

## Original Article

# Molecular Identification of Pre-Existing Immunity in Human against H9N2 Influenza Viruses Using HLA-A\*0201 Binding Peptides

Shahsavandi S\*

Razi Vaccine and Serum Research Institute, P.O.Box 31975-148, Karaj, Iran.

### Abstract

**Background and Aims:** The contribution genetic and antigenic diversity of H9N2 influenza viruses in evading from immune responses, cytotoxic T lymphocytes (CTL) epitopes in hemagglutinin (HA) protein restricted by HLA binding peptides was identified.

**Materials and Methods:** Phylogenetic analyses were carried out for all of full length HA and deduced amino acid sequences of H9N2 viruses available in GenBank from their emergence to now. The impact of amino acid substitutions on emergence of antigenic variants was evaluated by using *in silico* prediction of the HLA-A\*0201 binding peptides within HA protein. Potentially changes in structure and antigenic function of HA protein were quantified.

**Results:** The viruses phylogenetically clustered in Middle East and China. HA protein of the Middle East viruses represented a single sublineage, G1-like, while China isolates grouped to various sublineages and genotypes. Despite Middle Eastern viruses, more variations found in CTL epitopes of China isolates correlated to phylogenetic analysis. Quantify and scoring of HA epitopes revealed that structure and antigenic function of the protein was not changed within H9N2 viruses during the previous decades indicating viral recognition by host-specific CTL response was not affected.

**Conclusion:** Using an evolutionary and immunological approach, we showed that substantial levels of immunogenic peptide conservation for H9N2 HA protein was presented by HLA-A\*0201. The potentially pre-existing immunity to H9N2 viruses in human is important for determining the outcome of influenza infection and developing vaccine with a T cell-based component against the public health threat.

**Keywords:** H9N2 influenza virus; CTL epitope; HLA-A\*0201 binding peptide; immunity

### Introduction

H9N2 avian influenza viruses have been isolated in multiple avian species throughout Asia, Europe and Africa (1, 2). The widespread prevalence of H9N2

virus in poultry, continuous co-circulation and mixing of the virus with human H3N2, H5, H7 viruses, and acquiring the typical receptor specificity of human influenza viruses suggest that H9N2 has the potential to cause epidemics in the population (3, 5). Human cases of H9N2 virus infection were first documented in Hong Kong in 1999 and reported in 2003, 2007, 2008, and 2009 (6, 7). Moreover H5N1, H7N9 and H10N8, another novel avian influenza virus that recently caused human death in

\*Corresponding author: Shahla Shahsavandi, PhD.  
Razi Vaccine and Serum Research Institute, P.O.Box 31975-148, Karaj, Iran.  
E-mail: s.shahsavandi@rvsri.ac.ir

China, sharing the highest similarities to H9N2 internal genes (8, 9). The susceptibility of the human to influenza infection strongly depends on the genetic background. The major histocompatibility complex (MHC or HLA in human) is an important host genetic risk factor in viral disease. Specific binding of antigenic peptides to MHC class I molecules is a prerequisite for their recognition by cytotoxic T-cells (CTL). The HLA restriction with CTL plays a major role in the immune response which associated with the host response to the infection. Among the supertype HLA alleles the HLA-A\*0201-restricted epitopes in human influenza virus proteins are highly conserved (10, 11) considering the immunodominant nature of the epitope. The virus-specific CTL recognize both a viral antigen and a MHC restriction molecule resulted in immunologic susceptibility to the viral infection (12).

