

## Bioinformatics study of complete amino acid sequences of neuraminidase (NA) antigen of H1N1 influenza viruses from 2006 to 2013 in Iran

Farhangi A<sup>1</sup>, Goliaei B<sup>1\*</sup>, Kavousi K<sup>1</sup>, Ashtari A<sup>2</sup>, Bayatzadeh MA<sup>2</sup>, Pourbakhsh A<sup>2</sup>

<sup>1</sup>Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran.

<sup>2</sup>Razi Vaccine and Serum Research Institute, Karaj, Alborz, Iran.

### ABSTRACT

**Introduction:** Influenza is a contagious acute viral disease of the respiratory tract that causes fever, headache, muscle aches and cough. One of the unique features of influenza virus is antigenic variation in viral protein neuraminidase (NA) which causes emergence of new virus variants. NA is responsible for the release and spread of progeny virions. Due to the continuous changes of NA genes, vaccine strains must be re-selected annually. **Methods:** Complete NA amino acid sequences of 97 strains circulating from 2006 to 2013 in Iran were downloaded from NCBI. The sequences were edited and classified by the year of isolation and their diversity and important changes as well as changes in the predicted ligand binding sites and their resistance to anti-NA drugs, were analyzed. Bioinformatics software such as MEGA6.0, BioEdit, DNAsisMAX and DNAsstar were used for the sequence alignments and phylogenetic analyses. Web-based analysis such as SWISS-MODEL, Phyre2 and 3DLigandSite were used for evaluation of the second and third protein structures and prediction of the ligand binding sites. **Results:** The results showed that 2009 could be considered as an important transition year which caused to classify the isolates into two different distinct groups. This shows the importance of changes made during possible mutations in the genomic structure of the virus which have made it antigenically different from the previous years. Anti-NA drug resistance was observed in 2009. This pandemic strain has become dominant in the following years and is used as a standard vaccine strain from 2010 onwards. **Conclusion:** The results obtained in this study can aid in better understanding of the antigenic evolution of H1N1 influenza viruses and can potentially accelerate the selection of the vaccine strains.

**KEYWORDS:** H1N1 influenza virus, Neuraminidase, Antigenic drift, Sequence analysis, Anti-NA drug resistance.

### INTRODUCTION

Influenza is a contagious acute viral disease of the respiratory tract that causes fever, headache, muscle aches, sweating, runny nose, sore throat and cough. It is caused by infection with type A or B influenza virus almost every year [1]. Influenza A viruses are enveloped viruses which belong to the *Orthomyxoviridae* family with segmented genome made of eight single-stranded negative RNA segments from 890 to 2,341 nucleotides that encode 10 and in some cases, 11 proteins [2]. The classification and nomenclature of influenza A viruses are based on the type of hemagglutinin (HA) and neuraminidase (NA) glycoproteins which are presented on the surface membrane [3]. To date, 16 serologically-distinct HA types and 9 different NA types have been described. Recently, a new HA subtype (H17) has been identified in Bats in South America [4]. NA and HA are two main external transmembrane glycoproteins, extended from the surface of the virus envelope.

These two antigens are the most variable antigens that cause protective immunity. HA is the receptor binding glycoprotein which mediates membrane fusion process [5]. NA is responsible for assisting in the virus entry and the release and spread of the progeny virions [6-8]. In fact, NA is the component of the virus membrane with enzymatic activity which is involved in the spread of influenza infection from cell to cell [9, 10]. At the late stages of the viral life cycle, NA removes sialic acids from the virus progeny and the infected cells, thus prevents the virus aggregation and promotes its detachment and release [11].

Antigenic variation is a significant feature in influenza [1]. In other words, changes in proteins of virus specially (i.e. HA and NA) would cause variants that are new in the population. Every year or every few years, occurrences of point mutations that are minor antigenic changes leads to the antigenic drift. The antigenic shifts are major antigenic changes, leading to completely new virus variants that might have a pandemic potential. Vaccination and antiviral drugs are main methods for prevention and treatment of influenza disease. Continuous changes of HA and NA genes (antigenic changes) have made vaccine strains selection challenging and they must be selected

\*Corresponding Author: Bahram Goliaei, Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran.

Email: goliaei@ibb.ut.ac.ir

Tel: (+98) 2166498672

annually, based on the currently circulating strains. Antiviral drugs, especially NA inhibitors, are the valuable addition to the options used to control the influenza infection [12, 13].

Collecting, reporting and analysis of epidemiologic data obtained from inspection of antigenic changes in each year is an important procedure that global influenza surveillance program has emphasized on it and could be useful in selection of the vaccine strains [14, 15].

