**Original Article**

**Phylogenetic Analysis of Hemagglutinin Gene of H9N2 Avian Influenza Viruses Isolated from Chicken in Iran in 2010-2011: Emerging of a New Subgroup**

Norouzian H1*, Gholami SJ2, Vasfi-Marandi M2

1. Department of Clinical Sciences, Faculty of Veterinary Medicine, Lorestan University, Khorram Abad, Iran.
2. Department of Clinical Sciences, Faculty of Veterinary Medicine, University of Tehran, Tehran, Iran.

**Abstract**

**Background and Aims:** Hemagglutinin (HA) protein of Avian Influenza (AI) plays an essential role in the virus pathogenicity. AI H9N2 subtype causes significant economic loss in broiler and layer in poultry farms in Iran. AI viruses have a great involvement in evolutionary changes at nucleotide and amino acid levels and vaccines could induce faster rates of such changes. Up-dated understanding of the genetic changes of AI viruses circulating in Iran is necessary for controlling AI.

**Materials and Methods:** Sequence analysis and phylogenetic study of the HA gene of three H9N2 subtype of AI isolates in Iran in 2010-2011 were studied.

**Results:** Cleavage site of the Iranian 2010-2011 isolates possessed a different motif. Amino acid residue at position 226 at receptor binding site in these isolates was Leucine, which was similar to human viruses. The epitopes for HA showed a great variation related to the year of isolation. According to phylogenetic analysis, Iranian isolates were divided into two main subgroups. But, viruses isolated in this study formed a third minor subgroup. Degree of homology between the 2010-2011 isolates and former Iranian isolates was significantly low.

**Conclusion:** The results revealed that HA of new Iranian AI H9N2 isolates have undergone extensive genetic changes. Definitely, continuous monitoring of genetic changes is a useful tool for updating control strategy for AI outbreak in Iran.

**Keywords:** Avian Influenza; H9N2; Hemagglutinin; Phylogenetic analysis; Iran

**Introduction**

Avian influenza (AI) virus is a member of the family orthomyxoviridae. These viruses contain eight single-stranded RNA segments, which encode ten proteins, including two surface glycoproteins, hemagglutinin (HA) and neuraminidase (NA) (1). Up to now, 16 HA and 9 NA subtypes have been identified (2). HA protein plays an essential role in virus binding to sialic acid receptor on the host cell surface and fusion of virus to endosomal membrane, which facilitates entry of virus into the host cell (1). AI H9N2 subtype typically causes mild clinical signs in birds. However, significant mortality in broilers and significant decrease in egg production have been reported due to H9N2 subtype, particularly in co-infection with other respiratory agents (3, 4). H9N2 influenza viruses have become widespread in the poultry farms of the world (4-6). According to the phylogenetic and antigenic analysis, Asian H9N2 viruses are classified as three major...
sublineages represented by A/Quail/Hong Kong/G1/97 (G1 group), A/Chicken/Hong Kong/G9/97 (G9 group), and A/chicken/Korea/38349-p96323/96 (Korean group) (7).

To control AI infections in poultry, inactivated vaccines have been widely used in Iran, Pakistan and China (8-10). However, it is revealed that vaccines could induce faster rates of antigenic drift in influenza viruses due to providing a positive selective pressure (9). On the other hands, co-circulation of AI H9N2 viruses with other subtypes, including H5 and H7, may lead to the genetic reassortment in the viruses (11). For these reasons, up-dated understanding of the genetic changes of AI viruses circulating in Iran is important for controlling AI. In this study, the HA gene of three AI H9N2 strains isolated from poultry chicken farms in Iran during 2010 and 2011, as well as, other reported Iranian H9N2 isolates were genetically analyzed and nucleotide changes were evaluated.