Vaccination is one of the most effective and cost-benefit interventions that reduce influenza morbidity (13). Results on developing a universal vaccine revealed that the epitopes capable of binding HLA types provides an alternative strategy against influenza infection (14-16). Hemagglutinin (HA) protein of the viruses is a major antigenic site of protective immunity and more variable than other virus proteins. Since 2005 H9N2 viruses have undergone gradual evolution into several lineages and sublineages either by point mutations or extensive reassortments (17-19). The circulating viruses may possess different HA antigenicity and the induced antibodies may fail to neutralize the virus adequately. In the absence of a vaccine or pre-existing immunity, the emergence of a new strain of influenza virus places nearly every member of the population at infection risk (19). One of the major concerns about the ongoing influenza outbreak is that little protection may exist in human population which induced from the pre-existing CD4+ and CD8+T cell responses to the conserved epitopes. In this study, the molecular basis for pre-existing immunity against H9N2 virus in human population that may prime by circulating strains was examined. The *in silico* study was undertaken to screen HA H9N2 sequences as potential

CTL epitopes of HLA-A\*0201 from their emergence to now by using bioinformatics algorithms. The possible differences between HA protein structure and function relationships of the isolates were also explored.

## Methods

### Dataset of HA H9N2 virus and molecular characterization

The full length HA and deduced amino acid sequences of H9N2 viruses available in GenBank database from 1997 to now were downloaded and aligned using ClustalW with default parameters. The phylogenetic tree was constructed using the minimum evolution (ME) analysis with 1000 bootstrap replication by MEGA4 (20). To determine the evolutionary relationships the level of diversity was separately calculated for every aligned amino acid position according to Claude Shannon's information theory (21). The variability at amino acid sites at each alignment position was estimated in which the quantity of information or entropy can be interpreted as a level of diversity. An entropy plot was generated for HA amino acid sequences using BioEdit ver. 7.0.9 (22). To identify topology features of HA protein the amino acid mutation spectra between groups of sequences were compared.

### Prediction of HLA-A-restricted epitopes

The 9 mer sequences corresponding to the HLA-A\*0201 binding peptides within influenza HA protein were predicted by the web-based programs: BIMAS ([http://www.bimas.cit.nih.gov/molbio/hla\\_bind](http://www.bimas.cit.nih.gov/molbio/hla_bind) (23), and SYFPEITHI (<http://www.syfpeithi.de/Scripts/MHCServer.dll/EpitopePrediction.htm> (24).

### Protein structure and antigenicity prediction

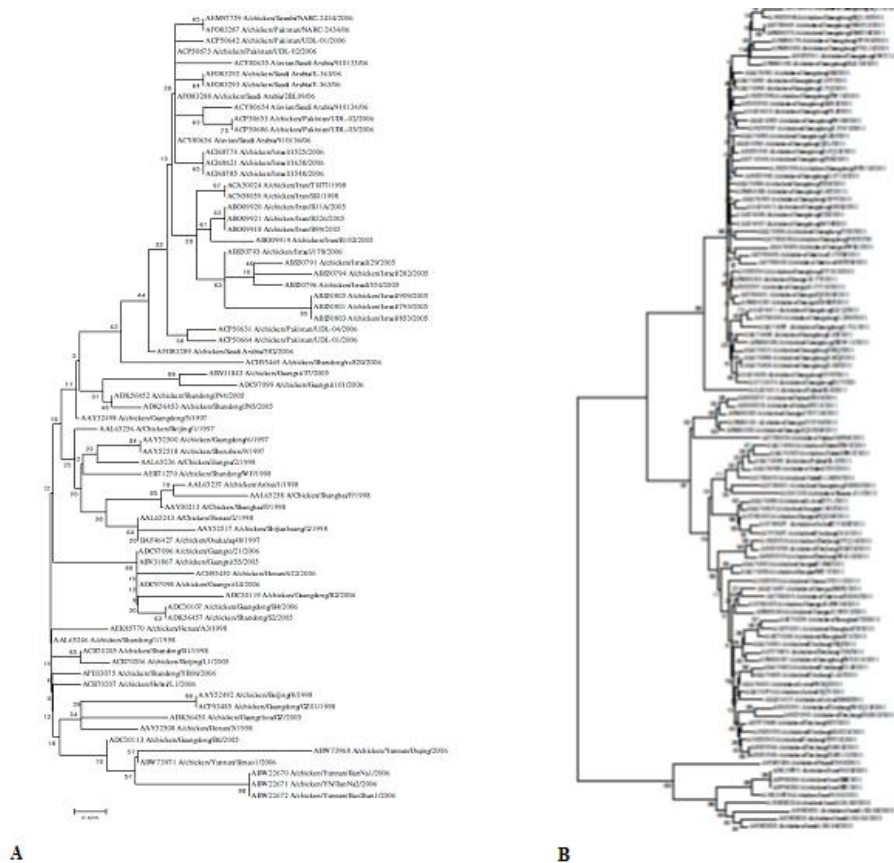
In case that HA sequences show differences, the secondary structure was predicted by using PSIPRED tool (<http://bioinf.cs.ucl.ac.uk/psipred>). The antigenic sites within HA proteins was predicted by Immune Epitope Database (IEDB) analysis server (<http://tools.immuneepitope.org/tools/bcell/tuto>

