Emerging Problems in Infectious Diseases

Comparative sequence, antigenic and phylogenetic analysis of avian influenza (H9N2) surface proteins isolated in Pakistan between 1999 and 2008

Azeem Mehmood Butt[1], Samerene Siddique[1], Shifa Tahir[2], Izza Nasrullah[2], Mureed Hussain[1], Muhammad Idrees[1], Jun Lu[3], Yigang Tong[4]

[1]*National Centre of Excellence in Molecular Biology (CEMB), University of the Punjab, Lahore, Pakistan*
[2]*Department of Bioinformatics and Biotechnology, International Islamic University, Islamabad, Pakistan*
[3]*Tumor Biotherapy Ward, Beijing YouAn Hospital, Capital Medical University, Beijing, PR China*
[4]*State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology, Beijing, PR China*

Abstract

Introduction: Influenza A viruses possess a unique genomic structure which leads to genetic instability, especially in products of neuraminidase and hemagglutinin genes. These surface proteins play major roles in viral entry and release, and in the activation of the host immune system.

Methodology: This study involved an *in silico* sequence, phylogenetic and antigenic analyses of hemagglutinin and neuraminidase proteins of avian influenza A (H9N2) strains that circulated in Pakistan's poultry flocks from 1999 to 2008 and determined variations among these sequences at different levels.

Results: Sequence and phylogenetic analysis revealed a large number of similar substitution mutations and close evolutionary relation among sequences of both proteins. Changes were observed in the N-glycosylation sites of both surface proteins, along with the appearance of a new glycosylation site in the neuraminidase sequence isolated in 2007. Epitopes for hemagglutinin remained conserved, whereas for neuraminidase, epitopes from older strains reappeared in present sequences.

Conclusions: Because of the rapid mutating nature of avian influenza subtype H9N2, constant surveillance of annual sequence variations is important. Preventive measures and vaccine products can be evaluated by keeping track of changes that may lead to reassortment among different circulating strains in Pakistan's commercial poultry flocks or in humans.

Key words: Avian Influenza (H9N2); Pakistan poultry flocks; phylogenetic, glycosylation and antigen; sequence comparison

J Infect Dev Ctries 2011; 5(6):413-424.

(Received 24 July 2010 – Accepted 24 November 2010)

Introduction

Influenza viruses are members of the orthomyxoviradae family of viruses, separated into three types; A, B and C. Type A influenza viruses can cause infection in a diverse range of hosts including birds and mammals, whereas type B influenza viruses are only known to affect humans [1]. Influenza type C viruses cause mild illness in humans and do not cause epidemics or pandemics. Both Influenza virus type A and B contain eight single-stranded RNA segments with negative polarity whereas type C, which comprises only seven single-stranded RNA segments that lack one of the envelope glycoproteins. These RNA segments encode ten proteins which include two surface glycoproteins, namely hemagglutinin (HA) and neuraminidase (NA), along with nucleoproteins (NP), three polymerase proteins (PA, PB1, PB2), and two matrix (M1, M2) and non-structural proteins (NS1, NS2) [2-6]. Avian influenza is caused by type A influenza virus and, based on the activity of the two surface glycoproteins; HA and NA [2,5,6]. There are 16 HA and 9 NA subtypes responsible for causing infection. HA plays an essential role in the early stages of infection and is responsible for the virus binding to its receptor, sialic acid, which is present on the host cell surface and promotes fusion of viral and endosomal membranes and eventually facilitates viral entry into the host cell [5]. The NA surface glycoprotein of influenza viruses prevents virus aggregation by cleaving the α-ketosodic linkage between sialic acid and adjacent sugar residue. This

Table 1. GenBank accession numbers of avian influenza A (H9N2) viral proteins, hemagglutinin and neuraminidase isolated from different regions of Pakistan between 1999 and 2008

/Proteins Year/	Hemagglutinin (HA)	Neuraminidase (NA)
1999	**CAC19694** A/chicken/Pakistan/2/99(H9N2)	**CAC19695** A/chicken/Pakistan/2/99(H9N2)
2005	**ACP50642** A/chicken/Pakistan/UDL-01/2005(H9N2)	**ACP50644** A/chicken/Pakistan/UDL-01/2005(H9N2)
	ACP50686 A/chicken/Pakistan/UDL-03/2005(H9N2)	**ACP50688** A/chicken/Pakistan/UDL-03/2005(H9N2)
	ACP50653 A/chicken/Pakistan/UDL-02/2005(H9N2)	**ACP50655** A/chicken/Pakistan/UDL-02/2005(H9N2)
2006	**ACP50664** A/chicken/Pakistan/UDL-01/2006(H9N2)	**ACP50666** A/chicken/Pakistan/UDL-01/2006(H9N2)
	ACP50631 A/chicken/Pakistan/UDL-04/2006(H9N2)	**ACP50633** A/chicken/Pakistan/UDL-04/2006(H9N2)
	ACP50675 A/chicken/Pakistan/UDL-02/2006(H9N2)	**ACP50677** A/chicken/Pakistan/UDL-02/2006(H9N2)
2007	**ACP50741** A/chicken/Pakistan/UDL-04/2007(H9N2)	**ACP50743** A/chicken/Pakistan/UDL-04/2007(H9N2)
	ACP50620 A/chicken/Pakistan/UDL-01/2007(H9N2)	**ACP50622** A/chicken/Pakistan/UDL-01/2007(H9N2)
	ACP50697 A/chicken/Pakistan/UDL-03/2007(H9N2)	**ACP50699** A/chicken/Pakistan/UDL-03/2007(H9N2)
2008	**ACP50719** A/chicken/Pakistan/UDL-02/2008(H9N2)	**ACP50710** A/chicken/Pakistan/UDL-01/2008(H9N2)
	ACP50730 A/chicken/Pakistan/UDL-03/2008(H9N2)	**ACP50721** A/chicken/Pakistan/UDL-02/2008(H9N2)
	ACP50708 A/chicken/Pakistan/UDL-01/2008(H9N2)	**ACP50732** A/chicken/Pakistan/UDL-03/2008(H9N2)

results in destruction of receptors recognized by HA and facilitates the to and fro movement of the virus from the site of infection [7].

Avian influenza virus (AIV) H9N2 is known to infect poultry populations throughout the world, are derived from the Eurasian and the North American influenza virus genes [2,8]. During 1994 to 1999, poultry populations in different countries, including Germany, Italy, Ireland, Iran, Pakistan, Saudi Arabia, South Africa and the USA were been found to be infected by H9N2 virus [9,10]. The social behavior and migratory routines of avian species is the reason why birds are the major hosts to these viruses rather than mammals and humans. This is more significant in the case of domestic poultry where there is extensive direct and indirect contact among the flocks and other living species [11].

