Identification of host encoded microRNAs interacting with novel swine-origin influenza A (H1N1) virus and swine influenza virus

Tao He, Guihai Feng, Huipeng Chen, Li Wang, Yumin Wang*

Beijing Institute of Biotechnology, Beijing 100071, China; Yumin Wang – Email: wym66@vip.sina.com; *Corresponding author

Abstract:
The discovery of microRNAs (miRNAs) is a remarkable breakthrough in the field of life science, and they are important actors which regulate gene expression in diverse cellular processes. Recently, several reports indicated that miRNAs can also target viruses and regulate virus replication. Here we discovered 36 pig-encoded miRNAs and 22 human-encoded miRNAs which have putative targets in swine influenza virus (SIV) and Swine-Origin 2009 A/H1N1 influenza virus (S-OIV) genes respectively. Interestingly, the putative interactions of ssc-miR-124a, ssc-miR-136 and ssc-miR-145 with their SIV target genes had been found to be maintained almost throughout all of the virus evolution. Enrichment analysis of previously reported miRNA gene expression profiles revealed that three miRNAs are expressed at higher levels in human lung or trachea tissue. The hsa-miR-145 and hsa-miR-92a putatively target the HA gene and hsa-miR-150 putatively targets the PB2 gene. Analysis results based on the location distribution from which virus was isolated and sequence conservation imply that some putative miRNA-mediated host-virus interactions may characterize the location-specificity.

Keywords: microRNAs (miRNAs), swine influenza virus, target prediction, virus-host interaction

Background:
MicroRNAs (miRNAs) are a class of ~22 nucleotide (nt) evolutionary conserved endogenous non-coding RNAs and believed to play important roles in gene regulation by either inactivating or degrading mRNA genes [1]. Up to date, more than 9,000 miRNAs have been reported in animals, plants and viruses by using computational and biological skills. They have been reported to participate in the regulation of almost every cellular process [2]. Some evidences support that miRNAs play an important role in the intricate host-virus interaction network. The early researches suggested that some plant miRNAs are involved in plant defense and viral offense systems [3]. Later, Lecellier *et al.* [4] reported that human miR-32 could inhibit retrovirus primate foamy virus (PFV) replication; Jopling *et al.* [5] presented that the liver-specific miRNA miR-122 could enhance the replication of hepatitis C virus (HCV); Murakami *et al.* [6] found that miR-199a could inhibit HCV genome replication; Hariharan *et al.* [7] predicted candidate targets of five human miRNAs in human immunodeficiency virus (HIV); Ahluwalia *et al.* [8] conformed that human miR-29a could down-regulate the expression of Nef protein and interfere with HIV-1 replication; And Hakim et al. [9] observed that many human microRNAs can target lentiviruses and retroviruses. On the other hand, viruses have also evolved to escape host's RNAi machinery by a variety of strategies [10-13]. Though studies into the microRNA mediated host-influenza virus interactions are still in its infancy, some reports have suggested that miRNAs may be involved in anti-flu viral response. Ge *et al.* [14, 15] showed that RNA interfering (RNAi) could decrease influenza virus production in mouse. Matskevich *et al.* [16] indicated that human Dicer, an enzyme essential for miRNA processing, might inhibit the influenza A virus replication. An unpublished data mentioned that two human encoded miRNAs miR-136 and miR-507 could target HA and PB2 genes in influenza A (H5N1) virus [17].