rial.jsp). The antigenic epitopes were determined using Kolaskar and Tongaonkar antigenicity prediction methods based on physico-chemical properties of amino acid residues (i.e. hydrophilicity, accessibility and flexibility) with about 75% accuracy (25). The impact of amino acid changes on HA structure and stability was evaluated using QMEAN4 global score (<http://swissmodel.expasy.org>) (26). Physico-chemical properties of the HLA-binding peptide sequences include the instability index, theoretical pI, aliphatic index and grand average of hydropathicity (GRAVY) were computed by ProtParam (<http://expasy.org/tool/protparam>).

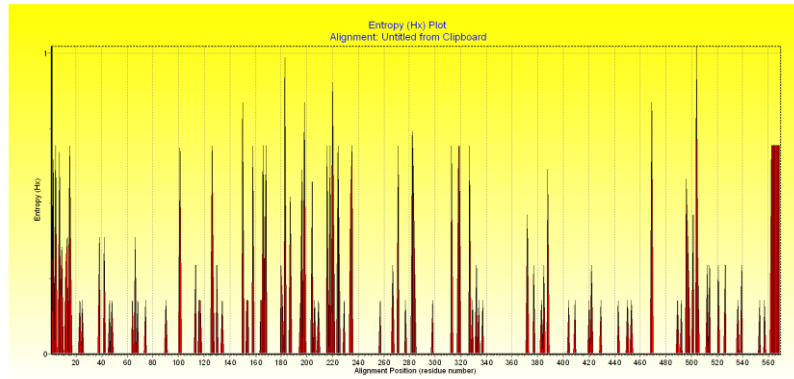
### Results

The HA viral gene segments of Asian H9N2

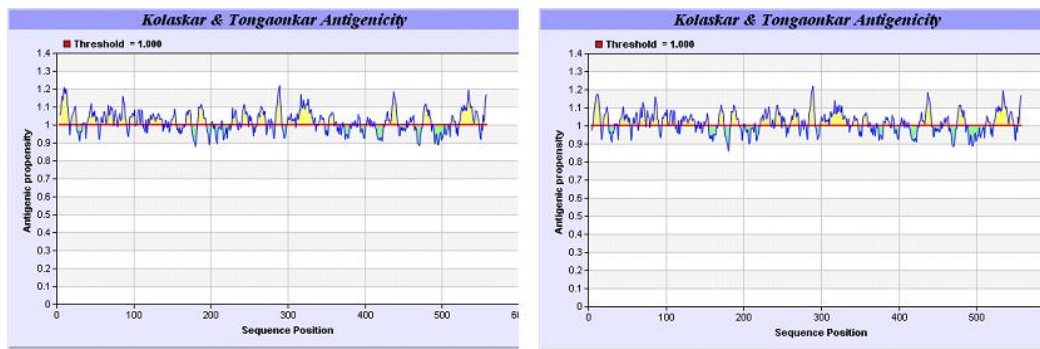
viruses were aligned and analyzed phylogenetically. Similar to our previous study (18) the sequences were divided into two periods 1997-2005 and 2006-now. The phylogenetic analysis of HA protein showed that Middle East isolates shared 96.3%–99.8% identity to each other. The China isolates divided into distinct lineages different from those of Middle East with 94.2%–97.7% identity. Wide distribution of H9N2 influenza viruses over distinct lineages occurred in China, while only one lineage was dominant in Middle Eastern isolates. A close relatedness was found when the phylogenetic tree constructed based on amino acids sequences (Figure 1) indicating synonymous nucleotide substitutions occurred in these sequences. So information entropy to analysis of synonymous mutations in the sequences and plotted panel