The recognition of the importance of anti-NA antibodies in immunity to influenza has led to studies of small molecules that can block the functions of NA, with a view to the antiviral drug development [16]. In the case of NA, the enzyme inhibitors; Zanamivir (GG167; Relenza, GlaxoSmithKline, Research Triangle Park, NC, USA) and Oseltamivir (GS4104; Tamiflu; Genentech, South San Francisco, California and Chugai Pharmaceutical Co., Japan) [17] are already licensed as anti-influenza drugs and are widely marketed for therapeutic and prophylactic uses [18-21]. They act specifically against all subtypes of influenza A and influenza B NAs that prevent the virus from being released from the infected cells.

Although the emergence of antiviral resistance is generally associated with drug treatments, due to the drug pressure, it may emerge coincidentally during a virus evolution when mutations are either neutral or confer an evolutionary advantage on the (epidemic) virus variant and are maintained in the absence of the drug pressure. The emergence and spread of anti-NA drug resistant viruses that should be monitored, emphasizes the importance of developing other effective agents against different viral targets and application of antiviral combinations to reduce the emergence of the resistance [22].

In this study which was conducted for the first time in Iran, the complete amino acid sequences of NA antigen of type A H1N1 influenza viruses available in NCBI GenBank from 2006 to 2013 in Iran were inspected to define the prevalence library and to study the trend of changes in NA sequences while determining their relation with the changes of the virus using bioinformatics methods. The relations of gene alterations and the tertiary structure (3D model) with the relevant regions in NA were also studied. Furthermore, the emergence of anti-NA drug resistance was inspected in each year. The relationship of changes of each year's isolates to the vaccine strain of that year was also investigated.

## MATERIALS and METHODS

Complete amino acid sequences (complete CDS) of NA antigen of human H1N1 influenza virus available in NCBI GenBank (<http://www.ncbi.nlm.nih.gov>) from 2006 to 2013 in Iran were downloaded and saved in FASTA format for creation of required data bank. All downloaded sequences were edited using DNASIS MAX 3.0 software (Hitachi Pharmacia, Hitachi Software Engineering Company, Yokohama, Japan) to have proper and identical sequences in terms of name and format for comparison and performing of the intended analyses. Totally, 97 sequences from 16 provinces of Iran were downloaded, edited and classified based on the year of isolation.

The sequences were aligned based on ClustalW method using BioEdit 7.1 program [23]. The Molecular Evolutionary Genetic Analysis (MEGA 6.0) software (<http://www.megasoftware.net>) was used to analyze the sequence alignments and the evolutionary distances between the strains were computed with the same software using the maximum composite likelihood method for evaluation and exact comparison of the sequences,

determination of conserved and variable regions and similarities and differences [24, 25]. Megaline-Dnastar software (Expert Analysis Software for PC, Inc., USA) was used for construction of phylogenetic trees by UPGMA (unweighted pair group method with arithmetic means) method.

Using the percent similarity/divergence matrix created by BioEdit program in each year, the sequences with the highest and lowest similarity with other sequences of that year were selected as distinct sequences for comparison and analysis. Then evaluation of secondary and tertiary protein structure and prediction/determination of the ligand binding sites, type and number of the ligand(s) of distinct sequences of each year were done using following web-based analysis software:

SWISS-MODEL (<http://swissmodel.expasy.org>) [26], Phyre2 (<http://www.sbg.bio.ic.ac.uk/phyre2>) [27] and 3DLigandSite (<http://www.sbg.bio.ic.ac.uk/3dligandsite>) [28].

## RESULTS

By investigating the results of sequence alignments and phylogenetic trees and percent similarity/divergence matrices and predicted secondary and tertiary structures (3D structures) of the ligand binding sites of distinct sequences of each year (data not shown), it was found that from 2006 to 2008, the sequences are identical and there is no change in the ligand binding sites, or the type and number of the ligands.

In 2009, with respect to the occurrence of 2009 H1N1 pandemic, by studying the available similarities and differences in alignments of the sequences (Fig.1) and by investigating the phylogenetic tree of the 2009 isolates (Fig. 2), it was found that influenza isolates of Iran, are located in two entirely distinct and separated branches and form two different groups in terms of amino acid properties.

The remarkable point is that, the first group, in terms of amino acid sequences and gene positions, is similar to the sequences of 2006-2008 isolates and in some cases is entirely identical while the second group is similar to the sequences of 2010-2013 isolates and in some cases is entirely identical. This similarity, in terms of alignment analyses of amino acid sequences and the structure of phylogenetic tree between the two different groups in 2009, shows the importance of changes in the genomic structure of the virus in this year which were possibly due to mutations in the genome and the occurrences of frame shift or gene shuffling, causing different sequences and a new group. This new group became dominant in the following years, with significant amino acid changes in NA genome that resulted in isolates with different clinical properties. It was also found that, in 2009, the distinct sequence which had the highest similarity (approximately 99%) with the other sequences of this year (ADJ18168) had also the highest similarity with the sequences of the following years isolates (i.e. 2010-2013) and the distinct sequence which had the lowest similarity (approximately 81%) with the other sequences of 2009 (ACY01439) had the highest similarity with the sequences of the previous year's isolates (i.e. 2006-2008) (Table 1). This can explain the domination of 2009 pandemic strain from 2010 onwards.