Swine-Origin 2009 A/H1N1 influenza virus (S-OIV) is a newly emerged virus which transmitted abroad quickly. More than 10,000 cases from 56 countries have been diagnosed as 2009 A(H1N1) influenza in the world within 2 months. It trends to be thought that this S-OIV are homogeneous and similar to North American swine A(H1N1) viruses [18]. Traditionally, swine influenza viruses (SIV) could cause the swine flu among pigs. But there has been no evidence of the S-OIV transmission among the swine up to date. Here we reported the discovery of pig miRNAs and human miRNAs which putatively target in SIV and S-OIV genes respectively. In all of 36 *Sus scrofa* miRNAs predicted to target specific SIV genes, the putative interactions of ssc-miR-124a, ssc-miR-136 and ssc-miR-145 with their SIV target genes had been found to be maintained throughout virus evolution. In all of 22 *Homo sapiens* miRNAs predicted to target specific S-OIV genes, hsa-miR-145 and hsa-miR-92a putatively target the HA gene and hsa-miR-150 putatively targets the PB2 gene. Analysis of previously reported gene expression profiles of miRNAs revealed that these three miRNAs are expressed at higher levels in human lung or trachea tissue. Analysis results based on the location distribution from which virus was isolated and sequence conservation imply that some putative miRNA-mediated host-virus interactions may characterize the location-specificity.

Methodology:
Figure 1 depicts the flowchart in the present study. And the following description is about the explanation for all the data and tools mentioned in this figure:

Datasets:
885 Homo sapiens and 74 Sus scrofa microRNA mature sequences were downloaded from the miRbase database (http://microrna.sanger.ac.uk/sequences/, Release 13.0) [19]. The Swine-Origin 2009 A/H1N1 influenza virus and the swine influenza virus sequences with annotations in GenBank flat file format were downloaded from the NCBI Influenza Virus Sequence Database (http://www.ncbi.nlm.nih.gov/genomes/FLU/) [20].

MicroRNA expression enrichment analysis:
Because the respiratory tract is the main organ through that influenza virus invades, replicate and spread , all human miRNAs expressed in the respiratory tract including oral cavity, larynx, pharynx, trachea, bronchus and lung tissue were collected from experimental data published in literatures [21, 22] and database [23, 24]. Data normalized by Z score transformation can be used directly in the calculation of significant changes in gene expression between different samples and conditions [25]. We used Z scores (Equation 1 in supplementary material) to compare the expression levels of the same miRNA in different tissues, and then sorted all microRNAs expressed in a specific tissue by their expression levels.
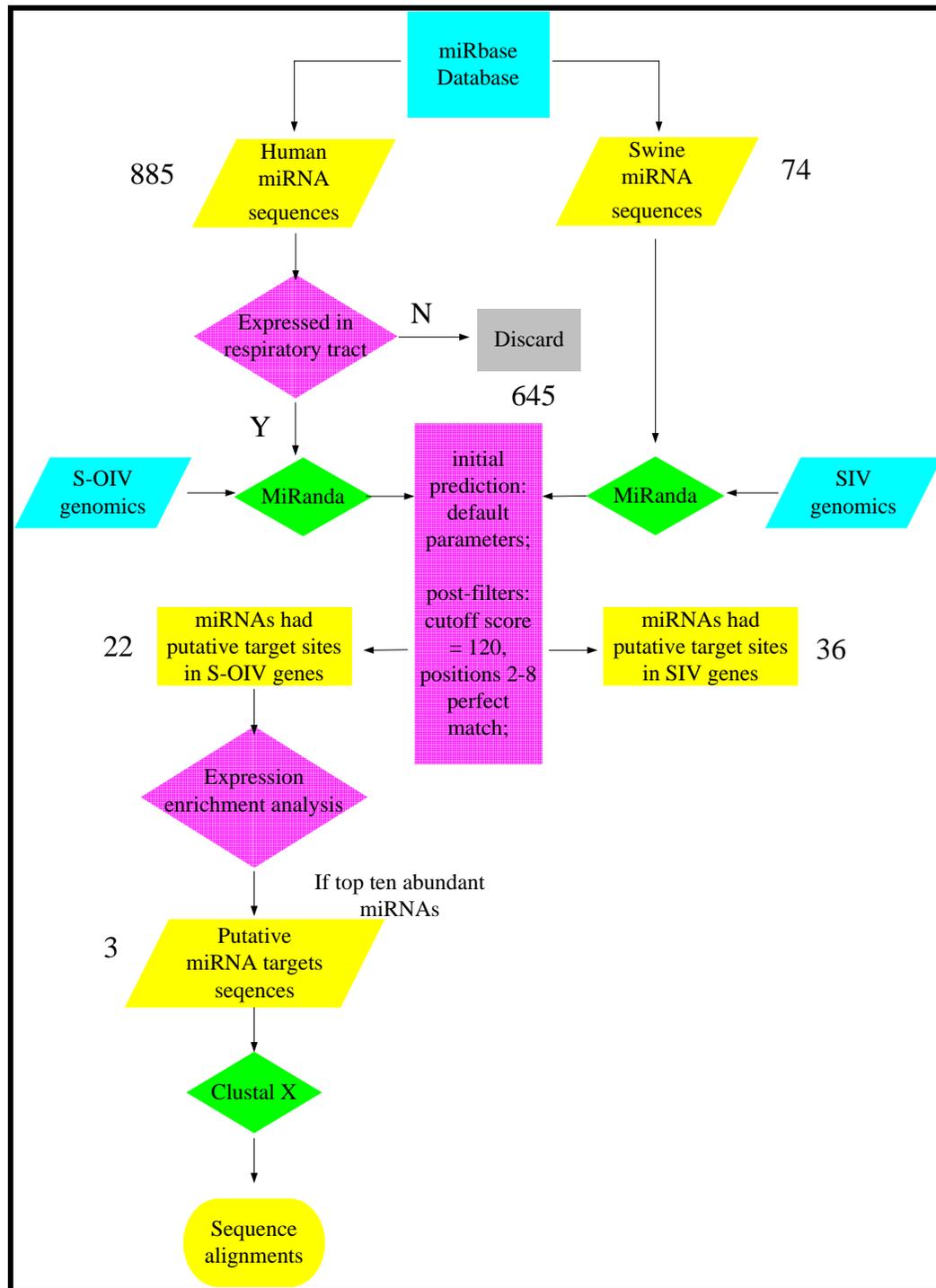
Figure 1: A systematic flowchart in the present study. All the data, software or databases are described in Methodology or Supplementary material.

