

## Identification of Chicken Pulmonary miRNAs Targeting PB1, PB1-F2, and N40 Genes of Highly Pathogenic Avian Influenza Virus H5N1 In Silico

Amod Kumar<sup>1</sup>, Muhasin Asaf V.N.<sup>1</sup>, Ashwin Ashok Raut<sup>2</sup>, Richa Sood<sup>2</sup> and Anamika Mishra<sup>2</sup>

<sup>1</sup>Division of Animal Genetics, Indian Veterinary Research Institute, Bareilly, India. <sup>2</sup>High Security Animal Disease Laboratory, Bhopal, India.

**ABSTRACT:** Highly pathogenic Avian influenza (HPAI) is a notifiable viral disease caused by avian influenza type A viruses of the Orthomyxoviridae family. Type A influenza genome consists of eight segments of negative-sense RNA. RNA segment 2 encodes three proteins, PB1, PB1-F2, and N40, which are translated from the same mRNA by ribosomal leaky scanning and reinitiation. Since these proteins are critical for viral replication and pathogenesis, targeting their expression can be one of the approaches to control and resist HPAI. MicroRNAs are short noncoding RNAs that regulate a variety of biological processes such as cell growth, tissue differentiation, apoptosis, and viral infection. In this study, a set of 300 miRNAs expressed in chicken lungs were screened against the HPAI virus (H5N1) segment 2 with different screening parameter like thermodynamic stability of heteroduplex, seed sequence complementarity, conserved target sequence, and target-site accessibility for identifying miRNAs that can potentially target the transcript of segment 2 of H5N1. Chicken miRNAs *gga-mir-133c*, *gga-mir-1710*, and *gga-mir-146c\** are predicted to target the expression of PB1, PB1-F2, and N40 proteins. This indicates that chicken has genetic potential to resist/tolerate H5N1 infection and these can be suitably exploited in designing strategies for control of avian influenza in chicken.

**KEYWORDS:** chicken, miRNA, H5N1, PB1, PB1-F2, N40

**CITATION:** Kumar et al. Identification of Chicken Pulmonary miRNAs Targeting PB1, PB1-F2, and N40 Genes of Highly Pathogenic Avian Influenza Virus H5N1 In Silico. *Bioinformatics and Biology Insights* 2014;8:135–145 doi: 10.4137/BBI.S14631

**RECEIVED:** February 1, 2014. **RESUBMITTED:** April 21, 2014. **ACCEPTED FOR PUBLICATION:** April 23, 2014.

**ACADEMIC EDITOR:** JT Efrid, Associate Editor

**TYPE:** Original Research

**FUNDING:** The authors thank Indian Council of Agricultural Research (ICAR) for financial assistance in the form of a Junior Research Fellowship during the course of this work.

**COMPETING INTERESTS:** Authors disclose no potential conflicts of interest.

**COPYRIGHT:** © the authors, publisher and licensee Libertas Academica Limited. This is an open-access article distributed under the terms of the Creative Commons CC-BY-NC 3.0 License.

**CORRESPONDENCE:** reach2anamika@yahoo.com

This paper was subject to independent, expert peer review by a minimum of two blind peer reviewers. All editorial decisions were made by the independent academic editor. All authors have provided signed confirmation of their compliance with ethical and legal obligations including (but not limited to) use of any copyrighted material, compliance with ICMJE authorship and competing interests disclosure guidelines and, where applicable, compliance with legal and ethical guidelines on human and animal research participants.

### Introduction

Highly pathogenic avian influenza (HPAI) is a notifiable viral disease caused by avian influenza type A viruses of the Orthomyxoviridae family. It affects poultry in which mortality may reach up to 100%.<sup>1,2</sup> This causes significant economic losses to the poultry industry and raises a great public health threat due to potential host jump from animals to humans.<sup>3</sup> Despite stringent control measures in the form of mass culling of poultry in and around the site of outbreak, re-emergence of avian influenza in poultry is frequently observed. Vaccination is unable to provide sufficient protection against HPAI virus (H5N1) of subsequent outbreaks, due to the characteristic genetic shift and drift seen in this virus.<sup>3</sup> In view of the fact that poultry lines are developed by several generations of selection and planned breeding to meet the production requirements, protecting such

valuable line against avian influenza is the need of the hour. Several anti-influenza drugs are used in case of humans; however, therapy is not recommended in poultry due to threat to effectiveness of the same antivirals as therapeutics for human influenza.<sup>4</sup> Development of influenza-resistant chicken lines is advocated as a possible measure to control avian influenza in poultry. Although transgenesis is reported in poultry, the expression of transgene may not be consistent over generations. Moreover, since chicken is a popular protein source to the human population, genetic modification of chicken would raise several issues pertaining to genetically modified foods, etc. It appears judicious to first screen the chicken genome for possible endogenous products that can confer protection or resistance to H5N1 infection so that further strategies can be developed to exploit them without introduction of exogenous genes.

MicroRNAs (miRNA), that are short, endogenous, noncoding RNAs, regulate a variety of biological processes such as cell growth, tissue differentiation, apoptosis, and viral infection by sequence-specific targeting of mRNAs.<sup>5,6</sup> They act at sequence level to regulate gene expression, either by degrading mRNA of target protein or by causing translational repression of target protein by inhibiting mRNA of target protein, thus act as master switches in many biological pathways.<sup>7-9</sup>

Recently, miR-181 has been identified as a positive regulator of immune response, and miR-181 mimics have been shown to strongly hinder porcine reproductive and respiratory syndrome virus replication in vitro by specifically binding to a highly (over 96%) conserved region in the downstream of open reading frame (ORF) 4 of the viral genomic RNA.<sup>10</sup> miR-181b is reported to inhibit the replication of mink enteritis virus (MEV) in the feline kidney (F18) cell line by targeting the MEV non-structural protein 1 (NS1) mRNA in the coding region.<sup>11</sup> The H5N1 genome contains eight segments of negative-sense RNA (13.5 kb) of which the segment 2 encodes three proteins, PB1, PB1-F2, and N40, which are translated from the same mRNA by ribosomal leaky scanning and reinitiation.<sup>12</sup> The sequence-specific targeting of this segment can result in the downregulation of three viral proteins and affect the viral replication in host cells. Strategies such as use of candidate miRNAs or the use of the mimics of these miRNAs can be used for providing protection to chicken against avian influenza. The lungs are a major site of replication of H5N1 in chicken.<sup>13</sup> In this study, we screened a set of 300 miRNAs expressed in lungs of chicken, with an intent to identify endogenously expressed chicken miRNAs that can potentially target the segment 2 of H5N1.