**Fig. 1.** The evolutionary history of hemagglutinin protein of H9N2 influenza viruses was inferred using the Minimum Evolution method. Phylogenetic analyses were conducted in MEGA4. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test [1000 replicates] is shown next to the branches. All positions containing gaps and missing data were eliminated from the dataset. A) The sequences of H9N2 viruses isolated in 1997-2005 and B) the sequences isolated in 2006-now.



**Fig. 2.** The entropy plot of hemagglutinin gene of H9N2 influenza viruses. High peaks show the more variations in these sequences.



**Fig. 3.** Antigenicity prediction plot of HA protein using Kolaskar-Tongaonkar method. Regions with antigenic propensity scale more than 1 are predicted as antigenic regions. Left: H9N2 isolate 1998, Right: H9N2 isolate 2012.

was applied (Fig. 2). The peaks in entropy indicate synonymous substitutions in the related nucleotide sequences. Those amino acid residues with an entropy value between 0 and -0.40 were identified as highly conserved and more negative value indicated more diversity. The entropy values of HA amino acid sequence were computed -0.32 and -0.26

for Middle East and China isolates, respectively. The value indicated no more variability in Asian H9N2 HAs, however the rate of diversity in China isolates was faster than the others.

The peptide sequence between HA1/HA2 subunits, specific amino acids residues at the receptor binding site, and the presence or

**Table 1:** The QMEAN4 score consists of four function terms estimated for homology modeling of HA H9N2 influenza viruses

Scoring function term	Middle East isolate		China isolate	
	1997-2005	2006-2011	1997-2005	2006-2011
C-beta interaction energy	-109.85	-99.66	-114.10	-105.41
AA-atom pairwise energy	-6526.56	-6127.73	-6644.72	-6578.47
Solvent energy	-20.53	-17.27	-20.86	-20.73
Torsion angle energy	-112.34	-101.14	-105.99	-101.60

**Table 2:** Impact of amino acid substitutions on physicochemical properties of DINRTFKPI motif within HA H9N2 viruses

Motif	Theoretical pI	Instability index	Aliphatic index	Grand average of hydropathicity
DINRTFKPL	8.75	-2.48	86.67	-0.733
D/EINRTFKPL	6.07	-2.48	43.33	-1.622
DIN/DRTFKPL	5.96	-10.86	86.67	-0.733
DINRT/VFKPL	8.75	-16.19	118.89	-0.189
EINRTFKPL	8.85	-18.92	86.67	-0.733

absence of glycosylation sites near the receptor binding site are molecular determinants of pathogenicity and virulence in HA molecule. All isolates shared an identical amino acid sequence of R/KSSR at HA cleavage site, a signature of low pathogenic avian influenza virus adapted to chicken. Seven potential glycosylation sites (with the N-X-T/S motif) were present in the HA of the isolates including five sites in HA1 and two in HA2. The HA sequence alignment data indicate that HA gene of H9N2 viruses circulating in Middle East during the two past decades did not undergo extensive diversity, while China isolates developed adaptive evolution from 1997 to 2005 and its rate was slow.

The contribution of epitope variations in HA diversity was assessed by identification of immunogenic H9N2-specific CTL epitopes restricted to the HLA-A\*0201 supertype. Despite Middle Eastern viruses that represented similar CTL epitopes, variation in China isolates was observed. The variable epitopes included changes in positions 5, 10, 50, 182, 268, 280, 305, 414, and 534 were predicted in recent China H9N2 isolates. Most of them were located at HA1 polypeptide outside the major antigenic sites. Among the conserved epitopes, the original <sup>216</sup>DINRTFKPL<sup>224</sup> motif showed D/EINRTFKPL, DIN/DRTFKPL, and DINRT/VFKPL alternations in 2005 Middle East viruses which continued to now. The motif was changed to EINRTFKPL in 2005 China viruses which replaced by NLDRTFKPL in recent isolates. The CTL escape mutation as amino acid substitutions

within epitopes restricted by HLA was not detected.