MicroRNA target prediction:

MiRanda is a dynamic programming global alignment algorithm, that places double score on bases 2 - 8 from the 5' end of the miRNA [26]. We used this software to initially predict miRNA targets with default parameters (Gap Open Penalty: -8.0; Gap Extend:-2.0; Score Threshold: 50.0; Energy Threshold: -20.0 kcal/mol; Scaling Parameter: 2.0). After this initial matching, several post-filters were applied: 1. a cutoff score of 120 was used according to ref. 7; 2. an exact matching to 5'end seed region (positions 2-8) of the mature miRNA was used and the G:U base pairing was not allowed. 3. Only the miRNAs which are expressed in the human respiratory tract were considered. The third criterion is not applied in target identification in SIV genes due to the lack of the pig miRNA expression profile data.
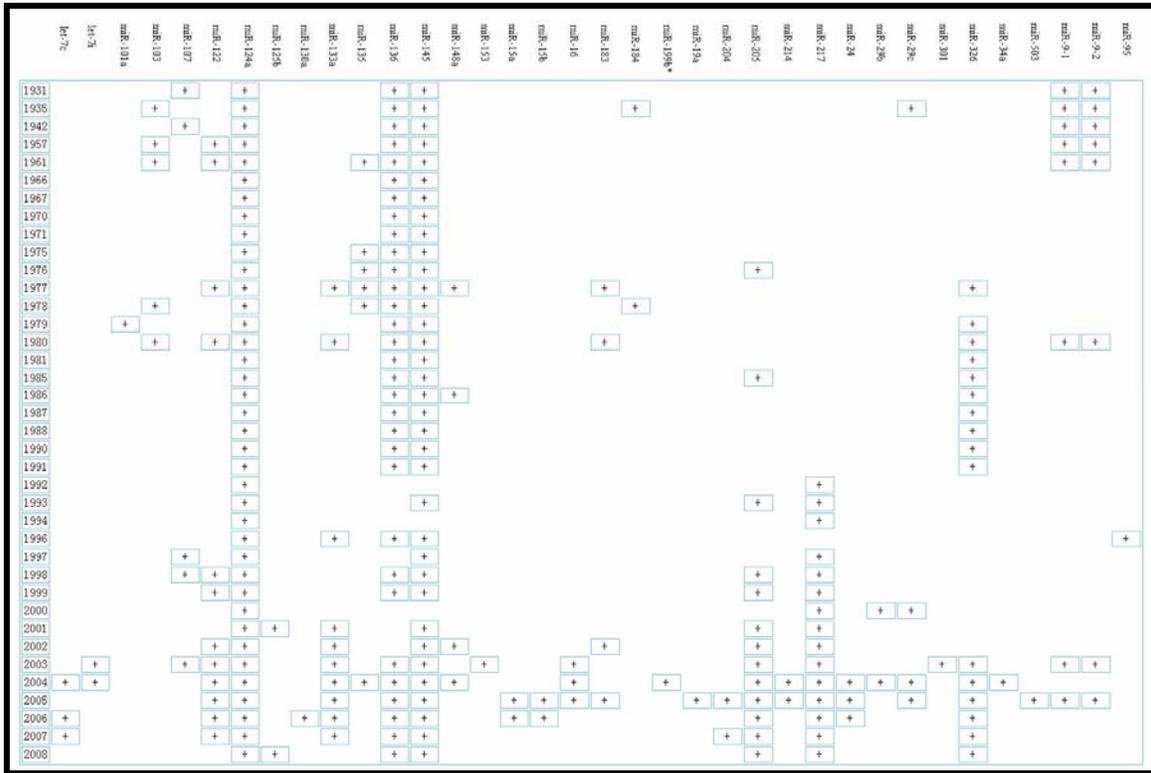
Figure 2: The putative interactions between swine influenza viruses isolated in different times with the pig miRNAs. The first row shows pig miRNAs with putative target sites in SIV genes. The first column shows the different 38 times in which swine influenza viruses were isolated (ranging in years from 1931 to 2008).