In chickens, H9N2 virus infections cause mild respiratory signs and loss of egg production [12-17]. The first case of AIV (H9N2) in poultry flocks in Pakistan was reported in 1998. Genome analysis of the virus isolated from Pakistan during 1999 (A/chicken/Pakistan/2/99) showed close similarity to viruses isolated from children in Hong Kong in 1997, and were phylogenetically grouped together within the Eurasian G1-lineage [14,16,18-20].

Human infections with H9N2 virus were first reported during 1999 when two children in Hong Kong with mild upper respiratory tract infections tested positive for H9N2 virus [17,21]. In 2003, a five-year-old child in Hong Kong was again confirmed to have H9N2 virus infection that was of purely avian origin [16]. Genetic studies of H9N2 viruses from Hong Kong live bird markets have shown the preferential binding of viruses to 2, 6-linked sialic acid, which is a human-like receptor [18]. These findings have proven the ability of H9N2 avian influenza viruses to emerge as new pandemic strains.

The aims of present study were: to carry out comparative sequence analyses to characterize and establish the phylogenetic relationships of Pakistani isolates with neighboring and Eurasian sublineages; to describe changes that occurred in avian (H9N2) viruses during 1999 – 2008; and to identify and predict epitopes for antigenic changes, variations in glycosylation sites, and comparison of circulating viruses on a yearly basis.

Methodology

Taxon sampling

For comparison among H9N2 viruses circulating from 1999 to 2008, we conducted a computational search of all available sequences of avian influenza Virus subtype H9N2 reported in poultry flocks in Pakistan. The search was performed with the National Centre for Biotechnology Information (NCBI) Flu Database [22] and sequences of NA and HA was retrieved and downloaded (Table 1).

Divergences in sequence patterns

The protein sequences of NA and HA from viral strains (Table 1) were aligned and compared with selected strains from neighboring countries and the Eurasian G1 sublineage reference strain by using multiple sequence alignment software ClustalW2 [23], to determine sequence similarities, variations, and phylogenetic relations.

Determination of H9N2 host binding and release factors

The attachment and release of viruses from their host cells exploit the phenomenon of glycosylation. To determine variations in the sites of viral attachment, we used an online server application, ScanProsite [24] to compare and identify N-glycosylation sites and determined whether there were any inter-strain differences or variations in previous years, which could have affected or changed the glycosylation sites.

Epitope analysis of HA and NA for antigenic variations

Epitope prediction was performed by the CTL (Cytotoxic T Lymphocyte) epitope prediction method [25]. The amino acid sequences of HA and NA proteins from selected strains were used for this step and the predicted antigenic sites from each sequence were then compared for composition similarities and differences.

Phylogenetic analysis and tree construction

Phylogenetic analysis of HA and NA protein sequences of avian influenza H9N2 viruses isolated from Pakistan between 1999 and 2008 was performed using MEGA 4.0.2 [26]. The sequences were first aligned by using a multiple sequence alignment tool, CLUSTALW2. Unrooted phylogenetic trees were constructed by the minimum evolutionary method. Internal branching probabilities were determined by bootstrap analysis of 1,000 replicates and are indicated by percentage value on each branch.

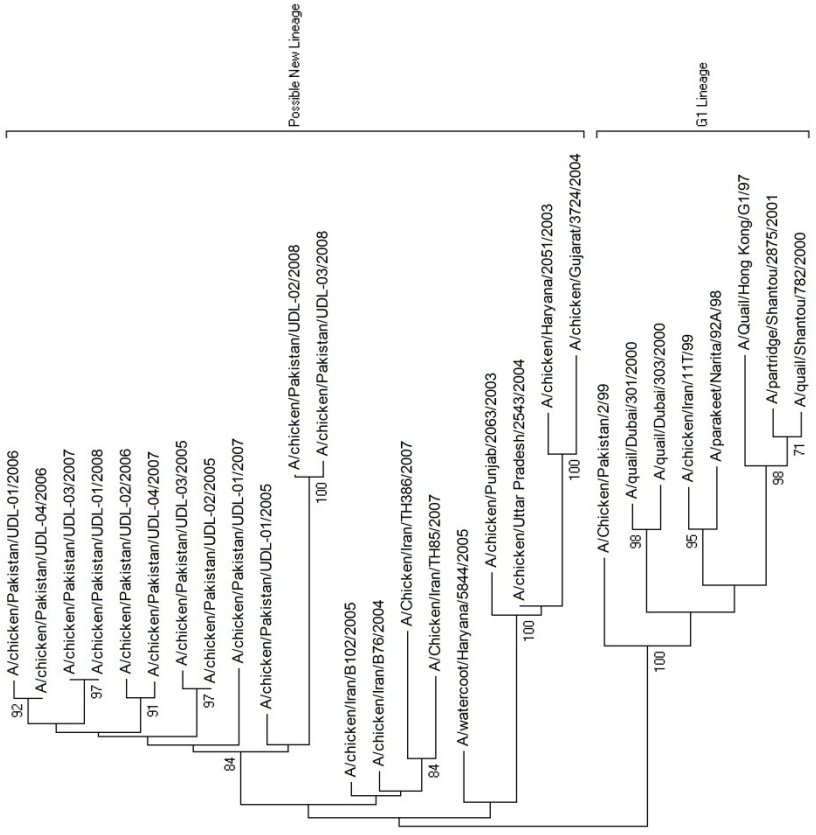
Results and discussion

Sequence and phylogenetic analysis

H9N2 viruses continue to circulate widely in domestic poultry in Asia [2,6,9,11]. The ability of H9N2 to transmit among humans poses a pandemic potential, which could arise by mutations in H9N2 itself or by reassortment between avian and human influenza viruses. Chances of this to occur are high in developing countries such as Pakistan and among people who are in direct contact with poultry flocks, such as farmers. Computational sequence analysis has the potential to be an effective approach towards vaccine development and measures against potential pandemics before they became a threat to poultry (H5N1) or human such as the pandemic 2009 Swine flu (H1N1). Therefore, the present study was performed to verify the extent of reassortment and antigenic shifts and drifts in surface proteins of the H9N2 avian virus isolated in Pakistan.