Prediction of secondary structure of a protein aims to determine the protein's stability and function. Analysis of the HA protein secondary structures for 1997-2005 and 2006-now H9N2 isolates showed an exact similarity between them indicating the location and percent of random coils,  $\alpha$ -helixes and  $\beta$ -sheets structures were not affected by amino acid substitutions. Also the substitutions in HA protein did not made up antigenic distinct domains. According to Kolaskar and Tangaonkar antigenicity scale both HA proteins showed similar antigenic determinants with average propensity 1.021 for 1997-2005 isolates and 1.019 for other isolates (Figure 3). To evaluate the impact of amino acid changes on HA structure and stability the QMEAN score was used. The local geometry analysis based on a torsion angle potential, atoms distance-dependent interaction potential and the burial status of the residues (solvation free energy) were summarized in table 1. The scoring function terms were not significantly varied among China isolates, whereas a little change was shown in 2006 Middle Eastern isolates which increased in recent isolates. The QMEAN score as an estimate of protein stability did not differ considerably between the isolates and ranged between 0-1.

Because the DINRTFKPL located adjacent to the left-edge of the receptor binding pocket of HA we focused on the impact of different features of the motif on HLA-A\*0201 peptide binding properties (Table 2). According to the calculated aliphatic and the instability indexes the motifs described as stable peptides. The

negative GRAVY values showing the motifs as hydrophilic in nature and predicting most of the residues to be present in the surface. The DINRVFKPL motif tended to be slightly basic (pI 7.04) due to the replacement of T to V and other motifs showed very high basic pI (10.27). The results indicated that the HLA-binding peptide sequence dynamics were not undergone change during the past decades.

## Discussion

The H9N2 virus characterized as the mild disease in human and animal species. The sporadic human infection cases have been reported (6, 7). Phylogenetic studies revealed that the majority of H9N2 viruses that have been sequenced belong either to the G1 clade or the Y280/G9 clade. During the last two decades, several distinct lineages and sublineages of H9N2 viruses have become established and the viruses classified in at least seven genotypes series A-G. Among them G1-like viruses (genotypes A1–A3, A5, and A9) have been dominated in Middle east whereas, the genotypic complexity of the viruses shown in China (27, 28). Reassortment between H9N2 and wild bird viruses could trigger adaptation to other hosts. It seems that H9N2 control might allow officials to prevent the next event. In this study, to investigate whether the evolution of H9N2 viruses coupled with emergence of antigenic variants, potential CTL epitope sequences were predicted for HA protein. At least twelve conserved epitope motifs between them, particularly in the HA2 polypeptide were scored. The degree of sequence diversity between the Middle East and China viruses is great in the HA1. More additional motifs are found in the region in recent Chinese isolates. The coexistence of different lineages in the same susceptible population in the country is likely to generate appropriate conditions for the emergence of novel reassortant strains. As shown in figure 1 these emerging viruses by late 2005 have not replaced the previous circulating strains and are instead co-circulating with them. The Middle East H9N2 viruses phylogenetically grouped together within the

G1 lineage with a high identity (18, 28). It is expected that the region has not been exposed by novel genotype. Otherwise, Sun and coworkers (29) demonstrated that H9N2 in China belong to the other lineages than G1 evolved into distinct antigenic groups C, D, and E. The antigenic groups D and E isolated from wild waterfowl species possessed greater numbers of amino acid substitutions in antigenic sites of HA protein compared to the current vaccine strain and the antigenic group C virus. It seems that extensive circulation of the viruses has resulted in a progressive and constant genetic evolution and emergence of antigenically distinct viruses. Different distribution of the epitopes among the H9N2 isolates may reveal variation within the antigenic determinants. So, we focused on the effect of changes in HLA-restricted epitope on structure and antigenic function of HA protein in H9N2 isolates. Protein structure and antigenicity prediction analysis indicated that amino acid substitution in HA of the Middle East isolates did not affect viral recognition by host-specific CTL response. The claim is confirmed by QMEAN scoring describes the major geometrical aspects of protein structures using the torsion angles and solvation energies. The terms play an important role in flexibility required for folding of the polypeptide backbone and protein–ligand binding (26). According to our analysis, the HA protein of H9N2 viruses exhibited a score between normal values indicate that the protein structure and its antigenic function in recent viruses were not differing from previous isolates.