## Materials and Methods

**Selection of target sequence.** The sequence of segment 2 of influenza A virus H5N1 (A/chicken/India/CA0301/2011(H5N1)) having accession number CY092122 was selected as the target sequence for the study, because this was a complete sequence of segment 2 of H5N1 and had the codon sequence for PB1, PB1-F2, and N40. The whole sequence was retrieved from the NCBI database.

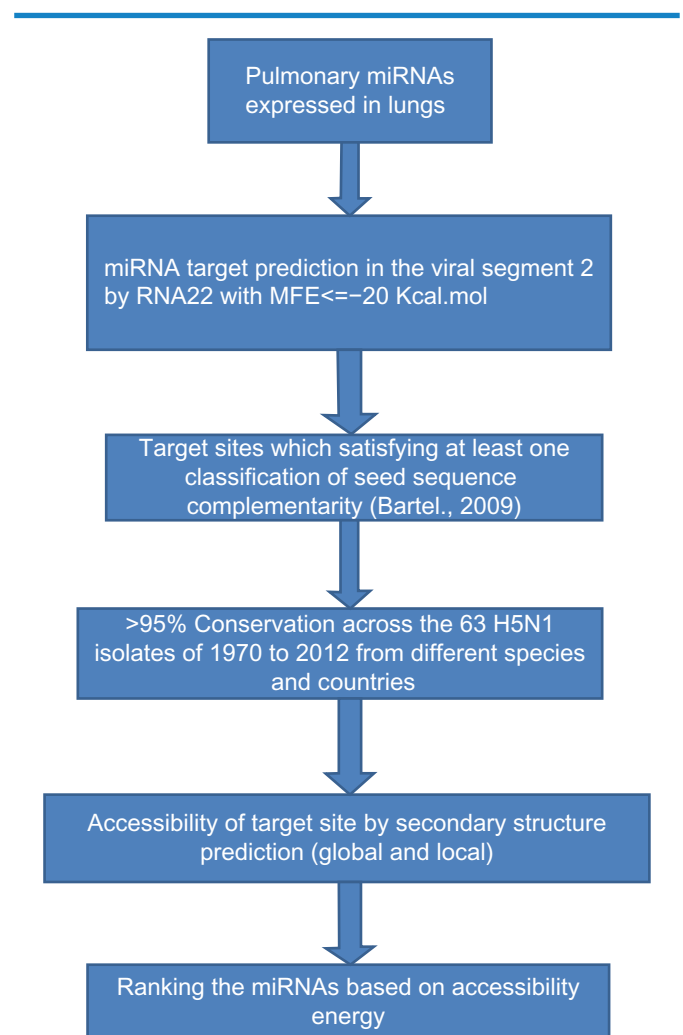
**Selection of miRNAs for targeting segment 2 of influenza A virus.** A set of 300 miRNA that were identified to be expressed in chicken lungs, by high throughput sequencing (unpublished), were included in the study and screened for presence of target site in the transcript of segment 2 of H5N1.

**Filtering of miRNAs for target identification.** Four successive filters were used to identify chicken miRNAs with potential binding site in segment 2 of H5N1 virus (Fig. 1). The miRNAs were first filtered for the thermodynamic stability of heteroduplex between miRNA and mRNA target site, based on minimum folding energy (MFE). The second criterion used was satisfaction of at least one type of seed sequence complementarity. This was followed by screening for conservation of miRNA target sites across 64 randomly selected

H5N1 isolates between 1970 and 2012. Finally, the accessibility of the target sites on the predicted secondary structure of the viral mRNA was also taken into consideration.

**Identification of miRNAs targets in segment 2 of influenza A virus.** In the first step of filtering, RNA22 miRNA target detection algorithm<sup>9</sup> (<http://cbcsrv.watson.ibm.com/rna22.html>) was used for identification of site in target mRNA. Keeping the maximum number of allowed UN-paired bases set to zero in a seed sequence of 6 bases and the minimum number of paired bases in heteroduplex as 14, each individual miRNA was screened for the presence of target sites in the segment 2 of H5N1. Based on the thermodynamics of the miRNA and target-site heteroduplex, all miRNAs with one or more target site in segment 2 of H5N1 with MFE of hybridization  $< -20$  kcal/mol were selected for further screening.

**Target sites with seed sequence complementary.** miRNA-mediated post-transcriptional regulation of gene expression is a sequence-specific process. The complementarity of the miRNA seed sequence and the target site are critical for effective targeting. Hence, the target sites of the miRNAs were



**Figure 1.** Work flow for the prediction of pulmonary chicken miRNAs that target expression of protein from segment 2 of HPAIV (H5N1).

further filtered for the type of seed sequence complementarity as described by Bartel (2009). Only those that showed at least one of the types of complementarity necessary for effectiveness of miRNA<sup>14</sup> were selected for further analysis.

**Conservation of target sequence.** Avian influenza virus H5N1 is known to undergo antigenic drift and shift due to changes in its gene sequences. In order to be targeted by a chicken miRNA, it is necessary that the target site should lie in the conserved regions of H5N1 genome. A total of 63 complete sequences of segments 2 of H5N1 isolated in 24 different Asian, European, American, and African nations from 1959 to 2012 were included. They were selected so as to include H5N1 isolated from different avian species including chicken, turkey, goose, bean goose, muscovy duck, mallard duck, wild duck pigeon, crow, swan, black-nested grebe, whooper swan, quail, and magpie. The sequences were downloaded and aligned using MEGA5 software. Only those miRNAs whose target sites were conserved across at least 95% of the isolates were selected for further analysis.