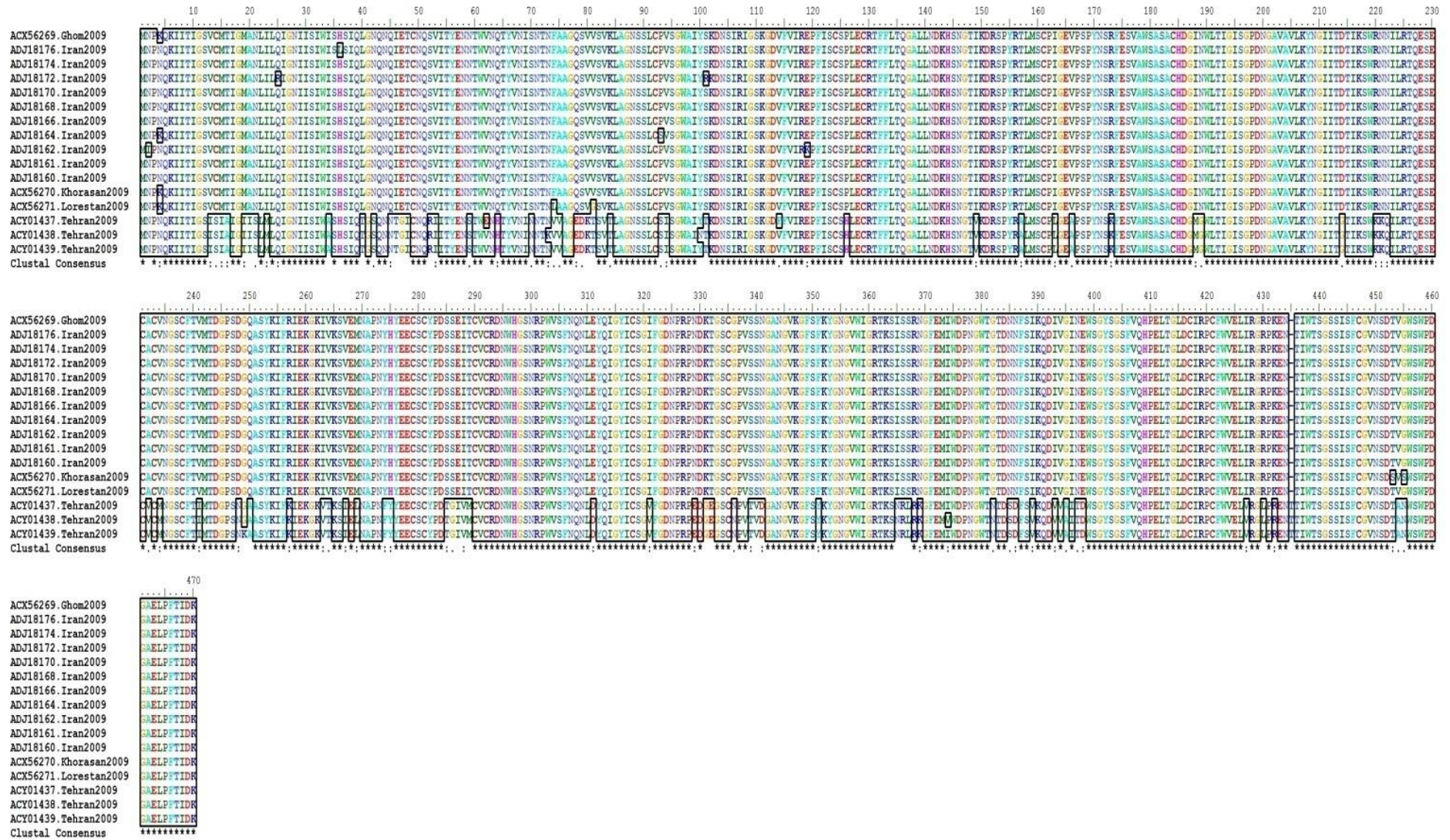
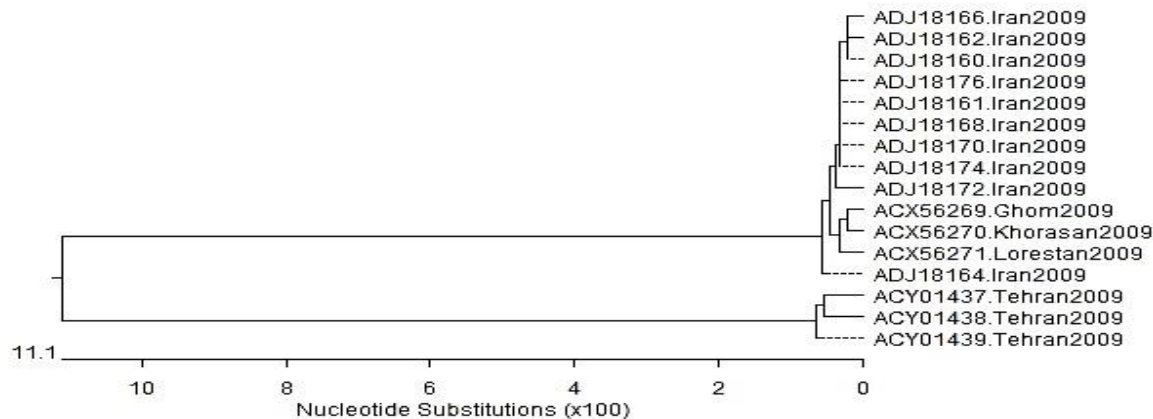