Methods

Viruses
Three strains, including: A/Chicken/Iran/N101/2011(H9N2), A/Chicken/Iran/N102/2011(H9N2) and A/Chicken/Iran/EBGV-88/2010 were isolated from three broiler farms affected by respiratory disease complex (RDC). Initial isolation and subtype identification of the virus was determined by standard methods (12) and PCR tests. All the other H9N2 viral sequences of hemagglutinin genes were present in the GenBank database.

Cell transfection
COS-7 and BHK-21 cell lines were received from the National Cell Bank of Iran (NCBI) affiliated with the Pasteur Institute of Iran, cultured and maintained in Dulbecco’s modified Eagle’s Medium (DMEM, Gibco) supplemented with 10% fetal bovine serum (FBS, Gibco). Briefly, one day before transfection, cells from confluent culture were transferred to 24 well plate (Nunc, Roskilde, Denmark) containing 10^5 cells per well. The cells were incubated for 24 h at 37°C in a humidified incubator containing 5% CO2 to achieve approximately 70-75% confluency on the day of transfection. Each well of cells was transfected with 0.8μg of plasmid DNA encoding the HA1 gene using lipofectamin 2000 reagent (Invitrogen, Carlsbad, CA; Cat. No. 11668-027) according to the manufacture’s instruction. After 48h incubation at 37°C, cells were examined by fluorescence microscope.

RNA Extraction and RT-PCR
Viral RNA was extracted from allantoic fluid by RNX plus kit (Cinnagen, Iran), according to the manufacturer instruction. Reverse transcription was done by using oligonucleotide influenza universal primer, Uni12. Amplification of the hemagglutinin full-length gene was carried out by PCR as described previously (13) using two pairs of specific primers as follows:

P-HA1-F: 5’-AGC AAA AGC AGG GG-3’ and  P-HA1-R: 5’-GTG YCC ATA CCA TGG RGC-3’ (1 to 832),
P-HA2-F: 5’- ACC AGG TCA GAC ATT GCG AG-3’ and P-HA2-R: 5’- AGT AGA AAC AAG GGT GTT TTT GC-3’ (766-1742).

Gene Sequence
PCR products were subjected to electrophoresis in a 1% (w/v) agarose gel and DNA fragments of the expected length were extracted and purified with GeneJET™ Gel Extraction kit (Fermentas, Canada, catalogue#k0691). The purified DNA fragments were cloned into pTZ57R/T cloning vector (Fermentas, Canada, catalog#k1213). Three clones of each fragment were sequenced using M13 forward and M13 reverse (promega, USA) at Sequetech Co. Ltd., USA. The sequences identified of at least two clones were kept for further usage.

Sequence Analysis and Phylogenetic Study
Pairwise sequence alignments were performed using the Clustal W alignment algorithm. Sequence similarity and phylogenetic relationship of different H9N2 subtype viruses was performed with MegAlign program (DNASTAR Inc., Madison, WI, USA).

Nucleotide Sequence Accession Numbers
The sequences determined in this study are available in the GenBank under accession numbers: JQ970436, JQ970437 and JN646748.

### Results

#### Analyses of the hemagglutinin sequence of the Iranian H9N2 isolates

In this study, 1742 nucleotides of the HA gene of the three isolates (A/Chicken/Iran/N101/2011(H9N2), Chicken/Iran/N102/2011(H9N2) and A/Chicken/Iran/EBGV-88/2010) were sequenced and amino acid sequences of the HA were deduced from the nucleotide sequence.

HA cleavage site motif sequence of the most Iranian isolates was: –R-S-R*S-G-. Amino acids of cleavage site of the isolates sequenced in this study possessed a –K-S-R*S-G- motif. Some other substitutions have happened in some Iranian isolates. Four isolates harbor –R-S-N-R*S-G- at the same position (Table 1).