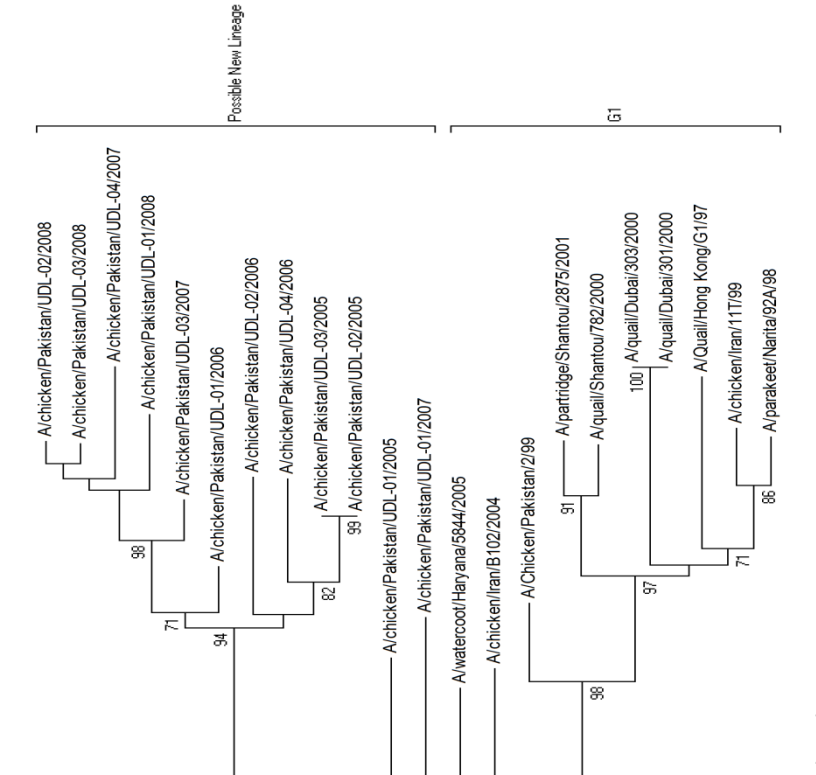
By studying the evolution of sequences, we tried to highlight how the selective pressure on the viral proteins changes with time, leading to changes in antigenicity and host specificity. A total of 13 protein sequences for each protein, NA and HA were

Fig. 1. Phylogenetic relationship of HA protein sequences of AIV (H9N2)



Unrooted phylogenetic tree of sequences isolated from Pakistan during 1999-2008 and its neighbouring countries including China, India, Iran, and also from Japan and Dubai. The tree was constructed by the Minimum evolution analysis, bootstrapped with 1,000 replicates using MEGA software version 4.0.2. Ck, chicken; Qa, quail; Pa, partridge; ST, Shantou.

Fig. 2. Phylogenetic relationship of NA protein sequences of AIV (H9N2)



Unrooted phylogenetic tree of sequences isolated from Pakistan during 1999-2008 and its neighbouring countries including China, India, Iran, and also from Japan and Dubai. The tree was constructed by the Minimum evolution analysis, bootstrapped with 1,000 replicates using MEGA software version 4.0.2. Ck, chicken; Qa, quail; Pa, partridge; ST, Shantou.

retrieved from the NCBI Flu Database by restricting the search query to Pakistan (Table 1). Sequence comparison was performed at two stages. First, full-length viral protein sequences of HA and NA isolated from poultry flocks in Pakistan (1999-2008) were compared with the Eurasian G1 sublineage reference strain (A/Quail/Hong Kong/G1/97), as shown in Tables 2 and 3 respectively. Sequence comparison showed the numbers of substitution mutations in both proteins. It has been observed that the H9N2 strain is mutating continuously, and the strains isolated during 2008 had many new mutations, which have not been found in previously reported entries in Pakistan. Second, amino acid variations in the HA receptor binding pocket and NA hemadsorbing site were speculated as shown in Table 4. Analysis of the HA receptor binding pocket in particular showed that, except for the isolate from 1999, all H9N2 isolates from Pakistan contained Leucine (L) at position 234, which has a preference for 2, 6-linked sialic acid (human receptors) instead of Glutamine (Q), which has a preference for 2, 3-linked sialic acid avian receptors. The role of this substitution mutation at position 234 (226; H3 numbering) has also been reported in other *in vivo* studies where avian H9N2 viruses showed replication with 100-fold higher peak titers in cultured human cells [27-30].

The evolutionary relationship of avian H9N2 virus HA and NA protein sequences was determined by analyzing selected isolates from Pakistan (Table 1) and representative H9N2 viruses from neighboring countries including China, India, Iran, Japan and Dubai, along with the established Eurasian H9N2 G1 lineage represented by its prototype strain (A/Hong Kong/G1/97). The unrooted phylogenetic tree of HA showed two distinct groups (Figure 1). G1 lineage reference virus and strain isolated from poultry flocks in Pakistan in 1999 clustered in one group, along with the strains from Dubai, Japan, and China. On the other hand, H9N2 avian virus isolates from Pakistan in 2005 - 2008 clustered together in one group along with H9N2 avian viruses isolated from neighboring countries Iran and India from 2003 to 2007. Clustering of all H9N2 Pakistani and neighboring viruses in one group postulates the emergence of a new lineage in the region of the subcontinent and Iran. This prediction could be related to another study in which it was postulated that a United Arab Emirates lineage of H9N2 viruses may have emerged [12]. In the case of NA sequences (Figure 2), the phylogenetic tree showed greater divergence. All Pakistani isolates from 2005 to 2008 formed one cluster, while the G1strain, viruses from Dubai, Japan, China, Iran and a single isolate from Pakistan in 1999 formed another cluster. It has been observed that the virus isolated from Pakistan in 1999 is closer to the G1 strain (97% identity) in comparison to other sequences isolated in later years.

Glycosylation and antigenic variations

Glycosylation of HA and NA represent the characteristic of the pathogen to escape from the host defense through co-evolution with the host and identification of the host receptor [31]. Glycosylation sites were predicted by ScanProsite for HA and NA viral proteins isolated from Pakistan only (1999-2008) as shown in Tables 5 and 6 respectively. A total of seven predicted glycosylation sites (PGS) (29-32; 105-108; 141-144; 298-301; 305-308; 492-495; 551-554) were obtained for each protein sequence of HA isolated during 2005-2008 except for a viral strain from 1999 which had an additional glycosylation site (218-221). This PGS was not found in other Pakistani isolates during sequence comparison and may represent the loss of a glycosylation site leading toward selected adaptation of avian H9N2 within poultry flocks. Results from ScanProsite revealed that amino acid variation E552G within the 2005 isolates resulted in alteration of the N-glycosylation site at positions 551-554 in all the later sequences. Two amino acid variations at position E32D (Glutamate to Aspartate) and Y144H (Tyrosine to Histidine) were found in all the 2008 isolates along with altered glycosylation sites, NSTD to NSTE (29-32) and NVTY to NVTH (141-144) respectively.