Figure 3: Heat maps of Z-score-normalized microRNA expression profiles in lung and trachea. The Table 3 in supplemental material contains miRNA name and corresponding Z scores with the same rank order in this figure.

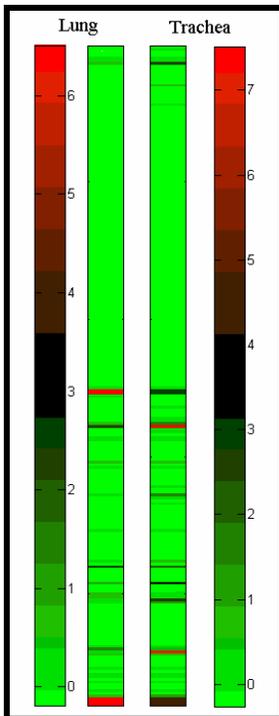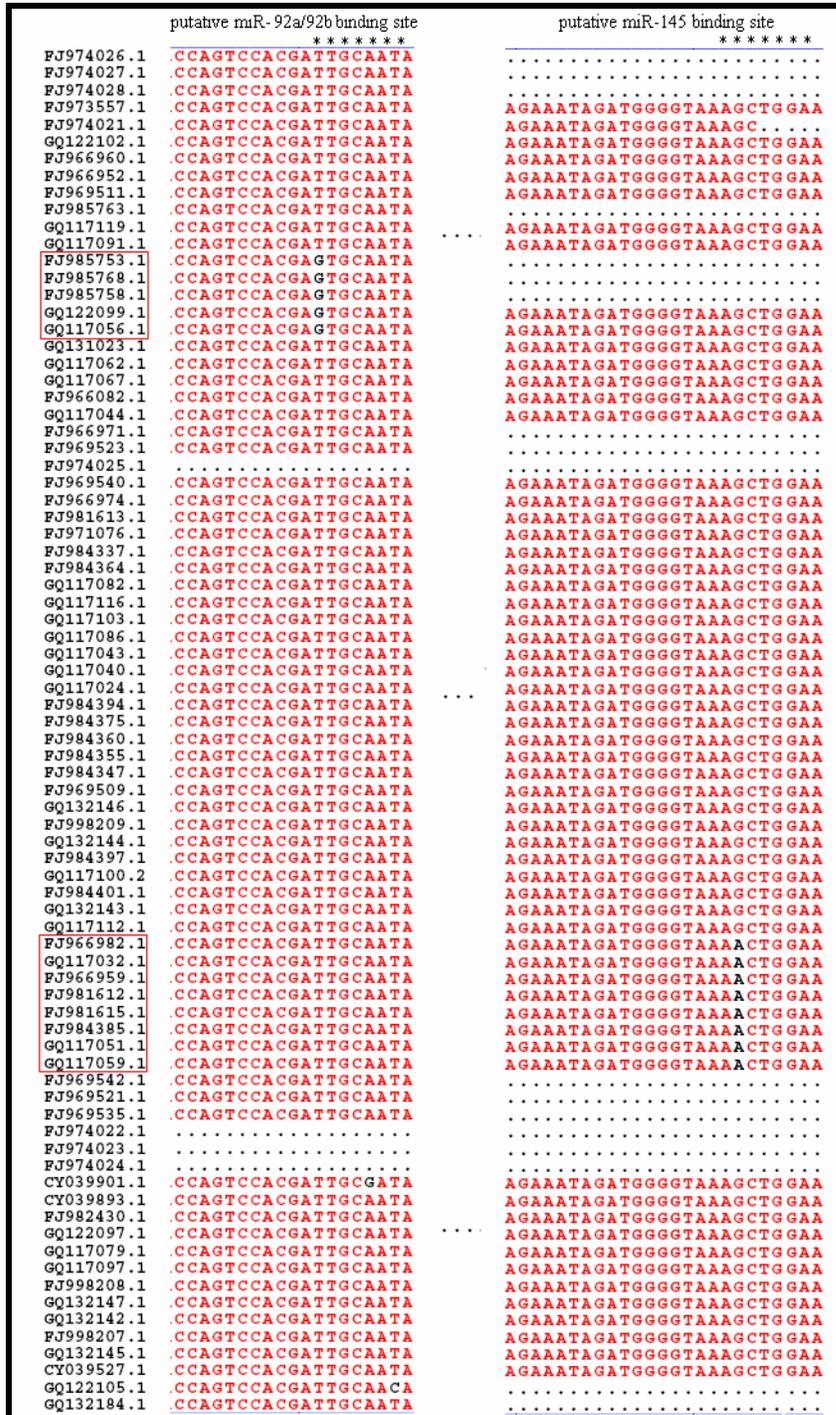**Figure 4:** Sequence alignments of the S-OIV HA genes derived from different isolated sources. The T to G, A to G and T to C mutations were identified in miR-92a/92b target region and a G to A mutation was identified in miR-145 target region.

Sequence conservation analysis:

The sequence segments of influenza viruses isolated from different patients were aligned using Clustal X [27]. The conservation of target sites for the respective miRNAs was analyzed based on these alignment results and the mutations in target sites were identified at the same time.

Discussion:

Pig microRNAs target in swine influenza virus:

The top scoring miRNA-target pairs were identified by miRanda software with a perfect heptamer match in positions 2-8 of the mature miRNA. We found that 36 *Sus scrofa* miRNAs had putative target sites in SIV genes (Figure 2 and Table 1 in supplementary material). We collected viral sequences isolated in different 38 years, ranging in years from 1931 to 2008. Three putative pig miRNA-SIV

interactions were found to be maintained almost throughout all the swine virus evolution (Figure 2). For example, ssc-miR-124a targets the NP genes isolated in all different 38 times, ssc-miR-145 targets the NP genes isolated in different 35 times of 38 times and ssc-miR-136 targets the NA and NP genes isolated in different 31 times of 38 times. The influenza A virus nucleoprotein (NP) is a single-stranded RNA-binding protein that encapsidates the virus genome and has essential functions in viral-RNA synthesis. Moreover, the sequences of the NP gene are highly conserved across subtypes of influenza virus. There have been reported that small interfering RNAs (siRNAs) specific to the nucleocapsid protein (NP) gene can inhibit influenza virus production [14, 15].