In the HA protein, the amino acids at positions 183, 190, 226, 227 and 228 (H3 numbering) are major receptor binding sites (RBS). Amino acid residue at position 226 in isolates sequenced in this study, was Leucine. In Iranian H9N2 viruses, the number of those possessed the human H9N2 virus-like motif, 226Leu, has increased after 2004. There is a substitution at position 227, Glu→Ile, in almost all Iranian H9N2 viruses isolated after 2004, including three viruses sequenced in this study. In one of the viruses sequenced in this study (A/Chicken/Iran/N101/2011), amino acid residue Thr replaced Ala at position 190. In Iranian isolates, substitutions in this position have happened, which Ala replaced by Thr, Ile and Val in some recent isolates (Table 1).

Potential glycosilation site (PGS) motifs –N-X-T/S- (X may be any amino acid except proline)

### Table 1. Amino acid substitutions in the cleavage site, Receptor binding site and PGS of H9 HAs of Iranian H9N2 isolates.

<table>
<thead>
<tr>
<th>virus</th>
<th>Accession no.</th>
<th>Cleavage site</th>
<th>190</th>
<th>226</th>
<th>227</th>
<th>105</th>
<th>166</th>
<th>206</th>
<th>218</th>
<th>283</th>
<th>492</th>
<th>551</th>
</tr>
</thead>
<tbody>
<tr>
<td>11T/99</td>
<td>AF508558</td>
<td>RSSR</td>
<td>A</td>
<td>Q</td>
<td>Q</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>B308A/04</td>
<td>EF063728</td>
<td>RSSR</td>
<td>A</td>
<td>L</td>
<td>Q</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>EBGV-88/10</td>
<td>JN646748</td>
<td>KSSR</td>
<td>T</td>
<td>L</td>
<td>I</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>N101/11</td>
<td>JQ970436</td>
<td>KSSR</td>
<td>T</td>
<td>L</td>
<td>I</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>N102/11</td>
<td>JQ970437</td>
<td>KSSR</td>
<td>A</td>
<td>L</td>
<td>I</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>G1/97</td>
<td></td>
<td>RSSR</td>
<td>E</td>
<td>L</td>
<td>Q</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

AF156378

### Table 2. Amino acid substitutions in the antigenic sites of H9 HAs of Iranian H9N2 isolates.

<table>
<thead>
<tr>
<th>virus</th>
<th>Accession no.</th>
<th>108</th>
<th>134</th>
<th>154</th>
<th>312</th>
<th>324</th>
<th>431</th>
</tr>
</thead>
<tbody>
<tr>
<td>11T/99</td>
<td>AF508558</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>B308A/04</td>
<td>EF063728</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>EBGV-88/10</td>
<td>JN646748</td>
<td>+</td>
<td>+</td>
<td>.T.IG.</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>N101/11</td>
<td>JQ970436</td>
<td>+</td>
<td>+</td>
<td>.I.IG</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>N102/11</td>
<td>JQ970437</td>
<td>+</td>
<td>+</td>
<td>.I.IG</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>
Table 3. Homology of the nucleotide sequences of HA of Iranian isolates belonging to different sublineages.

<table>
<thead>
<tr>
<th>sublineage</th>
<th>Similarity between viruses (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First sublineage</td>
</tr>
<tr>
<td></td>
<td>Max</td>
</tr>
<tr>
<td>First</td>
<td>100</td>
</tr>
<tr>
<td>Second</td>
<td>-</td>
</tr>
<tr>
<td>Third1</td>
<td>-</td>
</tr>
</tbody>
</table>

in the HA molecule of viruses sequenced in this study and most other Iranian H9N2 isolates were seven sites at positions 29, 105, 141, 298 and 305 in HA1 and 492 and 551 in HA2 region, C-terminal to the membrane anchor. However, some of these PGSs have been lost in some isolates. On the other hand, some Iranian H9N2 isolates possessed new PGSs at positions 158, 198, 210 and 275 due to point mutations (Table 1).