**Target-site accessibility by RNA folding structure.** As a measure to further strengthen the target-site prediction procedure, accessibility of the target site was the next criteria. Accessibility of the target sites was evaluated based on the secondary structure of the target mRNA. The “mfold Web Server” (<http://mfold.rna.albany.edu>) was used for RNA folding form with default parameters. Two types of folds, namely the global fold, which was the secondary structure of the complete coding sequence, and the regional fold, which was the secondary structure obtained by folding about 220 bases with 100 bases on either sides of the target site sequence, were included in the study. The presence of at least three unpaired nucleotides in the sequence complimentary to the seed sequence was taken as the filter criteria.

**Ranking of predicted miRNA target.** The accessibility energy of the identified miRNA targets was used for ranking the miRNAs identified to target the segment 2 of H5N1. The MFE of regional folding with constraint of unpairing at the target site ( $\Delta G_{\text{unpaired}}$ ) compared with that of the native fold ( $\Delta G_{\text{free}}$ ) to assess the free energy of unfolding ( $\Delta G_{\text{open}}$ ). The accessibility energy of the miRNA target sites were calculated as: accessibility energy ( $\Delta \Delta G$ ) =  $\Delta G_{\text{duplex}} - \Delta G_{\text{open}}$ , where,  $\Delta G_{\text{open}} = \Delta G_{\text{free}} - \Delta G_{\text{unpaired}}$ . The miRNA with the lower accessibility energy was ranked best.

## Results

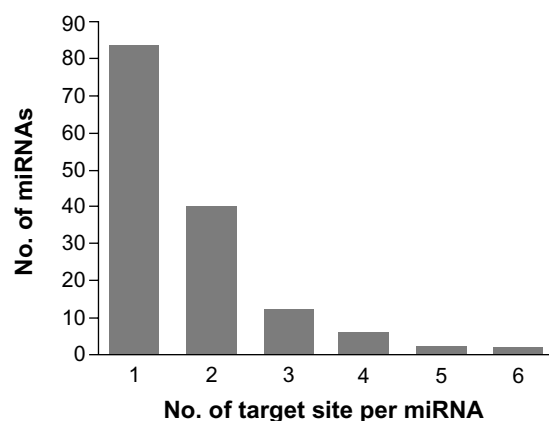
**Identification of miRNAs targets in segment 2 of influenza A virus.** On screening, the 300 miRNAs expressed in the chicken lungs, for target sites in the segment 2 of H5N1 virus, 146 miRNAs showed 253 targets (Supplementary Table 1) in the transcript of segment 2 with MFE of heteroduplex less than  $-20$  kcal/mol. While 84 miRNAs had single target sites, some of the miRNAs had more than one target site and even up to six target sites (Fig. 2). The target sites were identified all along the length of the segment

transcript including the 5'-UTR, 3'-UTR, and the Coding sequence (CDS) of the genes coded by the segment.

**Seed sequence complementarity and conservation of target sites.** Complementarity between the target site and the seed sequences of the miRNAs have been identified as an important factor for the effectiveness of miRNA-mediated knocking down. On screening, the target-miRNA hybrid predicted by RNA22 was reported for complementarity types.<sup>14</sup> Twenty-four miRNAs-mRNA hybrids were found to have target sites with seed sequence complementarity necessary for effective miRNA targeting (Table 1, Supplementary data file). Of these, 11 were found to be conserved across more than 95% of the sequence of segment 2 of the 63 different isolates of H5N1 included for the analysis (Fig. 3, Table 2). The miRNA gga-miR-1593 had two target sites that satisfied the criteria of seed sequence complementarity but only one of them was sufficiently conserved.

**Target site accessibility.** The miRNAs gga-miR-133c, gga-miR-137, gga-miR-146c\*, gga-miR-1593, and gga-miR-1710 had accessible targets on transcript-based global folding structures (Fig. 4), while gga-miR-133c, gga-miR-1553, gga-miR-1593, gga-miR-1671, and gga-miR-1710 show accessibility to target site based on the regional folding (Fig. 5). Those miRNAs that show accessibility to the target site by both folding methods may be assumed to have higher probability of being effective in down regulation of the target proteins as the secondary structure of RNA is a dynamic one in a live cell.

**Ranking of miRNAs by accessibility energy.** The accessibility energy is an indicator of the thermodynamic feasibility of the process of miRNA-mRNA hybridization while the former is in its secondary structure. Hence, this is a suitable parameter to rank the effectiveness of the predicted miRNAs (Fig. 5). The basis here is the difference between the MFE of native folding and the MFE of folding, such that the base pairing at the seed sequence complementary site is inhibited (Figs. 6 and 7). On ranking the miRNAs based on



**Figure 2.** Distribution of microRNAs against the number of target sites (MFE  $\leq -20$  kcal/mol) identified per miRNA in segment 2 of HPAIV (H5N1).