Human microRNAs target in Swine-Origin 2009 A/H1N1 influenza virus:

We used miRanda for identifying human microRNA target sites in S-OIV gene sequences with a perfect heptamer match in positions 2-8 of the mature miRNA. The miRNAs can act as effectors in the intricate host-virus interaction networks based on the premise that miRNAs expression and virus replication should happen at the same cell. Because influenza virus invades , replicates and spreads at the respiratory tract in humans, we filtered out miRNAs without evidence for expressing in human respiratory tract. And then, we identified 22 miRNAs with putative target sites in S-OIV genes. The human homologs of ssc-miR-145 and ssc-miR-136, hsa-miR-145 have putative target sites in S-OIV HA and NP gene whereas hsa-miR-136 also has putative target sites in S-OIV NP and NA genes (Table 2 in supplementary material). The imprinted miRNA miR-136 fully complementary to a retrotransposon-like gene Rtl1 in mouse may support this miRNA can also target exogenous nucleic acids, as in the case of viruses [28].

Based on Z-score analysis of previously reported gene expression profiles of miRNAs [23], three miRNAs with putatively target sites on S-OIV genes are expressed at higher levels (top ten significantly abundant miRNAs) in human lung or trachea tissue among all 240 previously studied miRNAs (Figure 3). The miR-145 and miR-92a putatively target the HA gene whereas miR-150 putatively targets the PB2 gene. HA is known to be critical for pathogenicity of influenza virus and immuno-suppression and PB2 is one component of ribonucleoprotein which is responsible for virus RNA replication and transcription [29]. Analysis results based on the isolated virus location distribution and sequence conservation imply that some putative miRNA-mediated host-virus interactions characterize the location-specificity. For example, putative regulation mediated by miR-145 may focus on the virus isolated from Texas State in the United States of America and putative regulations mediated by miR-92a/92b may focus on virus isolated from the location of Spain (**Table 1 in supplementary material**). Conservation analysis of putative miRNA target sites shows that a T to G mutation in miR-92a/92b target region and a G to A mutation in miR-145 target region affect the perfect match to the seed region of corresponding miRNAs (Figure 4).

Conclusion:

At present, our knowledge about the miRNAs' role in viral infections is still in its infancy. Genome-wide prediction of miRNA-virus interaction by computational skills may present some important clues for anti-virus defense. In summary, our study implies that human and pig miRNAs have the potential to affect the expression of S-OIV and SIV genes. Whether or not any of these interactions mediated by the host cellular miRNAs to enhance virus replication, or instead act to inhibit virus replication, is still unclear. Further studies focusing on these interactions are needed to reveal the interplay mechanism of host miRNAs and viruses, and the conserved interaction between miRNAs and S-OIV or SIV genes may be potential drug targets for anti-S-OIV or anti-SIV therapy.

Supplementary material:
The detailed supplementary materials are given at http://www.compbio.net.cn/siv.htm.