The epitopes for HA showed a great variation related to the year of isolation. The epitope (SFYRNMRLWL) from positions 154-162 and the epitope (KAIGLRNV) from positions 324-331 that were mostly absent in viruses isolated between 1998 and 2004, appeared in H9N2 viruses after 2004. Another epitope with high variation at position 312-320, which in most of the viruses isolated before 2004 was (GNCPKYVRV) and after that year it has changed to (GTCPKYIGV). This epitope in two of isolates sequenced in this study was changed to (GICPKYIGV). Some of the epitopes (position 431-439 (AYNAELLVL) and position 134-142 (IFPDTIWNV)) are present in all Iranian H9N2 isolates. In the epitope (CYPGNVENL), positions 108-116, there is an amino acid substitution at position 113 (Val113Ile) in a few Iranian isolates (table 2).

Phylogenetic analyses
The phylogenetic relationship between Iranian isolates and H9N2 viruses isolated in various part of Middle East countries were established. The analysis was based on the nucleotide sequence of full segment of HA. All of the Iranian isolates fell into the same main group (G1 sublineage), as the viruses isolated in Pakistan, Saudi Arabia, Japan and Israel (Fig 1a). The Iranian isolates can be further divided into two main subgroups. The first subgroup (11-T like sublineage) included viruses that were isolated between 1998 and 2004. The second subgroup comprised the strains isolated after 2004. Though, some exceptions were seen in both subgroups. Interestingly, three viruses isolated in this study formed a third minor subgroup (Fig 1b). Other Middle East isolates also clustered in two subgroups, with some exceptions, according to the country and time of isolation. Three viruses isolated in this study, placed near the new Pakistan isolates in a different subgroup (Fig 1a).

Analysis of the Iranian isolate nucleotide sequences showed that the difference within the first subgroup was relatively higher than the second subgroup: viruses belonging to the first subgroup exhibited 91.8-100% (average 96.4%) identity and viruses belonging to the second subgroup exhibited 93.7-99.9% (average 97.4%) identity. Three viruses of the third minor subgroup, the isolates sequenced in this study, exhibited higher identity 93.7-99.9% (average 97.4%) identity. Three viruses of the third minor subgroup, the isolates sequenced in this study, exhibited higher identity 98.7-99.1% (average 98.9%) (Table 3). Interestingly, these isolates also showed relatively high identity to some of the new Pakistan isolates (95%). Degree of homology between these new isolates and former Iranian isolates is relatively low (90.5% and 92.8% for first and second subgroup, respectively). The percent of homology of isolates sequenced in this study with prototype isolate of G1-like sublineage and human H9N2 isolates was 90.9% (Table 3).

Genetic variability amongst the Iranian isolates
In this study, we analyzed 1682 nucleotides encoding 560 amino acids (1-560). Of these 560 amino acids, 537 were characterized as highly conservative amongst the Iranian isolates...
isolates, whereas the other 23 amino acids showed diverse levels of variability. The subgroup specificity of the substitutions was observed in the following positions: amino acids A_{15}, S_{101}, A_{126}, A_{150}, S_{156}, S_{166}, S_{168}, S_{183}, D_{216}, N_{218}, M_{224}, Q_{235}, G_{271}, N_{313}, V_{318}, R_{319}, and V_{327} were typical of the viruses of the first subgroup, whereas T_{15}, P_{101}, S_{126}, S_{156}, N_{158}, N_{166}, L_{168}, D_{183}, N_{216}, D_{218}, L_{224}, I_{235}, E_{271}, T_{313}, I_{318}, G_{319}, and I_{327} were typical of those of the second subgroup.