Fig. 1. Sequence alignment of 2009 influenza virus isolates in Iran.



**Fig. 2.** Phylogenetic tree of the Neuraminidase (NA) gene by UPGMA method. Evolutionary relationships of individual coding regions from influenza A viruses sampled in Iran during 2009. Length of each branch pair represents the evolutionary distance between sequence pairs. Scale indicates the number of substitution. Megalign<sup>TM</sup> constructed phylogram (DNAS<sup>TM</sup>Star, Madison, WI).

**Table 1.** Percentage similarity/divergence of 2009 isolates. The triple asterisks indicate when rows and columns refer to the same isolate. A diagonal line formed by the triple asterisks divide the table into two triangles, above and under it. The percentage similarity values are indicated in the upper triangle and the percent divergence values are indicated in the lower triangle.

	ACX 56269	ADJ 18176	ADJ 18174	ADJ 18172	ADJ 18170	ADJ 18168	ADJ 18166	ADJ 18164	ADJ 18162	ADJ 18161	ADJ 18160	ACX 56270	ACX 56271	ACY 01437	ACY 01438	ACY 01439
ACX 56269	***	99.6	99.8	99.4	99.8	99.8	99.8	99.8	99.4	99.8	99.8	99.6	99.6	81	80.8	81.2
ADJ 18176	0.4	***	99.8	99.4	99.8	99.8	99.8	99.4	99.4	99.8	99.8	99.1	99.1	81	80.8	81.2
ADJ 18174	0.2	0.2	***	99.6	100	100	100	99.6	99.6	100	100	99.4	99.4	81.2	81	81.4
ADJ 18172	0.6	0.6	0.4	***	99.6	99.6	99.6	99.1	99.1	99.6	99.6	98.9	98.9	81	80.8	81.2
ADJ 18170	0.2	0.2	0	0.4	***	100	100	99.6	99.6	100	100	99.4	99.4	81.2	81	81.4
ADJ 18168	0.2	0.2	0	0.4	0	***	100	99.6	99.6	100	100	99.4	99.4	81.2	81	81.4
ADJ 18166	0.2	0.2	0	0.4	0	0	***	99.6	99.6	100	100	99.4	99.4	81.2	81	81.4
ADJ 18164	0.2	0.6	0.4	0.9	0.4	0.4	0.4	***	99.1	99.6	99.6	99.4	99.4	81.2	81	81.4
ADJ 18162	0.6	0.6	0.4	0.9	0.4	0.4	0.4	0.9	***	99.6	99.6	98.9	98.9	80.8	80.6	81
ADJ 18161	0.2	0.2	0	0.4	0	0	0	0.4	0.4	***	100	99.4	99.4	81.2	81	81.4
ADJ 18160	0.2	0.2	0	0.4	0	0	0	0.4	0.4	0	***	99.4	99.4	81.2	81	81.4
ACX 56270	0.4	0.9	0.6	1.1	0.6	0.6	0.6	0.6	1.1	0.6	0.6	***	99.1	80.8	80.6	81
ACX 56271	0.4	0.9	0.6	1.1	0.6	0.6	0.6	0.6	1.1	0.6	0.6	0.9	***	81	80.6	81.2
ACY 01437	21.9	21.9	21.7	21.9	21.7	21.7	21.7	21.7	22.2	21.7	21.7	22.2	21.9	***	98.9	99.4
ACY 01438	22.2	22.2	21.9	22.2	21.9	21.9	21.9	21.9	22.5	21.9	21.9	22.5	22.2	1.1	***	99.1
ACY 01439	21.7	21.7	21.4	21.7	21.4	21.4	21.4	21.4	21.9	21.4	21.4	21.9	21.7	0.6	0.9	***