Table 2. Comparison of variations in the HA protein sequence from A/Hong Kong/G1/97 (G1 lineage) and AIV (H9N2) isolates from poultry flocks in Pakistan from 1999 to 2008. Red colour indicates substitution in sequences. HK, Hong Kong; PK, Pakistan

Positions from 1 till 204

Amino Acid Variations / Positions																																					
Viruses	3	4	5	7	15	17	32	42	46	58	63	66	71	79	90	101	113	121	122	126	127	138	144	145	147	149	150	153	158	166	168	171	183	187	191	196	198	204
A/Hong Kong/G1/97	T	I	S	I	A	N	E	V	H	M	S	H	D	V	E	S	V	T	L	A	S	T	Y	T	I	R	A	G	S	S	F	V	S	V	H	Y	E	I
A/Pakistan/2/99	T	R	S	I	A	N	E	V	H	M	N	H	D	I	G	S	V	A	I	A	S	T	Y	T	T	K	A	V	N	G	L	V	S	V	H	D	A	V
A/UDL-01/2005	T	I	S	M	T	N	E	V	Q	M	N	R	D	I	G	P	V	T	L	S	S	T	Y	T	T	K	S	N	N	N	L	V	D	V	H	D	A	T
A/UDL-03/2005	T	I	S	M	T	N	E	V	H	M	N	R	D	I	G	P	V	T	L	S	S	T	Y	T	T	K	S	D	N	N	L	T	D	V	H	D	A	T
A/UDL-02/2005	T	I	S	M	T	N	E	V	H	M	N	R	G	I	G	P	V	T	L	S	S	T	Y	T	T	K	S	D	N	N	L	T	D	V	H	D	A	T
A/UDL-01/2006	T	I	S	M	T	N	E	V	H	M	N	R	G	I	G	P	V	T	L	S	S	A	Y	T	T	K	S	D	N	N	L	I	D	V	H	D	A	T
A/UDL-04/2006	T	I	S	M	T	N	E	V	H	M	N	R	D	I	G	P	V	T	L	S	S	A	Y	T	T	K	S	D	N	N	L	I	D	V	H	D	A	T
A/UDL-02/2006	T	I	S	M	T	N	E	V	H	L	N	R	D	I	G	P	V	T	L	S	R	T	Y	T	T	K	S	D	N	N	L	I	D	V	H	D	A	T
A/UDL-04/2007	T	I	S	M	T	N	E	V	H	L	N	R	D	I	G	P	I	T	L	S	R	T	Y	T	T	K	S	D	N	N	L	I	D	V	H	D	A	T
A/UDL-01/2007	T	I	S	M	T	I	E	V	H	M	D	R	D	I	G	P	V	T	L	S	S	T	Y	T	T	K	S	D	N	S	L	I	D	V	H	D	A	T
A/UDL-03/2007	K	I	P	M	T	N	E	V	H	M	N	R	D	I	G	P	V	T	L	S	S	T	Y	T	T	K	S	D	N	N	L	I	D	V	H	D	A	T
A/UDL-01/2008	K	I	P	M	T	N	E	V	H	M	N	R	D	I	G	P	V	T	L	S	S	T	Y	T	T	K	S	D	N	N	L	I	D	V	H	D	A	T
A/UDL-02/2008	A	I	S	M	T	N	D	I	Q	M	N	R	D	I	G	P	V	T	L	A	S	A	H	D	T	K	S	N	N	N	L	I	D	I	H	D	A	T
A/UDL-03/2008	A	I	S	M	T	N	D	I	Q	M	N	R	D	I	G	P	V	T	L	A	S	A	D	D	T	K	S	N	N	N	L	I	D	I	H	D	A	T

Positions from 206 till 553

Amino Acid Variations / Positions																																	
Viruses	206	216	218	224	233	234	235	239	264	267	271	282	285	294	312	318	319	327	334	335	377	383	392	394	399	420	466	469	498	501	539	548	553
A/Hong Kong/G1/97	N	D	N	V	D	L	Q	D	Y	V	G	K	N	K	G	V	R	V	A	R	D	I	N	V	K	N	S	M	R	R	M	-	-
A/Pakistan/2/99	T	D	N	V	G	Q	Q	N	Y	V	G	N	N	K	G	V	R	I	A	R	D	I	N	V	K	N	S	M	R	K	M	A	G
A/UDL-01/2005	T	N	D	L	G	L	I	N	F	V	E	N	N	K	G	I	G	I	A	R	D	V	N	V	K	N	S	M	R	K	M	A	E
A/UDL-03/2005	T	N	D	L	G	L	I	N	F	V	E	N	N	K	R	I	G	I	A	R	D	V	N	V	K	N	T	M	R	K	I	A	G
A/UDL-02/2005	T	N	D	L	G	L	I	N	F	V	E	N	N	K	G	I	G	I	A	R	D	V	N	V	K	N	T	M	R	K	I	A	G
A/UDL-01/2006	T	N	D	L	G	L	I	N	F	V	E	N	N	K	G	I	G	I	A	R	D	V	N	V	K	N	S	V	R	K	M	A	G
A/UDL-04/2006	T	N	D	L	G	L	I	N	F	V	E	N	N	K	G	I	G	I	D	R	D	V	N	V	K	N	S	V	R	K	M	T	G
A/UDL-02/2006	T	N	D	L	G	L	I	N	F	V	E	N	D	K	G	I	G	I	A	R	D	V	N	V	K	N	S	M	R	K	M	A	G
A/UDL-04/2007	T	N	D	L	G	L	I	N	F	V	E	N	N	K	G	I	G	I	A	R	D	V	N	V	K	N	S	M	R	K	M	A	G
A/UDL-01/2007	T	N	D	L	G	L	I	N	F	I	E	K	N	K	G	I	G	I	A	R	D	V	N	V	K	N	S	M	K	K	I	A	G
A/UDL-03/2007	T	N	D	L	G	L	I	N	F	V	E	N	N	K	G	I	G	I	A	K	D	V	T	V	K	N	S	I	R	K	M	A	G
A/UDL-01/2008	T	N	D	L	G	L	I	N	F	V	E	N	N	K	G	I	G	I	A	K	D	V	T	V	K	N	S	I	R	K	M	A	G
A/UDL-02/2008	T	N	D	L	G	L	I	N	Y	V	E	K	N	R	G	I	G	I	A	K	G	V	N	I	R	D	S	M	R	T	M	A	G
A/UDL-03/2008	T	N	D	L	G	L	I	N	Y	V	E	K	N	R	G	I	G	I	A	K	G	V	N	I	R	D	S	V	R	T	M	A	G

PGSs for NA protein sequences also showed amino substitutions and variations. Eight glycosylation sites were obtained for each sequence except for one isolate from 2007, which possessed a new glycosylation site NESG at position 342-345. Two out of three strains from 2008 also showed a new modification at N-glycosylation site, 61-64 (NITK). Comparison with previously reported entries shows that these sites and positions are conserved, but in 2008 isolates, variations occurred with the substitution of a single amino acid (E64K; negatively charged polar to positively charged polar). It is important to note that all the above-mentioned substitutions in HA and NA protein sequences fell within the favorable region of mutations, which could still result in stable protein expression by retaining the hydrophobic / hydrophilic interactions and the 3D conformation of glycosylation sites. Therefore, detailed study of viral protein 3D structure, hydrophobicity and *in vivo* analysis would be useful for understanding possible outcomes of such sequential changes on the activity of viruses.