**Discussion**

Changes in the HA amino acid content at the proteolytic cleavage site (PCS) by either site mutation or insertions of amino acids or both, resulting in an increase in the number of basic amino acids (Arginine and Lysine) at the site provides a template for cleavage by widely distributed furin-like or subtilisin-like endoproteases (11). This is central key to the switch from low to high pathogenicity in H5 and H7 subtypes (14, 15). Nucleotide sequence analysis corresponding to the cleavage site of Iranian isolates revealed no poly-basic motif. Most of the earliest Iranian isolates (since 1998) had the cleavage site motif –RSSR*G–with two basic amino acids, which is rather typical for H9 viruses, including those from Hong Kong (Qa/HK/G1/97 and Dk/HK/Y280/97), Pakistan, Israel, Saudi Arabia and Dubai (16). Some viruses with point mutations of amino acid in the cleavage site (RSNR, RSNK) were isolated on 2007. Then, on 2010 and 2011, all isolated viruses, including viruses sequenced in this study, showed another amino acid substitution (KSSR). This substitution of lysine instead of Arginine at the -4 position of the cleavage site is infrequently observed in H9N2 viruses, but have become fixed in all 2010 and 2011 Iranian and some recent Pakistani isolates (table 1) (17, 18, 19). The significance, if any, of these mutations is not known. Banks et al. (2000) (16) described 15 amino acid motifs of

![Phylogenetic Analysis of Hemagglutinin Gene of H9N2 Avian Influenza Viruses](image-url)
the cleavage sites of the H9 influenza viruses, but none of them had more than two basic amino acids. It should be emphasized, that all H9N2 viruses with two or three basic amino acids in the cleavage site have had low pathogenicity for chickens (19, 18). Some researchers suggested that -RSSR*G- motif is the signature of H9N2 LPAI viruses adapted to the chicken host (16, 18, 20). It is interesting that all of these mutations in Iranian isolates have happened after 2007. So, there may be a relationship between time and happening of mutations in AIVs, as observed in Israeli isolates (19).

AI H9N2 viruses have some capabilities to transmit from bird to human and are a candidate for next human flu pandemic (20). The RBS motif of HA is critical for cellular receptor specificity and determining virus host range (21, 22). Glutamine at position 226 indicates a preferential binding to sialic acid (SA) α2,6-linked to galactose found in avian host and is a major host range determinant for the H2 and the H3 subtypes associated with pandemic human flu infections (23, 24). The substitution of Glu226Leu makes a preferential receptor binding specificity for SA α2,6-linked galactose, allows H9N2 viruses to replicate about 100-fold more efficiently in human cell culture (23, 25). There is a general increasing trend in the proportion of Iranian H9N2 viruses that possessed the human virus-like motif 226Leu, particularly in viruses isolated after 2004 (Table 1) and in recent Pakistan, UAE and Chinese isolates (7, 18, 26). This result shows increasing potential of recent Iranian H9N2 isolates to transmit to human. The residue at position 190 within the RBS influences the affinity of virus binding to the SA receptor. Amino acids Valine, Threonine and Alanine at this position had high, intermediate and low affinity to SA α2, 6-linked galactose receptor, respectively (27). Alanine at this position is more typical of viruses isolated from chickens (28). In Iranian isolates, substitutions in this position (A→T, A→I, A→V) have happened in some recent isolates. In one of the viruses sequenced in this study (N101), amino acid Threonine has replaced Alanine, indicating high affinity for human airway epithelial receptors (Table 1). Nevertheless, the definitive significance of these substitutions is not completely understood and requires detailed structural analysis.

Seven PGSs were present in most of the Iranian H9N2 isolates, including isolates sequenced in this study. However, there were additional or lost PGSs in some of the isolates. Two PGSs at positions 218 and 206, present in representative reference strain of G1-lineage (A/Quail/HK/G1/97) have lost in most of the Iranian isolates (Table 1). The loss or creation of PGSs may represent a selected adaptation of H9N2 viruses transmitted from original host (e.g. Quail) to poultry (14, 29). Besides, these alterations may affect the receptor binding capacity of HA and contribute to the antigenic variations (30, 31). Anyway, the biological significance of these changes in PGSs pattern in the Iranian H9N2 isolates remains to be studied.

