn594784	Pepper mild mottle virus partial 126K gene for replicase small subunit, isolate P84/4, genomic RNA	51 ctgccaagct 41	1 ctgccaagct 11	hsa-let-7i
fn594785	Pepper mild mottle virus partial 126K gene for replicase small subunit, isolate P89/4, genomic RNA	51 ctgccaagct 41	1 ctgccaagct 11	hsa-let-7i
fr671374	Pepper mild mottle virus partial 126K gene for replicase, isolate P95/23, genomic RNA	51 ctgccaagct 41	1 ctgccaagct 11	hsa-let-7i

Selected plant viruses in order of TMV, PVY, CPMV, and PMMoV, showing matching between plant virus and miR sequences.

Materials and Methods

A data processing pipeline was developed for collecting sequences from plant virus genomes and hsa-miRs and for procuring high quality comparative data, as depicted in **Figure 1**.

Step 1

Sequences from two databases were downloaded: (1) descriptions of plant viruses (DPVweb),²⁹ including the complete set of independent files corresponding to plant virus genomes (**Fig. 1**, data set-A; downloaded on April 7th, 2011); and (2) the miRBase⁴⁸ sequence database (version 16, November 2nd, 2011), a repository of published miR sequences and associated annotations. From miRBase, we selected only those sequences for human miRs (hsa-miRs; **Fig. 1**, data set B). The set of plant viruses and the set of hsa-miRs were FAST formatted for alignment with Basic Length Alignment Sequence Tool for nucleotide-nucleotide matching (BLAST+ 2.2.25+), hosted by the National Center for Biotechnology Information (NCBI), and performed the BLASTN in a local server.

Step 2

BLASTN 2.2.25+^{31,32} was used to detect matches that demonstrated sequence similarity between plant viral RNAs and human miRs.³⁰ A sequence of six or more consecutive matching nucleotides was required. The bit-score was used to select the statistically significant subsets. This score is a numerical value that describes the overall quality of an alignment, with higher bit-scores associated with better alignments.⁴⁹

Step 3

All sequence matches were then filtered to select higher quality data. First, we used only those plant viruses with sequence lengths equal to or greater than 1,221 nucleotides to filter out partial viral sequences. In addition, we required any matching sequences to have a minimum of half the length of a mature miR which is 11 nucleotides.³³ A third filtering step required the presence of at least some part of the miR-seed sequence, nucleotides 2–7. 393 sequence similarities were found from filtering by selecting just those similarities that are in common from the second nucleotide in the human miR. The latter sequence is important for complementary interactions of miRs with their gene-targets and is reported to include the second through the

seventh¹¹ nucleotides of the mature miR sequence. In addition, we allowed searches of partial virus genomes larger than 1,221 nucleotides for studies of four specific crop plants (tobacco, potato, pepper and cowpea) for which there were only incomplete viral genomes available. Finally, we performed a statistical analysis on the results. GenRGens⁵⁰ software was utilized to randomly generate genomics sequences. We created randomized versions of every single plant virus genome in the DPV database. The settings used were the Markov model, with order and phase equal to one and to those in our original database. By using the same miRBase filtered human miRs as in our original study, we analyzed similarities to the new randomly generated plant virus-like sequences by the same pipeline. By filtering the results for similarities starting at the second nucleotide of the human miRs, we obtained 9,359 similarities (compared with only 393 found between naturally existing plant viruses and human miRs). We statistically compared both sets of analyses to determine if there were any similarities.

The database

The pipeline fed into a set of curated data that was subsequently used for a variety of additional analyses. For example, a clustering analysis of the information in these data provides an overview of the collected information from either the virus's or miR's point of view. In addition, some filtering was used to identify viral sequences that shared identity with the complete miR seed sequences. Finally, viruses associated with a few agricultural crops of importance to agronomy were examined.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank Ruma Raha-Chowdhury, PhD (University of Cambridge, UK), Yong Li, PhD (University of Louisville, KY), Guy Brock, PhD (University of Louisville) and Shivani Nandi, PhD (University of Louisville, KY) for their helpful discussions.