Results from CTLPred epitope prediction showed that HA protein sequences isolated from Pakistan tended to have amino acid Valine (V) at position 113. Nevertheless, in year 2007 one of the isolated strains had Isoleucine (I) at the same position. Another change was seen at position 138 where all previously isolated viruses had Threonine (T) at this position except entries from 2008 and two strains that were reported in 2006 (Table 7). It was also observed that in a strain from 1999 that an epitope was present at

Table 3: Comparison of variations in the NA protein sequence from A/Hong Kong/G1/97 (G1 lineage) and AIV (H9N2) isolates from Pakistan's poultry flocks from 1999 to 2008. Red colour shows substitution into sequences.

Positions from 1 till 155

	Amino Acid Variations / Positions																																				
viruses	6	7	13	15	20	24	33	36	38	39	42	45	47	48	49	50	52	54	57	59	61	63	64	65	70	72	81	82	83	86	98	105	106	125	127	152	154	155
A/Hong Kong/G1/97	K	I	V	I	I	M	M	H	-	-	C	P	N	N	Q	A	P	E	I	E	N	T	E	I	N	T	V	A	E	N	A	S	I	G	G	R	P	H
A/Pak/2/99	K	I	V	L	I	M	M	H	K	Q	C	P	N	N	Q	V	P	D	I	E	N	T	E	I	S	I	A	A	E	N	A	S	I	G	S	R	P	H
A/UDL-01/2005	R	I	A	L	I	I	M	H	K	Q	Y	S	K	N	Q	V	P	E	I	E	N	T	E	I	G	I	A	S	E	I	V	A	I	G	S	R	P	Y
A/UDL-03/2005	R	I	A	L	V	I	M	H	K	Q	Y	S	-	-	-	-	Q	D	I	E	N	T	E	V	G	I	A	A	E	N	V	S	I	G	S	R	S	Y
A/UDL-02/2005	R	I	A	L	V	I	M	H	K	Q	Y	S	-	-	-	-	Q	D	I	E	N	T	E	V	G	I	A	A	E	N	V	S	I	G	S	R	S	Y
A/UDL-01/2006	K	I	A	L	V	I	M	H	K	Q	Y	S	K	N	H	V	Q	D	V	E	N	T	E	I	G	I	A	A	E	N	V	S	I	G	S	R	P	Y
A/UDL-04/2006	R	I	A	L	V	I	M	H	K	Q	Y	S	K	N	N	V	Q	D	I	E	N	T	E	V	G	I	A	A	E	N	V	S	I	D	S	R	P	Y
A/UDL-02/2006	R	I	A	L	V	I	M	H	K	Q	Y	S	K	N	H	V	Q	D	I	E	N	T	E	V	G	I	A	A	E	N	V	S	I	G	S	R	P	Y
A/UDL-04/2007	K	I	A	L	V	I	M	L	S	Q	Y	S	K	N	H	V	Q	D	I	K	N	A	K	I	G	I	A	A	K	N	V	S	F	G	S	K	P	Y
A/UDL-01/2007	K	I	V	L	I	I	I	H	E	H	C	S	K	G	Q	V	P	G	I	E	N	T	E	I	G	I	A	A	E	N	V	S	I	G	S	R	P	Y
A/UDL-03/2007	K	I	A	L	V	I	M	H	N	Q	N	S	K	N	H	V	Q	D	I	E	D	T	E	I	G	I	A	A	E	N	V	S	I	G	S	R	P	Y
A/UDL-01/2008	K	M	A	L	V	I	M	L	N	Q	N	S	K	N	H	V	Q	D	I	K	N	T	E	I	G	I	A	A	K	N	V	S	I	G	S	R	P	Y
A/UDL-02/2008	K	I	A	L	V	I	M	L	N	Q	Y	S	K	N	H	V	Q	D	I	E	N	T	K	I	G	I	A	A	K	N	V	S	I	G	S	K	P	Y
A/UDL-03/2008	K	I	A	L	V	I	M	L	N	Q	Y	S	K	N	H	V	Q	D	I	K	N	T	K	I	G	I	A	A	K	N	V	S	I	G	S	R	P	Y