**Table 1.** MicroRNAs with seed sequence complementarity necessary for effective mRNA targeting.

miRNAs	miRNA-mRNA HYBRIDIZATION	TYPE OF SEED SEQUENCE COMPLEMENTARITY
gga-miR-27b	CTAGC-GAGG--AGTATCTGTGAG                  ---CGUCUUGAAUCGGUGACACUU	6mer site
gga-miR-30a-3p	A--GGAGATCATTCGAGCTGAAAA                   CGACGUUU-GUAGGCU-GACUUUC	7mer-A1 site
gga-miR-92	ACCAGGGCA-GACT-GTGCAATC                 --GUCCGGCCUGUUCACGUUAU	7mer-m8 site
gga-miR-133c	----TTGGTTTCAGATGGAGACCAA                CGUCGACCAA---CUUCC-CCUGGUU	7mer-A1 site
gga-miR-137	AGATTGTGTATTGG-AAGCAATG               GAUG-CGCAUAAGAAUUCGUUAU	3'-supplementary
gga-miR-146c*	AGGGACAGGTATACCATGGAC                UCUCUUGACU-UAUGGUACCUGA	7mer-m8 site
gga-miR-194	CGACATGAGCATTGGTGTTACA                    AGGUGUAC-CUCAACGACAAUGU	7mer-A1 site
gga-miR-204	GAGGA--AGAGAAGAGTAAGGGAC                    -UCCGUAUC-CUACUGUUUCCUU	6mer site
gga-miR-211	GAGGA--AGAGAAGAGTAAGGGAC                    -UCCGUAUC-CUACUGUUUCCUU	6mer site
gga-miR-367	-ACCAGGGCAGACTGTGCAATC                   GUGGUAACGAUUU-CACGUUAA	7mer-m8 site
gga-miR-1553	-CATGAAGATCTGTTCCACCATT                  UGUAGUGACAGG-AGGUGUGUGU	Offset 6mer site
gga-miR-1561	----GCGAGCAATTGCAACACCCGG                 UUUUCGC-CGG---CGUU-UGGGCCG	6mer site
gga-miR-1593	GACTGTGGTTTCAGATGGA                ACGCACAGGUUAGUCUACCA	8mer site
gga-miR-1593	AATGT-TCTCAAACAAGATGGC                ACGCACAG-GUUUAGUCUACCA	6mer site
gga-miR-1628	-GTTCCACCATTGAAGAGCTCAG                 CCAGUGUUGUC-CUUCUCGAGAA	Offset 6mer site
gga-miR-1629	AGATCAAACGGTCTGACAGCA                   UCUUGUUUGGGUGGUGUCGU	7mer-A1 site
gga-miR-1632	-CGAA-TCATGAAGGAATACAAGCA                   CGUUCGAGUAGGUUUU--UGUUCGU	7mer-A1 site
gga-miR-1662	AAACCATT--TGAATGGATGTCAA                UAGGGUUCAUACUA--CUACAGUU	7mer-A1 site
gga-miR-1670	AACATGCTGAG--TACAGTCCTAG                   G-GUGACACUCCGAUG-CAGGGUG	3'-compensatory
gga-miR-1671	---GGACGGCT-GATAGATTTCCTCAA                AAACC-GGUGAGUUGUC---AGGAGUG	7mer-A1 site

(Continued)



Table 1. (Continued)

miRNAs	miRNA-mRNA HYBRIDIZATION	TYPE OF SEED SEQUENCE COMPLEMENTARITY
gga-miR-1692	GCTGAACAAAAGGAGCTACC                  UGAGAUGGUUGACUCGAUGU	7mer-A1 site
gga-miR-1710	CAGGACTGTTG-GTTTCAGATGG                   A-CCUACCAAUACGUCGUCUACU	6mer site
gga-miR-1803	CCATTCTCAACACGAGTCAAA                 GGUUUGGGCGACUUCAGUUA	8mer site
gga-miR-3526	AGCTCAGATGGC-TCTTCAG                  -UGUGGUUGAAGUAGAAGUU	6mer site

the accessibility energy, gga-mir-133c was found to be the best followed by gga-mir-1710 and gga-mir-146c\* (Fig. 8). Of these, the first two also show accessibility to target site both on global folding and regional folding. Those miRNA that show

accessibility to the target site by both folding methods may be assumed to have higher probability of being effective in down regulation of the target proteins as the secondary structure of RNA is a dynamic one in a live cell.



Figure 3. Alignment report complementary site of the seed sequence of miRNAs across segment 2 of 63 H5N1 isolates.

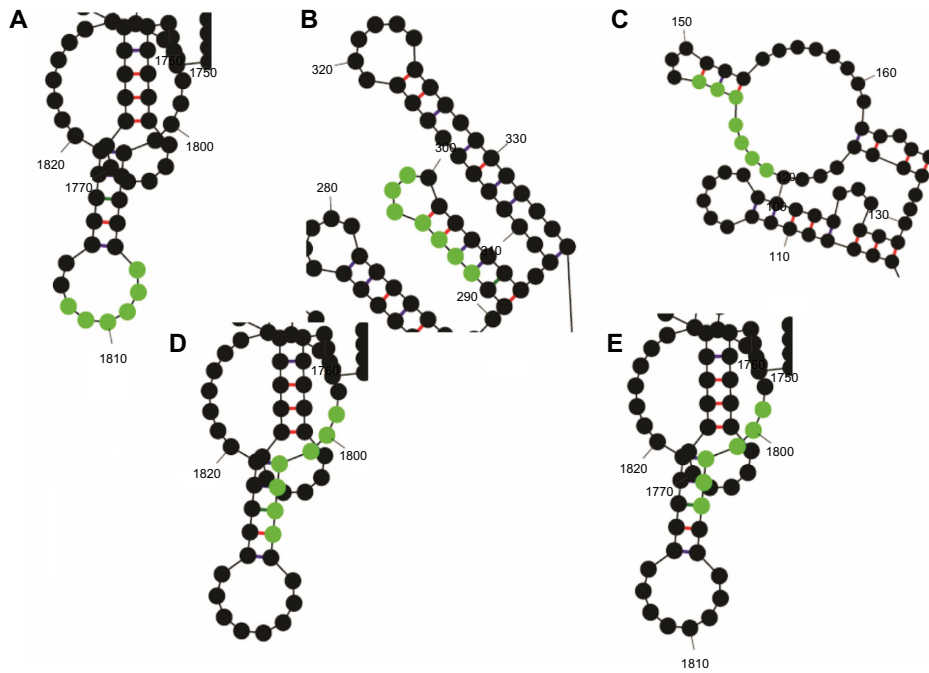


**Table 2.** Homology of target site sequence across segment 2 of 63 HPAIV (H5N1) isolates from different avian species in 24 different countries, reported between years 1970 and 2012. "+" indicates 100% homology of target sequence.