Supplementary Material

Supplementary material may be found here: <http://www.landesbioscience.com/journals/cib/article/24951/>

References

1. Lee RC, Feinbaum RL, Ambros V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 1993; 75:843-54; PMID:8252621; [http://dx.doi.org/10.1016/0092-8674\(93\)90529-Y](http://dx.doi.org/10.1016/0092-8674(93)90529-Y)
2. Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T. Identification of novel genes coding for small expressed RNAs. *Science* 2001; 294:853-8; PMID:11679670; <http://dx.doi.org/10.1126/science.1064921>
3. Bandi N, Zbinden S, Gugger M, Arnold M, Kocher V, Hasan L, et al. miR-15a and miR-16 are implicated in cell cycle regulation in a Rb-dependent manner and are frequently deleted or down-regulated in non-small cell lung cancer. *Cancer Res* 2009; 69:5553-9; PMID:19549910; <http://dx.doi.org/10.1158/0008-5472.CAN-08-4277>
4. Chang KW, Liu CJ, Chu TH, Cheng HW, Hung PS, Hu WY, et al. Association between high miR-211 microRNA expression and the poor prognosis of oral carcinoma. *J Dent Res* 2008; 87:1063-8; PMID:18946016; <http://dx.doi.org/10.1177/154405910808701116>
5. van Rooij E, Olson EN. MicroRNAs: powerful new regulators of heart disease and provocative therapeutic targets. *J Clin Invest* 2007; 117:2369-76; PMID:17786230; <http://dx.doi.org/10.1172/JCI33099>
6. Visvanathan J, Lee S, Lee B, Lee JW, Lee S-K. The microRNA miR-124 antagonizes the anti-neural REST/SCP1 pathway during embryonic CNS development. *Genes Dev* 2007; 21:744-9; PMID:17403776; <http://dx.doi.org/10.1101/gad.1519107>
7. Ambros V. microRNAs: tiny regulators with great potential. *Cell* 2001; 107:823-6; PMID:11779458; [http://dx.doi.org/10.1016/S0092-8674\(01\)00616-X](http://dx.doi.org/10.1016/S0092-8674(01)00616-X)
8. Hatzigeorgiou AG, Fiziev P, Reczko M. DIANA-EST: a statistical analysis. *Bioinformatics* 2001; 17:913-9; PMID:11673235; <http://dx.doi.org/10.1093/bioinformatics/17.10.913>
9. Rehmsmeier M, Steffen P, Hochsmann M, Giegerich R. Fast and effective prediction of microRNA/target duplexes. *RNA* 2004; 10:1507-17; PMID:15383676; <http://dx.doi.org/10.1261/rna.5248604>
10. Yoon S, De Micheli G. Prediction and Analysis of Human microRNA Regulatory Modules. *Conf Proc IEEE Eng Med Biol Soc* 2005; 5:4799-802; PMID:17281315
11. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 2005; 120:15-20; PMID:15652477; <http://dx.doi.org/10.1016/j.cell.2004.12.035>
12. Krek A, Grün D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, et al. Combinatorial microRNA target predictions. *Nat Genet* 2005; 37:495-500; PMID:15806104; <http://dx.doi.org/10.1038/ng1536>