By predicting and analyzing the secondary and tertiary structures in the amino acid sequences of 2009 and comparison of these structures and analysis and comparison of regions and amino acid sequences in the binding ligands and active sites of the protein in molecular interactions between the ligand and the receptor, the differences between these two groups in amino acid sequences analysis (bifurcation of phylogenetic tree) of 2009 isolates became more clear. It was found that in the ligand binding site 1, amino acid positions 83-84-235-411, in two distinct sequences of 2009 were changed from Val-Thr-Asn-Glu with 15 ligands to Val-Lys-Asn-Glu with 17 ligands (Fig. 3).

Since 2010 to 2013, sequences are identical, there is no change in their ligand binding sites and the type and number of the ligands (data not shown). This result shows that the predominant 2009 pandemic strain is also dominant in these years.

In general, studying changes and mutations in conserved positions and regions, especially those located in active/catalytic sites or the areas associated with anti-NA drug resistance, are very important because these regions usually have functional or structural values in the protein sequence and can be used to be developed as targets for the new drugs, the infection control, the treatments and design of new vaccines.

In the next phase of this study, the sequences of different isolates were studied in terms of their position and the number of glycosylation sites at Asn-X-Ser/Thr motif, where X can be any amino acid except proline [29]. These sites play an important role in some features of the protein such as folding, stability, antigenicity and immunogenicity [30]. The sequences of 2006-2008 isolates have 9 conserved glycosylation sites at positions: 44, 58, 63, 70, 88, 146, 235, 434 and 455. The sequences of 2009-2013 isolates have 8 glycosylation sites at positions 50, 58, 63, 68, 88, 146, 235 and 386 except for some cases in 2009 (creation of glycosylation site at positions 50, 68, 386 and deletions at positions 44, 434 and 455) and 2010 (the creation of glycosylation site at position 42) that some changes were detectable in different positions (Table 2). This implies a change in the number and location of the glycosylation sites of 2009 pandemic strain which became dominant in the next years.

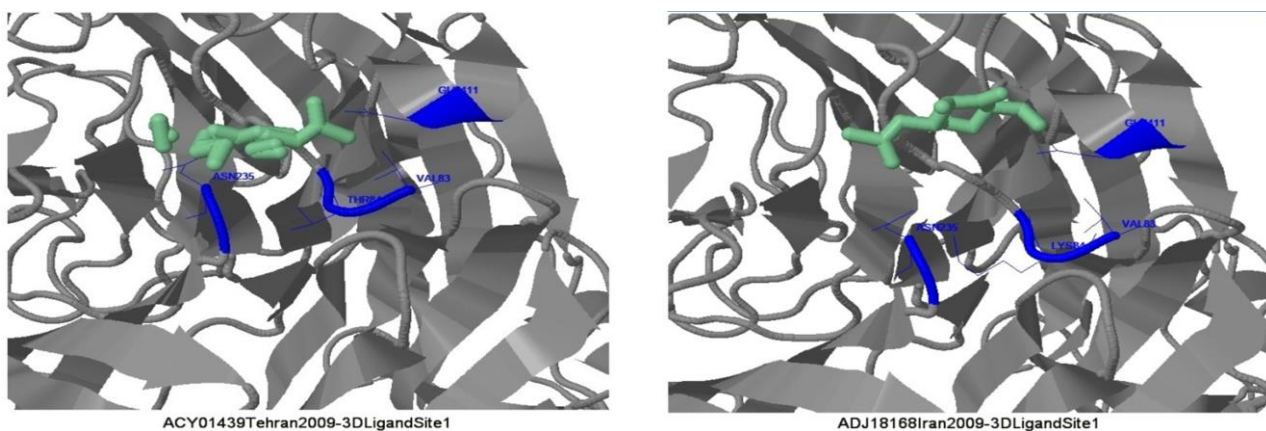
One of the important conserved regions in NA is 150 loop (cavity) which includes residues Gly147 to Asp151 surrounding

the active site of NA and functions as a laterally gate for the drug binding [31, 32]. It was found that this region was conserved from 2006 to 2008. In 2009, the mutation occurred in Val149Ile which became dominant from 2010 to 2013.

Residues within the active site are highly conserved among all of the NA subtypes, including 8 charged and polar residues (Arg118, Asp151, Arg152, Arg225, Glu277, Arg293, Arg368 and Tyr402). They interact with the substrate directly at the catalytic site. The structure of catalytic site is stabilized and supported by the conserved framework residues (Glu119, Arg156, Trp179, Ser180, Asp/Asn199, Ile223, Glu228, His275, Glu278, Asn295 and Glu425) [33, 34]. It was found that the catalytic site residues were conserved from 2006 to 2013 and the framework residues were also conserved except for some cases in 2009 that H275Y mutation, which is the most common molecular marker of NA drug resistance, has occurred.