ACC. NO.	HOST SPECIES	COUNTRY	YEAR	CHICKEN miRNAs (LOCATION OF TARGET SITE WITH REFERENCE TO CY092122)																																																	
				gga-miR-1671(505 TO 527)	gga-miR-3526(1653 TO 1671)	gga-miR-1803(2042 TO 2062)	gga-miR-1670(1261 TO 1282)	gga-miR-1553(2253 TO 2274)	gga-miR-1628(2264 TO 2285)	gga-miR-1692(657 TO 676)	gga-miR-1593(1050 TO 1070)	gga-miR-1593(1786 TO 1806)	gga-miR-1629(475 TO 495)	gga-miR-27B(796 TO 816)	gga-miR-30A-3P(1735 TO 1756)	gga-miR-92(1883 TO 1903)	gga-miR-133C(1792 TO 1813)	gga-miR-137(279 TO 300)	gga-miR-146C*(126 TO 147)	gga-miR-194(1587 TO 1608)	gga-miR-204(582 TO 603)	gga-miR-211(582 TO 603)	gga-miR-367(1883 TO 1903)	gga-miR-1561(738 TO 758)	gga-miR-1632(1385 TO 1407)	gga-miR-1662(14 TO 35)	gga-miR-1710(1784 TO 1805)																										
CY092122	Chicken	India	2011	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+																		
CY015087	Chicken	United Kingdom	1959	+	+			+	+			+						+	+	+														+	+	+	+																
GU052524	Chicken	United Kingdom	1959	+	+			+	+			+						+	+	+																+	+	+	+														
GU052009	Mallard	USA	1975	+	+			+	+			+						+	+	+					+												+	+	+	+													
EU871818	Turkey	Canada	1983	+	+						+							+	+					+														+	+	+	+												
CY015125	Turkey	United Kingdom	1991	+	+			+	+	+		+						+	+	+				+	+	+		+	+	+		+	+	+				+	+	+	+												
AF144301	Goose	China	1996	+	+			+	+	+	+	+						+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+						+	+	+	+											
GU052478	Avian	Italy	1997	+	+			+	+	+		+						+	+					+	+																+	+	+	+									
DQ997120	Chicken	China	1997	+							+							+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+										
GU052484	Duck	Hong Kong	1998	+	+			+	+	+		+						+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+								
AY585492	Duck	China	1999	+	+			+	+	+		+						+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+							
GU052025	Goose	Hong Kong	1999	+	+	+		+	+	+	+	+						+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+						
AY585486	Duck	China	2000	+	+			+	+	+		+						+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+						
GU052055	Goose	Hong Kong	2000	+	+	+	+	+	+	+	+	+						+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+					
GU051307	Mallard	USA	2000	+	+	+		+	+			+						+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+					
DQ351874	Chicken	China	2001	+				+	+	+	+		+					+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+					
GU186690	Duck	Hong Kong	2001	+	+			+	+	+	+	+						+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+				
GU052111	Goose	Viet Nam	2001	+	+	+	+	+	+	+	+							+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+				
DQ997084	Chicken	China	2002	+	+			+	+	+		+						+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+			
GU052463	Goose	Hong Kong	2002	+	+	+	+	+	+	+	+							+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+			
GU052416	Chicken	Indonesia	2003	+	+	+	+	+	+	+	+	+						+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		
AY676027	Chicken	South Korea	2003	+	+	+	+	+	+	+	+	+						+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		
AB212278	Duck	Japan	2003	+	+	+	+			+	+							+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
GQ122443	Chicken	Indonesia	2004		+	+	+	+	+	+	+	+						+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
GU052455	Magpie	South Korea	2004	+	+	+	+	+	+	+	+	+						+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
GU186791	Wild Duck	USA	2004						+	+								+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		
EU147291	Avian	Russia	2005	+	+	+		+	+	+		+						+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
CY095583	Mallard	Italy	2005	+	+			+	+	+		+						+	+					+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
EF517399	Mallard	Canada	2005	+	+	+		+	+									+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
AB450529	Pigeon	Thailand	2005	+	+	+	+	+	+	+		+						+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
DQ914811	Chicken	China	2006	+	+	+	+	+	+	+	+							+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
DQ464361	Swan	Germany	2006	+	+	+		+	+	+		+						+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
EF362424	Chicken	India	2006	+	+	+		+	+	+		+						+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
HM006749	Chicken	Laos	2006	+	+	+	+	+	+	+	+	+						+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
AM262525	Chicken	Nigeria	2006	+	+	+		+	+	+		+						+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
GU182140	Chicken	China	2007	+	+	+	+	+	+	+	+							+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

(Continued)

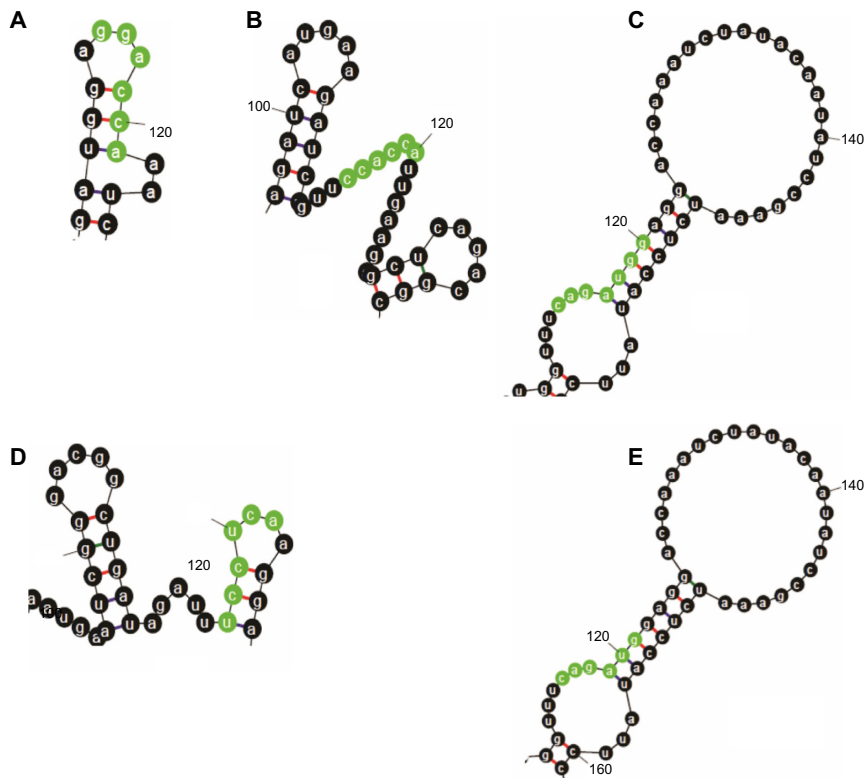




**Figure 4.** MicroRNA target sites of (A) miR-133c, (B) miR-137, (C) miR-146c\*, (D) miR-1593, and (E) miR-1710, highlighted from the secondary structure of mRNA from segment 2 of H5N1 in global folding.