13. Grün D, Wang Y-L, Langenberger D, Gunsalus KC, Rajewsky N. microRNA target predictions across seven *Drosophila* species and comparison to mammalian targets. *PLoS Comput Biol* 2005; 1:e13-13; PMID:16103902; <http://dx.doi.org/10.1371/journal.pcbi.0010013>
14. Hsu S-D, Chu C-H, Tsou A-P, Chen S-J, Chen H-C, Hsu P-W, et al. miRNAMap 2.0: genomic maps of microRNAs in metazoan genomes. *Nucleic Acids Res* 2008; 36(Database issue):D165-9; PMID:18029362; <http://dx.doi.org/10.1093/nar/gkm1012>
15. Mazière P, Enright AJ. Prediction of microRNA targets. *Drug Discov Today* 2007; 12:452-8; PMID:17532529; <http://dx.doi.org/10.1016/j.drudis.2007.04.002>
16. Yanju Z, de Bruin JS, Verbeek FJ. miRNA target prediction through mining of miRNA relationships; 2008.1-6
17. Barbato C, Arisi I, Frizzo ME, Brandi R, Da Sacco L, Masotti A. Computational challenges in miRNA target predictions: to be or not to be a true target? *J Biomed Biotechnol* 2009; 2009:803069; PMID:19551154; <http://dx.doi.org/10.1155/2009/803069>
18. Kumar S, Ansari FA, Scaria V. Prediction of viral microRNA precursors based on human microRNA precursor sequence and structural features. *Virology* 2009; 6:129; PMID:19691855; <http://dx.doi.org/10.1186/1743-422X-6-129>
19. John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS. Human MicroRNA targets. *PLoS Biol* 2004; 2:e363-363; PMID:15502875; <http://dx.doi.org/10.1371/journal.pbio.0020363>
20. Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP. MicroRNAs in plants. *Genes Dev* 2002; 16:1616-26; PMID:12101121; <http://dx.doi.org/10.1101/gad.1004402>
21. Vaucheret H. Post-transcriptional small RNA pathways in plants: mechanisms and regulations. *Genes Dev* 2006; 20:759-71; PMID:16600909; <http://dx.doi.org/10.1101/gad.1410506>
22. Zhang L, Hou D, Chen X, Li D, Zhu L, Zhang Y, et al. Exogenous plant MIR168a specifically targets mammalian LDLRAP1: evidence of cross-kingdom regulation by microRNA. *Cell Res* 2012; 22:107-26; PMID:21931358; <http://dx.doi.org/10.1038/cr.2011.158>
23. Pérez-Quintero ÁL, Neme R, Zapata A, López C. Plant microRNAs and their role in defense against viruses: a bioinformatics approach. *BMC Plant Biol* 2010; 10:138; PMID:20594353; <http://dx.doi.org/10.1186/1471-2229-10-138>
24. Chen F, Chen YP. Exploring cross-species-related miRNAs based on sequence and secondary structure. *IEEE Trans Biomed Eng* 2010; 57:1547-53; PMID:20199930; <http://dx.doi.org/10.1109/TBME.2010.2043734>
25. Myhre MR, Fenton KA, Eggert J, Nielsen KM, Traavik T. The 35S CaMV plant virus promoter is active in human enterocyte-like cells. *Eur Food Res Technol* 2006; 222:185-93; <http://dx.doi.org/10.1007/s00217-005-0154-3>
26. Shriver LP, Koudelka KJ, Manchester M. Viral nanoparticles associate with regions of inflammation and blood brain barrier disruption during CNS infection. *J Neuroimmunol* 2009; 211:66-72; PMID:19394707; <http://dx.doi.org/10.1016/j.jneuroim.2009.03.015>
27. Colson P, Richet H, Desnues C, Balique F, Moal V, Grob JJ, et al. *Pepper mild mottle virus*, a plant virus associated with specific immune responses, Fever, abdominal pains, and pruritus in humans. *PLoS One* 2010; 5:e10041; PMID:20386604; <http://dx.doi.org/10.1371/journal.pone.0010041>
28. Glazov EA, Horwood PF, Assavalapsakul W, Kongsuwan K, Mitchell RW, Mitter N, et al. Characterization of microRNAs encoded by the bovine herpesvirus 1 genome. *J Gen Virol* 2010; 91:32-41; PMID:19793906; <http://dx.doi.org/10.1099/vir.0.014290-0>
29. Adams MJ, Antoniw JF. DPVweb: a comprehensive database of plant and fungal virus genes and genomes. *Nucleic Acids Res* 2006; 34(Database issue):D382-5; PMID:16381892; <http://dx.doi.org/10.1093/nar/gkj023>
30. Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res* 2006; 34(Database issue):D140-4; PMID:16381832; <http://dx.doi.org/10.1093/nar/gkj112>
31. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990; 215:403-10; PMID:2231712