A study on evolution of AI H9N2 viruses in the Middle East and Indian sub-continent showed variability in antigenic sites of HA during 1994-2009 (20). Antigenic sites for HA showed a great deal of variation. Observation revealed that epitopes (SFYRNMRWL) from positions 154-162 and (KLAIGLRNV) from positions 324-332 that was absent before 2004 in first subgroup of Iranian isolates, appeared in the second subgroup after 2004 (Table 2). This finding is in agreement with phylogenetic analysis results. In the other epitopes, also, there are some variations between Iranian H9N2 isolates (Table 2). These results suggest that NA epitopes could have evolved due to point mutations and reassortment. Such changes would affect virus antigenicity and pathogenicity of the virus. It is revealed that vaccines could induce faster rates of antigenic drift in human and avian influenza viruses possibly due to inducing a positive selective pressure in the host (32, 33). Besides, there is a concern about that the antigenic differences between the vaccine strain and prevailing H9 viruses might worsen the continued antigenic drift of the viruses. Most of the isolated H9N2 viruses in China after 2003 were antigenically heterologous with the vaccine isolate due to
antigenic drift and commercial inactivated vaccine could not protect chickens against some of the H9N2 viruses in cases of virus shedding or appearing of clinical signs (34). It is better to compare the antigenic properties of HA protein of new H9N2 isolates with vaccine strain and re-evaluate of vaccine efficacy in Iran.

Previous studies on phylogenetic analysis of HA gene have determined that all Iranian H9N2 viruses fell into a special group and came from a single progenitor (13, 20, 35, 36). On the other hands, all Iranian isolates fell into the same main group (G1 sublineage) as the viruses isolated in Pakistan, Saudi Arabia, UAE, India and Israel, indicating they might have originated from a common source (Fig 1a). However, Middle East isolates showed a relationship that may indicate some ecological barriers to virus movement based around geographic location, with isolates grouped according to country of isolation (20). The Iranian isolates can be further divided into two main subgroups (Fig 1b). The first subgroup (11-T like subgroup) viruses isolated mainly between 1998 and 2004 and the second subgroup mainly after 2004, with some exceptions in both subgroups. The isolates sequenced in this study had more homology to some of the recent Pakistan isolates (2005-2008) compared to the former Iranian isolates and in the phylogenetic tree placed closer to new Pakistan isolates than other Iranian isolates. This may indicate appearance of a new sublineage of AI viruses in Iran and Pakistan (fig 1a). Xu et al (2007) demonstrated that antigenic diversity among H9N2 viruses from different location in China correspond with phylogenetic relationships (37). Therefore, the observed differences between Iranian H9N2 isolates in phylogenetic analysis reflect their antigenic differences, which cause important challenges in vaccine efficiency.

Most of the amino acids (96%) of Iranian H9N2 virus HAs were highly conservative, whereas the other amino acids showed diverse levels of variability. Certain amino acid substitutions have become fixed in every subgroup, but many are unique substitutions represented in individual viruses. This high frequency of substitutions in the HA gene of Iranian isolates is a sign of continued selection pressure in the poultry host. It is reasonable that AI vaccination in Iran could enhance this selection pressure and so the rate of the virus adaptation. This could be an important disadvantage of AI vaccination, considering that the AI vaccination cannot prevent occurrence of disease and its economic costs. It should be noted that differences mentioned in this study relate just to the HA gene.

A comprehensive evaluation of differences between Iranian isolates by complete genome sequencing is necessary to understand the probable reassortment of AI H9N2 virus in Iran. The present study highlights the need to continue with influenza virus surveys, to obtain more information that could lead to the prediction and control of future AI H9N2 virus epizooty in Iran.

Acknowledgement

We gratefully acknowledge the excellent technical support provided by Mrs Noroozi. We also thank Lorestan University and University of Tehran for providing part of the financial support.

References