Positions from 155 till 468

	Amino Acid Variations / Positions																																						
viruses	161	168	187	194	199	208	211	220	221	236	238	249	253	261	264	302	308	311	312	317	328	344	346	356	372	376	380	381	385	392	402	403	406	412	414	416	435	468
A/Hong Kong/G1/97	N	H	K	V	R	D	L	K	N	T	T	R	K	K	H	V	A	S	V	V	N	R	P	N	S	T	V	G	T	I	I	R	S	A	G	I	R	-
A/Pak/2/99	S	H	R	V	R	D	L	K	S	T	T	R	R	K	H	V	A	S	I	V	N	R	P	N	T	I	V	G	M	I	N	W	Y	V	G	I	R	S
A/UDL-01/2005	S	H	R	V	R	D	L	K	N	T	A	R	R	K	H	V	A	S	I	V	S	R	P	N	A	T	V	G	I	I	N	W	Y	V	G	I	R	A
A/UDL-03/2005	S	H	R	I	K	D	L	K	N	T	T	R	R	R	H	I	A	N	I	V	S	R	P	N	A	T	I	G	M	I	N	R	Y	V	G	T	R	A
A/UDL-02/2005	S	H	R	I	K	D	L	K	N	T	T	R	R	R	H	I	A	N	I	V	S	R	P	N	A	T	I	G	M	I	N	R	Y	V	G	T	R	A
A/UDL-01/2006	S	H	R	I	R	D	L	K	N	T	T	R	R	R	H	V	A	N	I	V	S	R	P	N	A	A	I	G	M	I	N	R	Y	V	G	T	R	A
A/UDL-04/2006	S	Y	R	I	R	D	L	K	T	T	T	R	R	R	H	V	A	N	I	V	S	R	P	D	A	T	I	G	M	I	S	R	Y	V	G	T	R	A
A/UDL-02/2006	S	H	R	V	R	D	L	K	N	I	T	R	R	R	H	V	A	S	I	L	S	R	P	N	A	T	I	D	V	I	N	R	Y	V	S	T	K	A
A/UDL-04/2007	S	H	R	I	R	D	I	N	N	T	T	R	R	R	H	V	A	N	I	V	S	R	P	N	T	T	I	G	M	T	N	R	Y	V	G	T	R	A
A/UDL-01/2007	S	H	R	V	R	N	L	N	N	T	T	R	R	K	Y	V	E	S	I	V	N	S	S	N	A	T	I	D	V	I	N	R	Y	V	G	I	R	A
A/UDL-03/2007	S	H	R	I	R	D	I	N	N	T	T	M	R	R	H	V	E	N	I	V	S	R	P	N	A	T	I	G	M	I	N	R	Y	V	G	T	R	A
A/UDL-01/2008	S	N	R	I	R	D	I	N	N	T	T	M	R	R	H	V	E	N	I	V	S	R	P	N	A	T	I	G	M	I	N	R	Y	V	G	T	R	A
A/UDL-02/2008	S	H	R	I	R	D	I	N	N	T	T	M	R	R	H	V	E	N	I	V	S	R	P	N	A	T	I	G	M	I	N	R	Y	V	G	T	R	A
A/UDL-03/2008	S	H	R	I	R	D	I	N	N	T	T	M	R	R	H	V	E	N	I	V	S	R	P	N	A	T	I	G	M	I	N	R	Y	V	G	T	R	A

position 312 (GTCPKYVRV), while in the rest of the entries, epitope (KLAIGLRNV) was observed at position 324. Epitopes for NA (Table 8) showed a great deal of variations in a yearly manner. Observations revealed that epitope (SCHDGRAWL) from positions 182-190 that was present in viruses isolated in 1999, was absent in 2005 but then reappeared in 2006/05/02, 2007/12/17, 2007/03/23 and all the sequences from 2008. Another Epitope (SCYPRYPEV) from the position 279-287 has not been found in previously reported strains in Pakistan except in one from 1999 and another from 2007, but it now seems to be in a stable position in present strains. These results suggest that NA epitopes could

Table 4: Amino acid residues at the receptor binding sites of HA, and HB of NA protein from avian H9N2 viruses isolated from Pakistan, China, Dubai, India, Iran, and Japan. Residues differences and similarities as compared to G1 lineage reference (A/Quail/Hong Kong/G1/97) are indicated. Amino acid Leucine (L) at position 234 has preference for 2, 6-linked sialic acid (human receptors). A, Avian; Ck, Chicken; Qa, Quail; Pa, Partridge; ST, Shantou; Wc, Watercoot.

Viruses	GenBank ID		Hemagglutinin (HA) Receptor Binding Sites					Neuraminidase (NA) Hemadsorbing Sites		
	HA	**NA**	**191**	**198**	**234**	**235**	**236**	**372**	**402**	**403**
A/Qa/Hong Kong/G1/97	AAF00706	AAD49006	H	E	L	Q	G	S	I	R
A/Ck/Iran/B76/04	ABO09917	**Not available**	H	A	L	I	G	-	-	-
A/Ck/Iran/B102/05	ABO09919	ACD47110	H	A	L	I	G	A	N	W
A/Ck/Iran/TH386/07	ACA50027	Not Available	H	T	L	I	G	-	-	-
A/Ck/Iran/TH85/07	ACA50025	**Not available**	H	A	Q	I	G	-	-	-
A/Ck/Iran/11T/99	AAQ04847	AAQ04868	H	A	Q	Q	G	A	N	L
A/Pa/Narita/92A/98 (Japan)	BAB39512	BAB39516	H	E	M	Q	G	A	N	L
A/Qa/Dubai/301/00 (Middle East)	ABM21877	ABM21882	H	E	L	Q	G	A	N	W
A/Qa/Dubai/303/00 (Middle East)	ABM21875	ABM21884	H	A	Q	Q	G	A	N	W
A/Pa/ST/2875/01 (China)	ABV46327	ABV46329	H	E	L	Q	G	A	N	W
A/Qa/ST/782/00 (China)	ABM46228	ABM46082	H	E	L	Q	G	A	N	W
A/Ck/Haryana/2051/03 (India)	ACF93481	**Not Available**	H	A	L	I	G	-	-	-
A/Ck/Punjab/2063/03 (India)	ACF93482	**Not Available**	H	A	L	I	G	-	-	-
A/Ck/Uttar Pradesh/2543/04 (India)	ACF93483	**Not Available**	H	A	L	I	G	-	-	-
A/Ck/Gujrat/3724/04 (India)	ACF93484	**Not Available**	H	A	L	I	G	-	N	W
A/Wc/Haryana/5844/05 (India)	ACL79894	ACL79895	H	A	L	I	G	A	N	W
A/Ck/Pakistan/2/99	CAC19694	CAC19695	H	A	Q	Q	G	T	N	W
A/Ck/Pakistan/UDL-01/05	ACP50642	ACP50644	H	A	L	I	G	A	N	W
A/Ck/Pakistan/UDL-03/05	ACP50686	ACP50688	H	A	L	I	G	A	N	R
A/Ck/Pakistan/UDL-02/05	ACP50653	ACP50655	H	A	L	I	G	A	N	R
A/Ck/Pakistan/UDL-01/06	ACP50664	ACP50666	H	A	L	I	G	A	N	R
A/Ck/Pakistan/UDL-04/06	ACP50631	ACP50633	H	A	L	I	G	A	S	R
A/Ck/Pakistan/UDL-02/06	ACP50675	ACP50677	H	A	L	I	G	A	N	R
A/Ck/Pakistan/UDL-04/07	ACP50741	ACP50743	H	A	L	I	G	T	N	R
A/Ck/Pakistan/UDL-01/07	ACP50620	ACP50622	H	A	L	I	G	A	N	R
A/Ck/Pakistan/UDL-03/07	ACP50697	ACP50699	H	A	L	I	G	A	N	R
A/Ck/Pakistan/UDL-01/08	ACP50719	ACP50710	H	A	L	I	G	A	N	R
A/Ck/Pakistan/UDL-02/08	ACP50730	ACP50721	H	A	L	I	G	A	N	R
A/Ck/Pakistan/UDL-03/08	ACP50708	ACP50732	H	A	L	I	G	A	N	R

have evolved due to possible reassortment or antigenic shift.