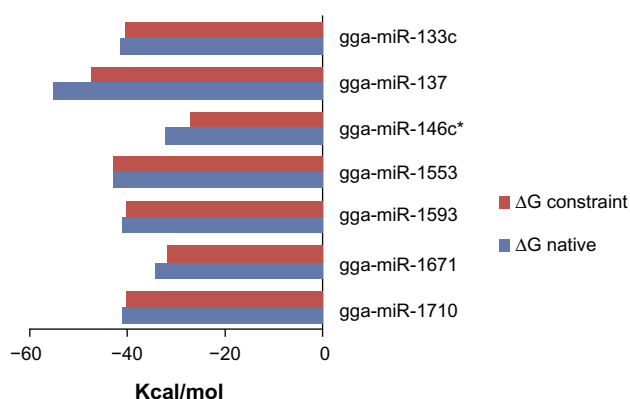
The first criteria for target identification as free energy below  $-20$  kcal/mol for miRNA-target hybridization were based on reports of similar target prediction studies with reference to influenza virus. Some researchers have used a cutoff

free energy of  $-17$  kcal/mol for predicting miRNA targets in PB1 gene of H1 N1<sup>21</sup> and some have used a threshold free energy of  $-15$  kcal/mol while predicting the binding site of miRNA to NS1 gene of H1 N1 and H5N1.<sup>33</sup> The complete length of



**Figure 5.** MicroRNA target sites of (A) miR-133c, (B) miR-1553, (C) miR-1593, (D) miR-1671 and (E) miR-1710, highlighted from the secondary structure of mRNA from segment 2 of H5N1 in regional folding. The regional folding was obtained by inclusion of 10 bases on either side of the target sequence for secondary structure prediction.



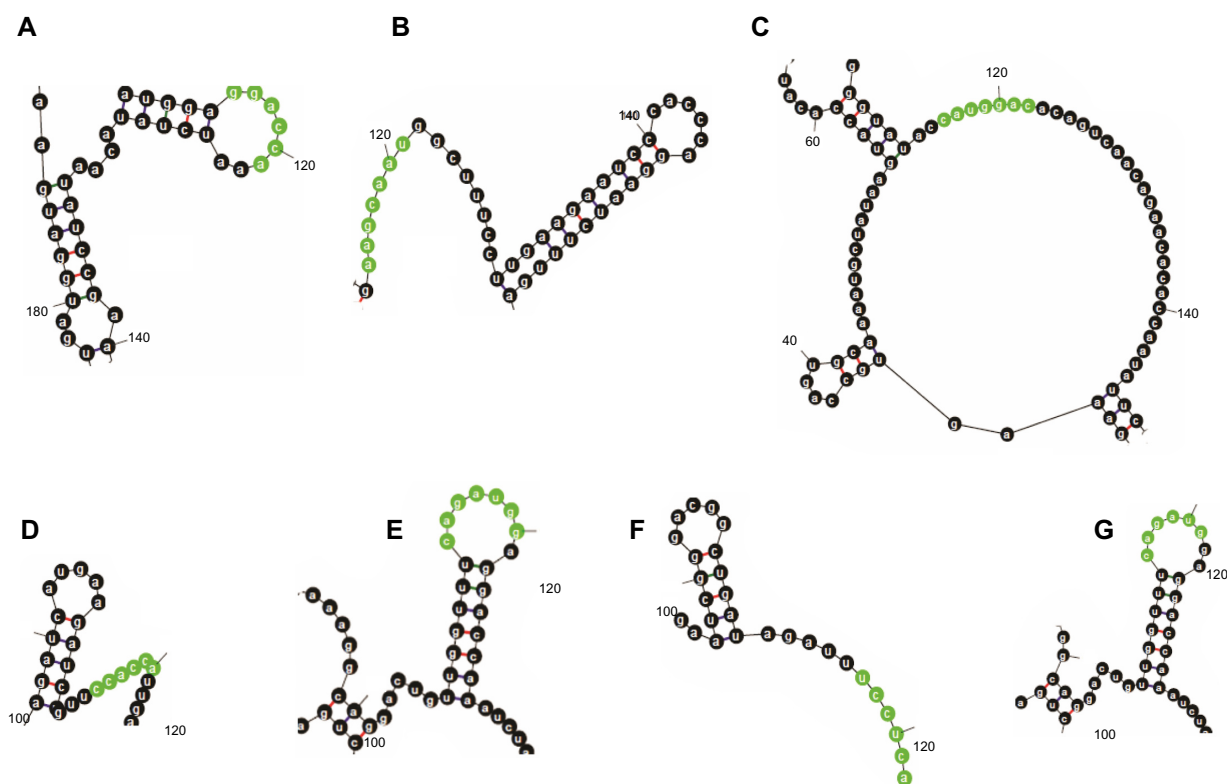


**Figure 6.** Comparison of the MFE value ( $\Delta G$ ) of native secondary structure and secondary structure with a constraint to prohibit base pairing at the miRNA target site.

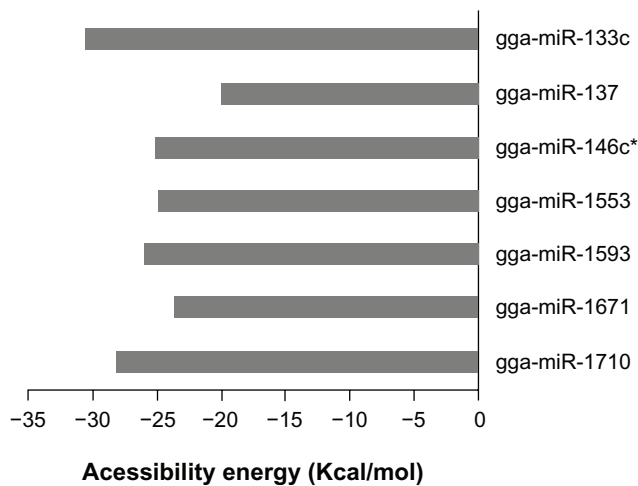
segment 2 was screened for target sites of the miRNAs as three different proteins were expressed from the same segment and the same target site may have a different impact on the different proteins. Although most investigation into metazoan miRNA function has been for sites in 3'-UTRs, experiments using artificial sites show that targeting can occur in 5'-UTRs and open ORFs.<sup>22,23</sup> The mRNA and protein expression of genes containing target sites both in coding regions and 3'-UTRs are reported to be mildly but significantly more regulated than those containing target sites in 3'-UTRs only.<sup>24</sup>