References:
[1] DP Bartel *Cell* **116**: 281 (2004) [PMID: 14744438]
[2] W Filipowicz *et al., Nat Rev Genet.* **9**:102 (2008) [PMID: 18197166]
[3] YD Lu *et al., Plant Cell Rep.* **27**:1571 (2008) [PMID: 18626646]
[4] CH Lecellier *et al., Science* **308**:557 (2005) [PMID: 15845854]
[5] CL Jopling *et al., Science* **309**:1577 (2005) [PMID: 16141076]
[6] Y Murakami *et al., J Hepatol.* **50**:453 (2009) [PMID: 19144437]
[7] M Hariharan *et al., Biochem Biophys Res Commun.* **337**:1214 (2005) [PMID: 16236258]
[8] JK Ahluwalia *et al., Retrovirology* **5**:117 (2008) [PMID: 19102781]
[9] ST Hakim *et al., Biochem Biophys Res Commun.* **369**:357 (2008) [PMID: 18282469]
[10] D Silhavy *et al., Embo J.* **21**:3070 (2002) [PMID: 12065420]
[11] E Bucher *et al., J Gen Virol.* **85**:983 (2004) [PMID: 15039540]
[12] S Lu, BR Cullen, *J Virol.* **78**:12868 (2004) [PMID: 15542639]
[13] R Triboulet *et al., Science* **315**:1579 (2007) [PMID: 17322031]
[14] Q Ge *et al., Proc Natl Acad Sci U S A.* **100**:2718 (2003) [PMID: 12594334]
[15] Q Ge *et al., Proc Natl Acad Sci U S A.* **101**:8676 (2004) [PMID: 15173599]
[16] AA Matskevich, K Moelling, *J Gen Virol.* **88**:2627 (2007) [PMID: 17872512]
[17] V Scaria *et al., Retrovirology.* **3**:68 (2006) [PMID: 17032463]
[18] RJ Garten *et al., Science.* **325**:197 (2009) [PMID: 19465683]
[19] S Griffiths-Jones *et al., Nucleic Acids Res.* **34**:D140 (2006) [PMID: 16381832]
[20] Y Bao *et al., J Virol.* **82**:596 (2008) [PMID: 17942553]
[21] LF Sempere *et al., Genome Biol.* **5**:R13 (2004) [PMID: 15003116]
[22] F Schembri *et al., Proc Natl Acad Sci U S A.* **106**: 2319 (2009) [PMID: 19168627]
[23] PW Hsu *et al., Nucleic Acids Res.* **34**:D135 (2006) [PMID: 16381831]
[24] X Wang *Rna* **14**:1012 (2008) [PMID: 18426918]
[25] C Cheadle *et al., J Mol Diagn.* **5**:73 (2003) [PMID: 12707371]
[26] B John *et al., PLoS Biol.* **2**:e363 (2004) [PMID: 15502875]
[27] MA Larkin *et al., Bioinformatics* **23**:2947 (2007) [PMID: 17846036]
[28] H Seitz *et al., Nat Genet.* **34**:261 (2003) [PMID: 12796779]
[29] NM Bouvier, P Palese, *Vaccine* **26**:D49 (2008) [PMID: 19230160]

Edited by P. Kangueane
Citation: He et al., Bioinformation 4(3): 112-118 (2009)

Supplementary material

$$Z\ score = \frac{X_m - \overline{X_M}}{\delta_M}$$

(Equation 1)

Where X_m is the expression value of any miRNA, $\overline{X_M}$ and δ_M are the mean and standard deviation of all of the measured miRNA expression values, respectively.

Table 1: the putative interactions between S-OIV HA genes derived from different isolated sources with the human miRNAs.