32. Tatusova TA, Madden TL. BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences. *FEMS Microbiol Lett* 1999; 174:247-50; PMID:10339815; <http://dx.doi.org/10.1111/j.1574-6968.1999.tb13575.x>
33. Lambert NJ, Gu SG, Zahler AM. The conformation of microRNA seed regions in native microRNPs is pre-arranged for presentation to mRNA targets. *Nucleic Acids Res* 2011; 39:4827-35; PMID:21335607; <http://dx.doi.org/10.1093/nar/gkr077>
34. Friedland RP, Tedesco JM, Wilson AC, Atwood CS, Smith MA, Perry G, et al. Antibodies to *Potato virus Y* bind the amyloid beta peptide: immunohistochemical and NMR studies. *J Biol Chem* 2008; 283:22550-6; PMID:18505725; <http://dx.doi.org/10.1074/jbc.M802088200>
35. Pfleger FL, Zeyen RJ. (2008) *Tomato-Tobacco mosaic virus* Disease. Extension. 01168 ed. Minnesota: University of Minnesota.
36. Balique F, Colson P, Raoult D. *Tobacco mosaic virus* in cigarettes and saliva of smokers. *J Clin Virol* 2012; 55:374-6; PMID:22959216; <http://dx.doi.org/10.1016/j.jcv.2012.08.012>
37. Liu R, Vaishnav RA, Roberts AM, Friedland RP. Humans Have Antibodies against a Plant Virus: Evidence from *Tobacco mosaic virus*. *PLoS One* 2013; 8:e60621; PMID:23573274; <http://dx.doi.org/10.1371/journal.pone.0060621>
38. Koudelka KJ, Destito G, Plummer EM, Trauger SA, Siuzdak G, Manchester M. Endothelial targeting of *Coupea mosaic virus* (CPMV) via surface vimentin. *PLoS Pathog* 2009; 5:e1000417; PMID:19412526; <http://dx.doi.org/10.1371/journal.ppat.1000417>
39. Wang F-Z, Weber F, Croce C, Liu C-G, Liao X, Pellett PE. Human cytomegalovirus infection alters the expression of cellular microRNA species that affect its replication. *J Virol* 2008; 82:9065-74; PMID:18596100; <http://dx.doi.org/10.1128/JVI.00961-08>
40. Umbach JL, Cullen BR. The role of RNAi and microRNAs in animal virus replication and antiviral immunity. *Genes Dev* 2009; 23:1151-64; PMID:19451215; <http://dx.doi.org/10.1101/gad.1793309>
41. Pfeffer S, Sewer A, Lagos-Quintana M, Sheridan R, Sander C, Grässer FA, et al. Identification of microRNAs of the herpesvirus family. *Nat Methods* 2005; 2:269-76; PMID:15782219; <http://dx.doi.org/10.1038/nmeth746>
42. Watanabe Y, Tomita M, Kanai A. (2007) Computational Methods for MicroRNA Target Prediction. In: John JR, Gregory JH, editors. *Methods in Enzymology*: Academic Press:65-86.
43. Watanabe Y, Kishi A, Yachie N, Kanai A, Tomita M. Computational analysis of microRNA-mediated antiviral defense in humans. *FEBS Lett* 2007; 581:4603-10; PMID:17825824; <http://dx.doi.org/10.1016/j.febslet.2007.08.049>
44. Othumpangat S, Walton C, Piedimonte G. MicroRNA-221 modulates RSV replication in human bronchial epithelium by targeting NGF expression. *PLoS One* 2012; 7:e30030; PMID:22272270; <http://dx.doi.org/10.1371/journal.pone.0030030>
45. Salmena L, Polisenio L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell* 2011; 146:353-8; PMID:21802130; <http://dx.doi.org/10.1016/j.cell.2011.07.014>
46. Cazalla D, Yario T, Steitz JA. Down-regulation of a host microRNA by a Herpesvirus saimiri noncoding RNA. *Science* 2010; 328:1563-6; PMID:20558719; <http://dx.doi.org/10.1126/science.1187197>
47. Libri V, Helwak A, Miesen P, Santhakumar D, Borger JG, Kudla G, et al. Murine cytomegalovirus encodes a miR-27 inhibitor disguised as a target. *Proc Natl Acad Sci U S A* 2012; 109:279-84; PMID:22184245; <http://dx.doi.org/10.1073/pnas.1114204109>
48. Kozomara A, Griffiths-Jones S. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res* 2011; 39(Database issue):D152-7; PMID:21037258; <http://dx.doi.org/10.1093/nar/gkq1027>
49. Karlin S, Altschul SF. Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proc Natl Acad Sci U S A* 1990; 87:2264-8; PMID:2315319; <http://dx.doi.org/10.1073/pnas.87.6.2264>
50. Ponty Y, Termier M, Denise A. GenRGenS: software for generating random genomic sequences and structures. *Bioinformatics* 2006; 22:1534-5; PMID:16574695; <http://dx.doi.org/10.1093/bioinformatics/btl113>