Seed sequence complementarity is critical for effective miRNA targeting. Although perfect complementarity is not necessary for effective silencing of expression in animal, specific pattern of seed sequence complementarity have been reported to be effective on the basis of experimentation and prediction.<sup>12</sup> Hence, the target sites predicted by the RNA22 were analyzed to classify them as per seed sequence complementarities. AIV is known to show genetic drift before an miRNA can be identified to target its gene, it was important that it should be conserved. Also change in the sequence at the target site would alter the thermodynamics of hybridization as well as seed sequence complementarity thus affecting the effectiveness. It is conventional to screen for conservation of the target site in majority of target-prediction databases like TargetScan, miRanda, etc. This filter was also necessary as the purpose of the study was to identify genetic components of chicken genome that can target the viral gene expression, and unless the target site is conserved, there would not be consistency in results.

The targets identified by our approach were not predicted by the ViTa database, which predicts targets of host miRNAs in viral genes.<sup>25</sup> This might be due to the fact that the target identification tools in its pipeline is miRanda<sup>26</sup> and TargetScan,<sup>27</sup> while we have used RNA22 here. Although miRanda is an efficient tool, it underestimates miRNAs with single but perfect base pairing as against those with multiple target sites. On the other hand, TargetScan identifies only those sites that are conserved across several species. Conservation across several



**Figure 7.** MicroRNA target sites of (A) miR-133c, (B) miR-137, (C) miR-146c\*, (D) miR-1553, (E) miR-1593, (F) miR-1671, and (G) miR-1710, highlighted from the secondary structure of mRNA from segment 2 of H5N1 in regional folding with a constraint to prohibit base pairing at the target site.



**Figure 8.** Accessibility energy of miRNAs target site on the mRNA from segment 2 of H5N1.

species was not an important factor here, as the rationale of this study was data mining of chicken miRNAs only for the presence of genetic factors in chicken, if any, which can potentially affect the expression of critical H5N1 genes.

Accessibility is an important factor for targeted down-regulation by miRNA. Long et al (2007) suggested that the presence of at least four consecutive nucleotides complementary to any part of miRNA in unpaired state in secondary structure of target RNA is required for initiation of the binding process between the target and the miRNA.<sup>28</sup> Robin et al (2005) proposed screening based on the presence of only three consecutive nucleotides complementary to the seed sequence as unpaired.<sup>29</sup> Here, the presence of only three consecutive unpaired nucleotides complementarity to the seed sequence was used as filter criteria to improve precision of the target prediction, because this gives due importance to the seed sequence complementarity while screening for the unpaired stretches in the secondary structure of target sequence. The folding algorithm cannot warranty the reliable structures of long sequences.<sup>30</sup> The protein involved in the recognition of RNA strand are expected to hinder the formation of long-range interactions, thus making target accessibility a matter of local secondary structure.<sup>31</sup> However, the predicted local folding may also not represent the exact secondary structure of an RNA molecule. Hence, for the study, global as well as regional folding of the mRNA was considered.

## Conclusion

The chicken genome has miRNAs that can potentially target the expression of the genes expressed from the segment 2 of H5N1 virus. The position of target of gga-miR-133c on segment 2 transcript is 1792–1813. Hence, it can target the PB1 and N40 protein by binding in their coding region and PB1-F2 by binding at 3'-UTR. The gga-miR-146c\* targets position 126–147 and hence can inhibit PB1 and PB1-F2 at

the coding regions and N40 at the 5'-UTR. This indicates that these host miRNAs have potential for regulating the expression of multiple viral proteins namely, PB1, PB1-F2, and N40. Since these miRNA are being expressed in the chicken lungs, they may be useful for inhibition of H5N1 replication or the development of resistance against H5N1 infection in chicken. However, the miRNAs identified to target genes of segment 2 are targeting the different proteins at different location with reference to their ORF; hence, they might have variable extent of impact on the regulation of these gene expressions. Those miRNAs that show accessibility to the target site by both folding methods may be assumed to have higher probability of being effective in down regulation of the target proteins as the secondary structure of RNA is a dynamic one in a live cell.

Owing to the fact that miRNAs are pleiotropic in their effect, these shortlisted miRNAs may have other off-target effects as well. Several databases like the miRDB<sup>32</sup> or TargetScan.<sup>27</sup> provide list of target genes for each of these miRNAs in chicken, but they bioinformatically predicted based on orthology and other parameters. Here, the effective impact of the miRNAs on these predicted host gene targets cannot be compared with their inhibitory impact on segment 2 genes of H5N1 as the pipeline for prediction are different. Hence, the biological usefulness and effectiveness of these miRNAs for the inhibition of H5N1 replication in chicken needs to be evaluated further by biological experimentations such as in vitro virus challenge and luciferase reporter assays, etc. However, the bioinformatics approach of this study has aided in narrowing down on three miRNAs for targeting the segment 2 of H5N1, from the given 300 miRNAs expressed in chicken lungs. This approach of in silico screening and short listing, before taking up biological experimentation, would save enormous resources that go into the conducting of wet laboratory experiments especially in the case of H5N1, which is a BSL3 agent.

## Author Contributions

Conceived and designed the experiments: AM, AK. Analyzed the data: AM. Wrote the first draft of the manuscript: AK. Contributed to the writing of the manuscript: AM, AAR. Agree with manuscript results and conclusions: AK, MAVN, AM, AAR, RS, AM. Jointly developed the structure and arguments for the paper: AK, MAVN, AM, AAR. Made critical revisions and approved final version: AM. All authors reviewed and approved of the final manuscript.