Due to the small number of sequences available, our findings present preliminary results; however, they show interesting mutants. Greater sequencing/surveillance efforts are needed in the region.

Conclusion

High variation in amino acid sequences and reassortment phenomena of Influenza A viruses propose that although the H9N2 virus infection currently is not severe, it has a further pandemic potential. Co-infections during bouts of influenza might play a crucial role in the evolution of H9N2 in birds possessing highly virulent H5N1 virus and may cause the development of resistant viruses. Apart from viral factors, host factors might play an important role in the onset of new subtypes, for example, poultry flocks infected with H5N1 or H7N3, or workers infected with or previously exposed to H5N1 or 2009 pandemic H1N1 subtype. Collectively, our analysis highlights the need for focused studies on the evolution of avian influenza A (H9N2) by developing a continuous surveillance system for the effective management of viral

Table 5. Comparison of predicted N-glycosylation sites (PGS) within amino acid sequences of HA proteins isolated between 1999 and 2008 (H9N2). Red colour indicates the differences between N-glycosylation sites of isolates. A, Avian; Ck, Chicken; PK, Pakistan.

1999		2005		2006		2007		2008	
A/Ck/PK/2/99		A/Ck/PK/UDL-01/05		A/Ck/PK/UDL-01/06		A/Ck/PK/UDL-01/07		A/Ck/PK/UDL-01/08	
N-glycosylation sites		N-glycosylation sites		N-glycosylation sites		N-glycosylation sites		N-glycosylation sites	
29 - 32	NSTE	29 – 32	NSTE	29 - 32	NSTE	29 - 32	NSTE	29 - 32	NSTD
105 - 108	NGTC	105 – 108	NGTC	105 – 108	NGTC	105 – 108	NGTC	105 – 108	NGTC
141 - 144	NVTY	141 – 144	NVTY	141 – 144	NVTY	141 – 144	NVTY	141 – 144	NVTH
218 - 221	NRTF	-	-	-	-	-	-	-	-
298 - 301	NSTL	298 – 301	NSTL	298 – 301	NSTL	298 – 301	NSTL	298 – 301	NSTL
305 - 308	NISK	305 – 308	NISK	305 – 308	NISK	305 – 308	NISK	305 – 308	NISK
492 - 495	NGTY	492 – 495	NGTY	492 – 495	NGTY	492 – 495	NGTY	492 – 495	NGTY
551 - 554	NGSC	551 - 554	NESC	551 – 554	NGSC	551 – 554	NGSC	551 – 554	NGSC
		A/Ck/PK/UDL-02/05		A/Ck/PK/UDL-04/06		A/Ck/PK/UDL-04/07		A/Ck/PK/UDL-02/08	
		A/Ck/PK/UDL-03/05		A/Ck/PK/UDL-02/06		A/Ck/PK/UDL-03/07		A/Ck/PK/UDL-03/08	
		N-glycosylation sites		N-glycosylation sites		N-glycosylation sites		N-glycosylation sites	
		29 - 32	NSTE	29 - 32	NSTE	29 - 32	NSTE	29 - 32	NSTD
		105 – 108	NGTC	105 – 108	NGTC	105 – 108	NGTC	105 – 108	NGTC
		141 – 144	NVTY	141 – 144	NVTY	141 – 144	NVTY	141 – 144	NVTH
		-	-	-	-	-	-	-	-
		298 – 301	NSTL	298 – 301	NSTL	298 – 301	NSTL	298 – 301	NSTL
		305 – 308	NISK	305 – 308	NISK	305 – 308	NISK	305 – 308	NISK
		492 – 495	NGTY	492 – 495	NGTY	492 – 495	NGTY	492 – 495	NGTY
		551 – 554	NGSC	551 – 554	NGSC	551 – 554	NGSC	551 – 554	NGSC

Table 6. Comparison of predicted N-glycosylation sites within amino acid sequences of NA protein isolated between 1999 and 2008 (H9N2). Sequences with same glycosylation sites are grouped together. A, Avian; Ck, Chicken; PK, Pakistan.

1999		2005		2006		2007		2008	
A/Ck/PK/2/99		A/Ck/PK/UDL-01/05		A/Ck/PK/UDL-01/06		A/Ck/PK/UDL-01/07		A/Ck/PK/UDL-01/08	
N-glycosylation sites		N-glycosylation sites		N-glycosylation sites		N-glycosylation sites		N-glycosylation sites	
N/A		44 - 47	NSSK	44 – 47	NSSK	44 – 47	NSSK	44 – 47	NSSK
61 – 64	NITE	61 – 64	NITE	61 – 64	NITE	61 – 64	NITE	61 – 64	NITE
69 - 72	NSTI	69 – 72	NGTI	69 – 72	NGTI	69 – 72	NGTI	69 – 72	NGTI
86 – 89	NWSK	N/A		86 – 89	NWSK	86 – 89	NWSK	86 – 89	NWSK
146 – 149	NGTT	146 – 149	NGTT	146 – 149	NGTT	146 – 149	NGTT	146 – 149	NGTT
200 – 203	NATA	200 – 203	NATA	200 – 203	NATA	200 – 203	NATA	200 – 203	NATA
234 – 237	NGTC	234 – 237	NGTC	234 – 237	NGTC	234 – 237	NGTC	234 – 237	NGTC
N/A		N/A		N/A		342 – 345	NESG	N/A	
402 – 405	NWSG	402 – 405	NWSG	402 – 405	NRSG	402 – 405	NRSG	402 – 405	NRSG
		A/Ck/PK/UDL-02/05		A/Ck/PK/UDL-04/06		A/Ck/PK/UDL-04/07		A/Ck/PK/UDL-02/08	
		A/Ck/PK/UDL-03/05		A/Ck/PK/UDL-02/06		A/Ck/PK/UDL-03/07		A/Ck/PK/UDL-03/08	
		N-glycosylation sites		N-glycosylation sites		N-glycosylation sites		N-glycosylation sites	
		44 – 47	NSSV	44 – 47	NSSK	44 – 47	NSSK	44 - 47	NSSK
		57 – 60	NITE	61 – 64	NITE	N/A		61 – 64	NITK
		65 – 68	NGTI	69 – 72	NGTI	69 – 72	NGTI	69 – 72	NGTI
		82 – 85	NWSK	86 – 89	NWSK	86 - 89	NWSK	86 – 89	NWSK
		142 – 145	NGTT	146 – 149	NGTT	146 – 149	NGTT	146 – 149	NGTT
		196 – 199	NATA	200 – 203	NATA	200 – 203	NATA	200 – 203	NATA
		230 – 233	NGTC	234 - 237	NGTC	234 – 237	NGTC	234 – 237	NGTC
		398 - 401	NRSG			402 – 405	NRSG	402 – 405	NRSG

Table 7: Comparison of antigenic sites in the amino acid sequences of HA protein (H9N2) from 1999 to 2008. Red colour indicates amino acid differences within antigenic sites of isolates. Sequences isolated each year are grouped in columns. A, Avian; Ck, Chicken; PK, Pakistan.