Besides monitoring of the naturally occurring changes in NA and its respective genes, monitoring of mutations especially those associated with resistance to the NA inhibitors is necessary. Antigenic changes and mutations in the gene of NA protein often cause resistance to the available anti-influenza drugs in influenza viruses. Zanamivir and Oseltamivir that inhibit the function of NA are used as antiviral drugs for the prevention and treatment of influenza [35]. As mentioned above, H275Y mutation is the most prevalent molecular marker of drug resistance to Oseltamivir that occurs in or near the active site of NA and decrease the sensitivity of NA and virus to this drug [36]. We found that this mutation was occurred in 3 cases in 2009. Other mutations associated with the resistance to anti-NA drugs were also inspected; however none of them was observed (Table 3).

Molecular and phylogenetic analyses were also performed to reveal the relationship of changes of each year's isolates to the vaccine strains (Table 4). It was found that all changes and the obtained results of each year were in compliance with and were observed in the standard vaccine strain of that year and the seasonal isolates turned out to be closely related to the corresponding vaccine strains. This means that studying and identifying the antigenic changes of the virus in each year can be used for the prediction of changes in future and designing new vaccines or anti-influenza drugs.



**Fig. 3.** Predicted 3D structure of ligand binding site of 2009 distinct sequences. 3DLigandSite visualization of the protein structure with the residues which form part of the binding site are colored blue. The ligands that form the cluster used for the prediction are also displayed as wireframes.

**Table 2.** Changes (creation/deletion) in glycosylation sites in 2009 and 2010.

Year	Positions						
	42	44	50	68	386	434	455
2009	-	NNT→NQI	NQR→NQS	NIN→NIS	DFS→NFS	NTT→NTI	NWS→GWS
2010	NQN→NQS	-	-	-	-	-	-

**Table 3.** Mutations associated with resistance to anti-NA drugs.

Year	Mutations associated with resistance to anti-NA drugs							
	E19V	R152K	D199N	I223V/R	S247N	H275Y	R293K	N295S
2006	×	×	×	×	×	×	×	×
2007	×	×	×	×	×	×	×	×
2008	×	×	×	×	×	×	×	×
2009	×	×	×	×	×	☑	×	×
2010	×	×	×	×	×	×	×	×
2011	×	×	×	×	×	×	×	×
2012	×	×	×	×	×	×	×	×
2013	×	×	×	×	×	×	×	×

☑: Observed    ×: Not observed

**Table 4.** Standard strains used in influenza vaccine from 2006 to 2013

Year	Standard Vaccine Strain	gi Number
2006-2007	A/New Caledonia/20/99(H1N1)	397776409
2007-2008	A/Solomon Islands/03/2006(H1N1)	156992241
2008-2009	A/Brisbane/59/2007(H1N1)	357016821
2009-2010	A/Brisbane/59/2007(H1N1)	357016821
2010-2011	A/California/07/2009(H1N1)	758899356
2011-2012	A/California/07/2009(H1N1)	758899356
2012-2013	A/California/07/2009(H1N1)	758899356
2013-2014	A/California/07/2009(H1N1)	758899356

## DISCUSSION

The selection of influenza vaccine strains and the development of new vaccines have been always a demanding task because of the antigenic changes of influenza viruses. Due to continuous and high mutation rate in NA, studying the changes in important sites of NA in the available sequences of different virus strains in a specified time period using conventional bioinformatics methods, can provide useful and important information about the trend of changes for the virologists and the public health authorities. Such insights can help to monitor the emergence of new influenza variants and to detect new antigenic changes occurring in NA protein which are important for updating the vaccine composition.

In this study, by computational investigation of sequences isolated between 2006 to 2013 from Iran and the inspection of the amount of phylogenetic similarity amongst them and by determination of the binding sites and relevant ligands of NA antigen, it was found that the amounts of similarities in sequence alignment analyses in 2006- 2008 in one side and in 2010-2013 from the other side was significant while in 2009, a set of two series of sequences of each of the two groups was detectable. The sequences of 97 NA antigens of human H1N1 influenza virus from 2006 to 2013 in Iran were downloaded. Sequences were aligned and the phylogenetic trees and percent similarity/divergence matrices were created to determine the conserved and variable regions and the similarities and differences in them. Moreover, changes/mutations in the predicted ligand binding sites were also inspected. Significant differences were observed in this period. It was found that 2009