## Supplementary Material

**Supplementary File 1.** List of target site identified by RNA22 with MFE  $\leq -20$  kcal/mol.

**Supplementary File 2.** RNA22 output data of miRNA-mRNA hybridization.

## REFERENCES

- Alexander DJ. A review of avian influenza in different bird species. *Vet Microbiol.* 2000;74:3–13.



2. OIE. *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals* (Mammals, Birds, and Bees). 5th ed. Paris, France: OIE; 2004:258–69.
3. Rebmann T, Zelicoff A. Vaccination against influenza: role and limitations in pandemic intervention plans. *Expert Rev Vaccines*. 2012;11(8):1009–19.
4. WHO. Use of antiviral drugs in poultry, a threat to their effectiveness for the treatment of human avian influenza. Available at [http://www.who.int/food-safety/micro/avian\\_antiviral/en/](http://www.who.int/food-safety/micro/avian_antiviral/en/). 2005.
5. Ambros V. The functions of animal microRNAs. *Nature*. 2004;431:350–5.
6. Miska EA. How microRNAs control cell division, differentiation and death. *Curr Opin Genet Dev*. 2005;15:563–8.
7. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenines, indicates that thousands of human genes are microRNA targets. *Cell*. 2005;120:15–20.
8. Lim LP, Lau NC, Garrett-Engele P, et al. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature*. 2005;433:769–73.
9. Miranda KC, Huynh T, Tay Y, et al. A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell*. 2006;126:1203–17.
10. Guo X, Zhang Q, Gao L, Li N, Chen X, Feng W. Increasing expression of microRNA 181 inhibits porcine reproductive and respiratory syndrome virus replication and has implications for controlling virus infection. *J Virol*. 2013;87(2):1159.
11. Sun JZ, Wang J, Yuan D, et al. Cellular microRNA miR-181b inhibits replication of mink enteritis virus by repression of non-structural protein 1 translation. *PLoS One*. 2013;8(12):e81515.
12. Tauber S, Ligertwood Y, Quigg-Nicol M, Dutia BM, Elliott RM. Behaviour of influenza A viruses differentially expressing segment 2 gene products in vitro and in vivo. *J Gen Virol*. 2012;93:840–9.
13. Chamnanpood C, Sanguanserm Sri D, Pongcharoen S, Sanguanserm Sri P. Detection of distribution of avian influenza H5N1 virus by immunohistochemistry, chromogenic in situ hybridization and real-time PCR techniques in experimentally infected chickens. *Southeast Asian J Trop Med Public Health*. 2011;42(2):303–10.
14. Bartel DP. MicroRNAs: target recognition and regulatory functions. *Cell*. 2009;136(2):215–33.
15. Taubenberger JK, Kash JC. Influenza virus evolution, host adaptation, and pandemic formation. *Cell Host Microbe*. 2010;7:440–51.
16. Naffakh N, Tomoiu A, Rameix-Welti MA, van der Werf S. Host restriction of avian influenza viruses at the level of the ribonucleoproteins. *Annu Rev Microbiol*. 2008;62:403–24.
17. Schmolke M, Manicassamy B, Pena L, et al. Differential contribution of PB1-F2 to the virulence of highly pathogenic H5N1 influenza A virus in mammalian and avian species. *PLoS Pathog*. 2011;7(8):e1002186.
18. Conenello GM, Zamarin D, Perrone LA, Tumpey T, Palese P. A single mutation in the PB1-F2 of H5N1 (HK/97) and 1918 influenza A viruses contributes to increased virulence. *PLoS Pathog*. 2007;3(10):e141.
19. Zamarin D, Ortigoza MB, Palese P. Influenza A virus PB1-F2 protein contributes to viral pathogenesis in mice. *J Virol*. 2006;80:7976–83.
20. Wise HM, Foeglein A, Sun J, et al. A complicated message: identification of a novel PB1-related protein translated from influenza A virus segment 2 mRNA. *J Virol*. 2009;83:8021–31.
21. Song L, Liu H, Gao S, Jiang W, Huang W. Cellular microRNAs inhibit replication of the H1 N1 influenza A virus in infected cells. *J Virol*. 2010;84(17):8849–60.
22. Kloosterman WP, Wienholds E, Ketting RF, Plasterk RH. Substrate requirements for let-7 function in the developing zebrafish embryo. *Nucleic Acids Res*. 2004;32:6284–91.
23. Lytle JR, Yario TA, Steitz JA. Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5' UTR as in the 3'UTR. *Proc Natl Acad Sci USA*. 2007;104:9667–72.
24. Fang Z, Rajewsky N. The impact of miRNA target sites in coding sequences and in 3'UTRs. *PLoS One*. 2011;6(3):e18067.
25. Hsu PW, Lin LZ, Hsu SD, Hsu JB, Huang HD. ViTa: prediction of host microRNAs targets on viruses. *Nucleic Acids Res*. 2007;35(Database issue):D381–5.
26. Enright A, John B, Gaul U, Tuschl T, Sander C, Marks D. MicroRNA targets in *Drosophila*. *Genome Biol*. 2003;5(1):R1.
27. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB. Prediction of mammalian microRNA targets. *Cell*. 2003;115(7):787–98.
28. Long D, Lee R, Williams P, Chan CY, Ambros V, Ding Y. Potent effect of target structure on microRNA function. *Nat Struct Mol Biol*. 2007;14:287–94.
29. Robins H, Li Y, Padgett RW. Incorporating structure to predict microRNA targets. *Proc Natl Acad Sci USA*. 2005;102(11):4006–9.
30. Doshi K, Cannone J, Cobaugh C, Gutell R. Evaluation of the suitability of free-energy minimization using nearest-neighbor energy parameters for RNA secondary structure prediction. *BMC Bioinformatics*. 2004;5:105.
31. Tafer H, Ameres SL, Obernosterer G, et al. The impact of target site accessibility on the design of effective siRNAs. *Nat Biotechnol*. 2008;26:578–83.
32. Wang X. miRDB: a microRNA target prediction and functional annotation database with a wiki interface. *RNA*. 2008;14(6):1012–7.
33. Koparde P, Singh S. Prediction of micro RNAs against H5N1 and H1 N1 NS1 protein: a window to sequence specific therapeutic development. *J Data Mining Genomics Proteomics*. 2010;1:2–8.