GenBank ID	Date	Country	hsa-miR-145	hsa-miR-92b	hsa-miR-92a	hsa-miR-339-5p
FJ966952	30-Mar-09	USA: California state				+
FJ966082	1-Apr-09	USA: California				+
GQ117044	1-Apr-09	USA: California state				+
FJ969511	8-Apr-09	USA: California state				+
FJ966974	9-Apr-09	USA: California state				+
FJ969540	9-Apr-09	USA: California state				+
FJ971076	9-Apr-09	USA: California state				+
FJ981613	9-Apr-09	USA: California state				+
FJ966982	14-Apr-09	USA: Texas state	+			+
FJ981612	14-Apr-09	USA: Texas state	+			+
FJ981615	14-Apr-09	USA: Texas state	+			+
FJ966959	15-Apr-09	USA: Texas state	+			+
FJ966971	15-Apr-09	USA: California state				
GQ122097	15-Apr-09	USA: Texas state				+
FJ966960	16-Apr-09	USA: California state				+
GQ117067	22-Apr-09	USA: Arizona state				+
GQ117097	22-Apr-09	USA: Indiana state				+
FJ984385	23-Apr-09	USA: Texas state	+			+
FJ969521	24-Apr-09	USA: Ohio state				
FJ969523	24-Apr-09	USA: Kansas state				
FJ969535	24-Apr-09	USA: Ohio state				
FJ984360	24-Apr-09	USA: New York state				+
FJ984364	24-Apr-09	USA: New York state				+
FJ984397	24-Apr-09	USA: Ohio state				+
FJ984401	24-Apr-09	USA: Ohio state				+
GQ117024	24-Apr-09	USA: New York state				+
GQ117051	24-Apr-09	USA: Texas state	+			+
GQ117059	24-Apr-09	USA: Kansas state	+			+
GQ117062	24-Apr-09	USA: Kansas state				+
GQ117100	24-Apr-09	USA: Ohio state				+
FJ973557	25-Apr-09	New Zealand: Auckland				+
FJ984347	25-Apr-09	USA: New York state				+
FJ984355	25-Apr-09	USA: New York state				+
FJ984375	25-Apr-09	USA: New York state				+
FJ984394	25-Apr-09	USA: New York state				+
FJ984337	25-Apr-09	USA: New York state				+
GQ117032	25-Apr-09	USA: Texas state	+			+
GQ117040	25-Apr-09	USA: California state				+
GQ117082	25-Apr-09	USA: New York state				+
GQ117086	25-Apr-09	USA: New York state				+
GQ117091	25-Apr-09	USA: Texas state				+
GQ117103	25-Apr-09	USA: Massachusetts state				+
GQ117116	25-Apr-09	USA: New York state				+
CY039893	26-Apr-09	USA: Orange County, NY				+
FJ985753	26-Apr-09	Spain: Albacete, Castilla-La Mancha		+	+	
FJ985758	26-Apr-09	Spain: Valencia		+	+	
FJ985763	26-Apr-09	Spain: Donostia, Pais Vasco				
FJ985768	26-Apr-09	Spain: Albacete, Castilla-La Mancha		+	+	
GQ117043	26-Apr-09	USA: Massachusetts state				+
GQ117056	26-Apr-09	USA: South Carolina state		+	+	+
GQ117079	26-Apr-09	USA: Arizona state				+
GQ117112	26-Apr-09	USA: Michigan state				+
FJ974021	27-Apr-09	Germany: Bavaria, Regensburg				+

GenBank ID	Date	Location			
CY039901	27-Apr-09	USA: Suffolk County, NY			+
GQ117119	27-Apr-09	USA: Colorado state			+
CY039527	29-Apr-09	Netherlands: Rotterdam			+
FJ982430	29-Apr-09	Denmark			+
GQ122105	30-Apr-09	Sweden: Stockholms Lan, Taby Kommun			+
GQ131023	2-May-09	South Korea			+
FJ974022	9-Apr	Canada			
FJ974023	9-Apr	Canada			
FJ974024	9-Apr	Canada			
FJ974025	9-Apr	Canada			
FJ974026	9-Apr	Canada			
FJ974027	9-Apr	Canada			
FJ974028	9-Apr	Canada			
GQ122099	9-Apr	Spain: Catalonia	+	+	+
GQ122102	9-Apr	Spain: Catalonia			+
FJ969509	2009	USA: New York state			+
FJ969542	2009	USA: New York state			
FJ998207	2009	Canada: Nova Scotia			+
FJ998208	2009	Mexico			+
FJ998209	2009	Canada: Ontario			+
GQ132142	2009	Canada: Nova Scotia			+
GQ132143	2009	Canada: Ontario			+
GQ132144	2009	Canada: Ontario			+
GQ132145	2009	Mexico			+
GQ132146	2009	Canada: Alberta			+
GQ132147	2009	Canada: Nova Scotia			+
GQ132184	2009	Thailand			+

+: this HA gene is putatively targeted by the corresponding miRNAs; the GenBank ID highlighted in grey indicates that this gene is partly sequenced and submitted to NCBI; the miRNA highlighted in yellow indicates that this putative miRNA-mediated host-virus interaction shows location-specificity.