was an important transition year which almost classified the isolates into two groups, namely the 2006-2008 and the 2010-2013 isolates while each group had similar changes/mutations. The 2009 phylogenetic tree was also divided into two distinct branches (due to the occurrence of pandemic) with two groups of isolates. One group was similar to 2006-2008 isolates in terms of amino acid properties and changes/mutations and the other group was similar to 2010-2013 isolates. This shows the importance of changes in the genomic structure of the virus in 2009 which were made possibly during mutations in the genome and likely by the occurrence of frame shifts or gene shuffling, leading to different sequences and a new group. Therefore, it can be assumed this new group and its relevant mutations became dominant and stable in the following years. The 2009 H1N1 pandemic strain is used as a standard vaccine strain from 2010 onwards and mutations associated with anti-NA drug resistance were observed in 2009 strains.

The results of our study are consistent with the results of others in terms of changes in number and location of glycosylation sites, mutations/conservation in important locations such as 150 loop, catalytic sites (NA active sites), framework residues and inspection of mutations associated with resistance to anti-NA drugs. The evaluations of changes in the predicted ligand binding sites, and the type and number of ligands were done for the first time in our study which indicated a significant change in ligand binding site of the 2009 isolates. This also emphasized the importance of 2009 as a transition year.

In conclusion, our findings suggest that the 2009 H1N1 pandemic strain might continue to mutate in its NA gene and create new dominant strains with the potential to cause pandemics in the future. Therefore, conduction of this study was important to assess the extent of diversity of the circulating

strains and to analyze the amino acid changes and the tertiary structures which can change the function of the virus and its virulence or pathogenicity. This also emphasized the importance of continuous monitoring of influenza A H1N1 virus strains (specially the A (H1N1) pdm09 strain) in different geographical areas in order to get better selection of the vaccine strain for the upcoming seasons. To the best of our knowledge, this study was done for the first time in Iran and the obtained results can be potentially useful for the epidemiological studies, as well as for better understanding of the antigenic evolution of H1N1 influenza viruses and the prediction of changes in the future that together can help to select the proper vaccine strains or to design new effective anti-influenza drugs.

## ACKNOWLEDGEMENT

Authors would like to thank Dr. Dariush Norouzian, Head of Pilot NanoBiotechnology Department of Pasteur Institute of Iran for his kind supports and advices.

## CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

## REFERENCES

- Treanor JJ. Influenza Virus. In: Mandell GL, Bennett JE, Dolin R, eds. Principles and Practice of Infectious Diseases. Philadelphia: Elsevier/Churchill Livingstone. 2005; (6th. ed.) :2060-85.
- Gürtler L. Virology of Human Influenza. In: Influenza Report. Kamps BS, Hoffmann C, Preiser W. Flying Publisher, Wuppertal. 2006.
- WHO, B. "WHO Memorandum. A revised system of influenza virus nomenclature". 1980; (58): 585-91.
- Tong S, Li Y, Rivailier P, Conrardy C, Castillo DA, Chen LM et al. A distinct lineage of influenza A virus from bats. *Proc Natl Acad Sci U S A*. 2012;109(11):4269-74. doi:10.1073/pnas.1116200109.
- Skehel JJ, Wiley DC. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annu Rev Biochem*. 2000;69:531-69. doi:10.1146/annurev.biochem.69.1.531.
- Gottschalk A. Neuraminidase: the specific enzyme of influenza virus and *Vibrio cholerae*. *Biochim Biophys Acta*. 1957;23(3):645-6.
- Matrosovich MN, Matrosovich TY, Gray T, Roberts NA, Klenk HD. Neuraminidase is important for the initiation of influenza virus infection in human airway epithelium. *J Virol*. 2004;78(22):12665-7. doi:10.1128/JVI.78.22.12665-12667.2004.
- Palese P, Schulman JL, Bodo G, Meindl P. Inhibition of influenza and parainfluenza virus replication in tissue culture by 2-deoxy-2,3-dehydro-N-trifluoroacetylneuraminic acid (FANA). *Virology*. 1974;59(2):490-8. doi:10.1016/0042-6822(74)90458-9.
- Palese P, Tobita K, Ueda M, Compans RW. Characterization of temperature sensitive influenza virus mutants defective in neuraminidase. *Virology*. 1974;61(2):397-410.
- Webster RG, Laver WG, Kilbourne ED. Reactions of antibodies with surface antigens of influenza virus. *J Gen Virol*. 1968;3(3):315-26. doi:10.1099/0022-1317-3-3-315.
- Wagner R, Matrosovich M, Klenk HD. Functional balance between haemagglutinin and neuraminidase in influenza virus infections. *Rev Med Virol*. 2002;12(3):159-66. doi:10.1002/rmv.352.
- Moscona A. Neuraminidase inhibitors for influenza. *N Engl J Med*. 2005;353(13):1363-73. doi:10.1056/NEJMr050740.
- Sheu TG, Deyde VM, Okomo-Adhiambo M, Garten RJ, Xu X, Bright RA et al. Surveillance for neuraminidase inhibitor resistance among human influenza A and B viruses circulating worldwide from 2004 to 2008. *Antimicrob Agents Chemother*. 2008;52(9):3284-92. doi:10.1128/AAC.00555-08.
- World Health Organization, WHO Global Technical Consultation: global standards and tools for influenza surveillance (2012) Available from: [http://www.who.int/influenza/resources/documents/technical\\_consultation/en/index.html](http://www.who.int/influenza/resources/documents/technical_consultation/en/index.html) (accessed 20 February 2013).
- Group WHO, Ampofo WK, Baylor N, Cobey S, Cox NJ, Daves S et al. Improving influenza vaccine virus selection: report of a WHO informal consultation held at WHO headquarters, Geneva, Switzerland, 14-16 June 2010. *Influenza Other Respir Viruses*. 2012;6(2):142-52, e1-5. doi:10.1111/j.1750-2659.2011.00277.x.
- Hayden FG. Antivirals for influenza: historical perspectives and lessons learned. *Antiviral Res*. 2006;71(2-3):372-8. doi:10.1016/j.antiviral.2006.05.016.
- Ison MG. Antivirals and resistance: influenza virus. *Curr Opin Virol*. 2011;1(6):563-73. doi:10.1016/j.coviro.2011.09.002.
- Colman PM. Influenza virus neuraminidase: structure, antibodies, and inhibitors. *Protein Sci*. 1994;3(10):1687-96. doi:10.1002/pro.5560031007.
- Colman PM. Neuraminidase inhibitors as antivirals. *Vaccine*. 2002;20 Suppl 2:S55-8.
- Colman PM. New antivirals and drug resistance. *Annu Rev Biochem*. 2009;78:95-118. doi:10.1146/annurev.biochem.78.082207.084029.
- Kim CU, Lew W, Williams MA, Liu H, Zhang L, Swaminathan S et al. Influenza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: design, synthesis, and structural analysis of carbocyclic sialic acid analogues with potent anti-influenza activity. *J Am Chem Soc*. 1997;119(4):681-90.
- Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symposium Series No. 41*. 1999: 95- 98.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol*. 2011;28(10):2731-9. doi:10.1093/molbev/msr121.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol*. 2013;30(12):2725-9. doi:10.1093/molbev/mst197.
- Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res*. 2003;31(13):3381-5.
- Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc*. 2015;10(6):845-58. doi:10.1038/nprot.2015.053.
- Wass MN, Kelley LA, Sternberg MJ. 3DLigandSite: predicting ligand-binding sites using similar structures. *Nucleic Acids Res*. 2010;38(Web Server issue):W469-73. doi:10.1093/nar/gkq406.
- Helenius A, Aebi M. Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem*. 2004;73:1019-49. doi:10.1146/annurev.biochem.73.011303.073752.
- Chutinimitkul S, Chieochansin T, Payungporn S, Samransamruajkit R, Hiranras T, Theamboonlers A et al. Molecular characterization and phylogenetic analysis of H1N1 and H3N2 human influenza A viruses among infants and children in Thailand. *Virus Res*. 2008;132(1-2):122-31. doi:10.1016/j.virusres.2007.11.007.
- Xu X, Zhu X, Dwek RA, Stevens J, Wilson IA. Structural characterization of the 1918 influenza virus H1N1 neuraminidase. *J Virol*. 2008;82(21):10493-501. doi:10.1128/JVI.00959-08.
- Le L, Lee E, Schulten K, Truong TN. Molecular modeling of swine influenza A/H1N1, Spanish H1N1, and avian H5N1 flu N1 neuraminidases bound to Tamiflu and Relenza. *PLoS Curr*. 2009;1:RRN1015. doi:10.1371/currents.RRN1015.
- Colman PM, Varghese JN, Laver WG. Structure of the catalytic and antigenic sites in influenza virus neuraminidase. *Nature*. 1983;303(5912):41-4.
- Ives JA, Carr JA, Mendel DB, Tai CY, Lambkin R, Kelly L et al. The H274Y mutation in the influenza A/H1N1 neuraminidase active site following oseltamivir phosphate treatment leave virus severely compromised both in vitro and in vivo. *Antiviral Res*. 2002;55(2):307-17.
- Dharan NJ, Gubareva LV, Meyer JJ, Okomo-Adhiambo M, McClinton RC, Marshall SA et al. Infections with oseltamivir-resistant influenza A(H1N1) virus in the United States. *JAMA*. 2009;301(10):1034-41. doi:10.1001/jama.2009.294.