Based on our epitope distribution analysis, the altered amino acid were not representing within HA receptor binding pocket. It has been shown that the considerable amino acid mutations involve in evading of influenza viruses from immune system response if the changed epitopes is present in the globular head of HA1, in or adjacent to CTL epitopes (30, 31). The high incidence of amino acid substitutions in HA as the main immunogenic protein was not surprising. Based on the studies on human influenza A virus a number of amino acid substitutions in epitopes of

internal proteins have been associated with escape from CTL (32, 33).

## Conclusion

In this study the antigenic CTL epitopes within HA proteins of H9N2 viruses recognized by HLA-A\*0201 to plan the development of a vaccine against emerging avian influenza epidemic has identified. These data may provide valuable insights into potentially pre-existing immunity to H9N2 viruses in human and emphasize the importance of studying the evolutionary changes occurring in the viruses during their emergence to now.

## References

- Alexander PE, De P, Rave S. Is H9N2 avian influenza virus a pandemic potential? *Can J Infect Dis Med Microbiol.* 2009;20:e35.
- Brown I, Banks J, Manvell R, Essen S, Shell W, Slomka M, et al. Recent epidemiology and ecology of influenza A viruses in avian species in Europe and the Middle East. *Dev Biol.* 2005;124:45-50.
- Park KJ, Kwon H-i, Song M-S, Pascua PNQ, Baek YH, Lee JH, et al. Rapid evolution of low-pathogenic H9N2 avian influenza viruses following poultry vaccination programmes. *J Gen Virol.* 2011;92:36-50.
- Choi Y, Ozaki H, Webby R, Webster R, Peiris J, Poon L, et al. Continuing evolution of H9N2 influenza viruses in Southeastern China. *J Virol.* 2004;78:8609-14.
- Matrosovich MN, Krauss S, Webster RG. H9N2 influenza A viruses from poultry in Asia have human virus-like receptor specificity. *Virology.* 2001;281:156-162.
- Anonymous Influenza. Hong Kong Special Administrative Region of China. WHO Weekly. 1999;Contract No.: 111.
- Butt K, Smith GJ, Chen H, Zhang L, Leung YC, Xu K, et al. Human infection with an avian H9N2 influenza A virus in Hong Kong in 2003. *J Clin Microbiol.* 2005;43:5760-7.
- Feng Y, Mao H, Xu C, Jiang J, Chen Y, Yan J, et al. Origin and Characteristics of Internal Genes Affect Infectivity of the Novel Avian-Origin Influenza A (H7N9) Virus. *PLoS ONE.* 2013;8:e81136.
- Quiñones-Parra S, Grant E, Loh L, Nguyen TH, Campbell K-A, Tong SY, et al. Preexisting CD8+ T-cell immunity to the H7N9 influenza A virus varies across ethnicities. *PNAS.* 2014;111:1049-54.
- Stewart-Jones GB, McMichael AJ, Bell JI, Stuart DI, Jones EY. A structural basis for immunodominant human T cell receptor recognition. *Nature Immunol.* 2003;4:657-63.
- Boon A, de Mutsert G, Graus Y, Fouchier R, Sintnicolaas K, Osterhaus A, et al. The magnitude and specificity of influenza A virus-specific cytotoxic T-lymphocyte responses in humans is related to HLA-A and-B phenotype. *J Virol.* 2002;76:582-90.
- Gotch F, McMichael A, Smith G, Moss B. Identification of viral molecules recognized by influenza-specific human cytotoxic T lymphocytes. *J Exp Med.* 1987;165:408-16.
- Lu X, Renshaw M, Tumpey TM, Kelly GD, Hu-Primmer J, Katz JM. Immunity to influenza A H9N2 viruses induced by infection and vaccination. *J Virol.* 2001;75:4896-901.
- Wang TT, Palese P. Universal epitopes of influenza virus hemagglutinins? *Nature Struct Mol Biol.* 2009;16:233-4.
- Ekiert DC, Bhabha G, Elsliger M-A, Friesen RH, Jongeneelen M, Throsby M, et al. Antibody recognition of a highly conserved influenza virus epitope. *Science.* 2009;324:246-51.
- El Bakkouri K, Descamps F, De Filette M, Smet A, Festjens E, Birkett A, et al. Universal vaccine based on ectodomain of matrix protein 2 of influenza A: Fc receptors and alveolar macrophages mediate protection. *J Immunol.* 2011;186:1022-31.
- Bi Y, Lu L, Li J, Yin Y, Zhang Y, Gao H, et al. Novel genetic reassortants in H9N2 influenza A viruses and their diverse pathogenicity to mice. *Virol J.* 2011;8:505.
- Shahsavandi S, Salmanian A-H, Ghorashi SA, Masoudi S, Ebrahimi MM. Evolutionary characterization of hemagglutinin gene of H9N2 influenza viruses isolated from Asia. *Res Vet Sci.* 2012;93:234-9.
- World Health Organization. Antigenic and genetic characteristics of zoonotic influenza viruses and development of candidate vaccine viruses for pandemic preparedness. 2013.
- Tamura K, Dudley J, Nei M, Kumar S. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol.* 2007;24:1596-9.
- Zeeberg B. Shannon information theoretic computation of synonymous codon usage biases in coding regions of human and mouse genomes. *Genome Res.* 2002;12:944-55.

22. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser.* 1999;41:95-8.
23. Parker KC, Bednarek MA, Coligan JE. Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains. *J Immunol.* 1994;152:163-75.
24. Rammensee H-G, Bachmann J, Emmerich NPN, Bachor OA, Stevanović S. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics.* 1999;50:213-9.
25. Kolaskar A, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett.* 1990;276:172-4.
26. Benkert P, Tosatto SC, Schomburg D. QMEAN: A comprehensive scoring function for model quality assessment. *Proteins.* 2008;71:261-77.
27. Li K, Xu K, Peiris J, Poon L, Yu K, Yuen K, et al. Characterization of H9 subtype influenza viruses from the ducks of southern China: a candidate for the next influenza pandemic in humans? *J Virol.* 2003;77:6988-94.
28. Dong G, Luo J, Zhang H, Wang C, Duan M, Deliberto TJ, et al. Phylogenetic diversity and genotypical complexity of H9N2 influenza A viruses revealed by genomic sequence analysis. *PLoS ONE.* 2011;6:e17212.
29. Sun Y, Pu J, Fan L, Sun H, Wang J, Zhang Y, et al. Evaluation of the protective efficacy of a commercial vaccine against different antigenic groups of H9N2 influenza viruses in chickens. *Vet Microb.* 2012;156:193-9.
30. Steel J, Lowen AC, Wang TT, Yondola M, Gao Q, Haye K, et al. Influenza virus vaccine based on the conserved hemagglutinin stalk domain. *MBio.* 2010;1:e00018-10.
31. Kaverin NV, Rudneva IA, Ilyushina NA, Lipatov AS, Krauss S, Webster RG. Structural differences among hemagglutinins of influenza A virus subtypes are reflected in their antigenic architecture: analysis of H9 escape mutants. *J Virol.* 2004;78:240-9.
32. Oldstone M. How viruses escape from cytotoxic T lymphocytes: molecular parameters and players. *Virology.* 1997;234:179-85.
33. Rimmelzwaan G, Berkhoff E, Nieuwkoop N, Smith D, Fouchier R, Osterhaus A. Full restoration of viral fitness by multiple compensatory co-mutations in the nucleoprotein of influenza A virus cytotoxic T-lymphocyte escape mutants. *J Gen Virol.* 2005;86:1801-5.