1999		2005		2006		2007		2008	
A/Ck/PK/2/99		A/Ck/PK/UDL-01/05		A/Ck/PK/UDL-01/06		A/Ck/PK/UDL-01/07		A/Ck/PK/UDL-01/08	
Position	Sequence	Position	Sequence	Position	Sequence	Position	Sequence	Position	Sequence
431	AYNAELLVL	431	AYNAELLVL	431	AYNAELLVL	431	AYNAELLVL	431	AYNAELLVL
108	CYPGNVENL	108	CYPGNVENL	108	CYPGNVENL	108	CYPGNIENL	108	CYPGNVENL
312	GTCPKYVRV	134	IFPDTIWNV	134	IFPDAIWNV	134	IFPDTIWNV	134	IFPDAIWNV
134	IFPDTIWNV	154	SFYRNMRWL	154	SFYRNMRWL	154	SFYRNMRWL	154	SFYRNMRWL
154	SFYRNMRWL	324	KLAIGLRNV	324	KLAIGLRNV	324	KLAIGLRNV	324	KLAIGLRNV
		A/Ck/PK/UDL-02/05		A/Ck/PK/UDL-04/06		A/Ck/PK/UDL-04/07		A/Ck/PK/UDL-02/08	
		Position	Sequence	Position	Sequence	Position	Sequence	Position	Sequence
		431	AYNAELLVL	431	AYNAELLVL	431	AYNAELLVL	431	AYNAELLVL
		108	CYPGNVENL	108	CYPGNVENL	108	CYPGNVENL	108	CYPGNVENL
		134	IFPDTIWNV	134	IFPDAIWNV	134	IFPDTIWNV	134	IFPDAIWNV
		154	SFYRNMRWL	154	SFYRNMRWL	154	SFYRNMRWL	154	SFYRNMRWL
		324	KLAIGLRNV	324	KLAIGLRNV	324	KLAIGLRNV	324	KLAIGLRNV
		A/Ck/PK/UDL-03/05		A/Ck/PK/UDL-02/06		A/Ck/PK/UDL-03/07		A/Ck/PK/UDL-03/08	
		Position	Sequence	Position	Sequence	Position	Sequence	Position	Sequence
		431	AYNAELLVL	431	AYNAELLVL	431	AYNAELLVL	431	AYNAELLVL
		108	CYPGNVENL	108	CYPGNVENL	108	CYPGNVENL	108	CYPGNVENL
		134	IFPDTIWNV	134	IFPDTIWNV	134	IFPDTIWNV	134	IFPDTIWNV
		154	SFYRNMRWL	154	SFYRNMRWL	154	SFYRNMRWL	154	SFYRNMRWL
		324	KLAIGLRNV	324	KLAIGLRNV	324	KLAIGLRNV	324	KLAIGLRNV

Table: 8. Comparison of antigenic sites in the amino acid sequences of NA protein (H9N2) from 1999 to 2008. Red colour indicates amino acid differences within antigenic sites of isolates. Sequences isolated each year are grouped in columns. A, Avian; Ck, Chicken; PK, Pakistan.

1999		2005		2006		2007		2008	
A/Ck/PK/2/99		A/Ck/PK/UDL-01/05		A/Ck/PK/UDL-01/06		A/Ck/PK/UDL-01/07		A/Ck/PK/UDL-01/08	
Position	Sequence	Position	Sequence	Position	Sequence	Position	Sequence	Position	Sequence
100	FSKDNSIRL	100	FSKDNAIRL	100	FSKDNSIRL	297	GSNRPVLYI	100	FSKDNSIRL
297	GSNRPVLYI	297	GSNRPVLYI	297	GSNRPVLYI	461	GANINFMAI	297	GSNRPVLYI
248	GRADTRILF	2	NPNQRIIAL	461	GANINFMAI	248	GRADTRILF	461	GANINFMAI
182	SCHDGRAWL	461	GANINFMAI	248	GRADTRILF	100	FSKDNSFRL	2	NPNQKMIAL
279	SCYPRYPEV	248	GRADTRILF	182	SCHDGRAWL	182	SCHDGRAWL	182	SCHDGRAWL
		A/Ck/PK/UDL-02/05		A/Ck/PK/UDL-04/06		A/Ck/PK/UDL-04/07		A/Ck/PK/UDL-02/08	
		Position	Position	Position	Sequence	Position	Sequence	Position	Sequence
		96	FSKDNSIRL	100	FSKDNSIRL	100	FSKDNSIRL	100	FSKDNSIRL
		2	NPNQRIIAL	297	GSNRPVLYI	297	GSNRPVLYI	297	GSNRPVLYI
		457	GANINFMAI	2	NPNQRIIAL	461	GANINFMAI	461	GANINFMAI
		244	GRADTRILF	461	GANINFMAI	248	GRADTRILF	182	SCHDGRAWL
		178	SCHDGRAWL	248	GRADTRILF	257	IREGKIVYI	279	SCYPRYPEV
		A/Ck/PK/UDL-03/05		A/Ck/PK/UDL-02/06		A/Ck/PK/UDL-03/07		A/Ck/PK/UDL-03/08	
		Position	Position	Position	Sequence	Position	Sequence	Position	Sequence
		96	FSKDNSIRL	100	FSKDNSIRL	100	FSKDNSIRL	100	FSKDNSIRL
		2	NPNQRIIAL	297	GSNRPVLYI	297	GSNRPVLYI	297	GSNRPVLYI
		457	GANINFMAI	2	NPNQRIIAL	461	GANINFMAI	461	GANINFMAI
		244	GRADTRILF	461	GANINFMAI	182	SCHDGRAWL	182	SCHDGRAWL
		178	SCHDGRAWL	248	GRADTRILF	279	SCYPRYPEV	279	SCYPRYPEV

epidemics, particularly in developing countries such as Pakistan. There is a strong possibility that many of the cases are still unreported and no new data is yet available for 2009 – 2010. Variations in the sequence, glycosylation and epitopes of NA are needed, which might be an indication of virus activity leading to complete replacement of NA segments. Computational and experimental studies of currently reported NA segments from Influenza A viruses, especially NA segments from H5N1, H1N1 and H7N3, against N2 will be useful in determining possible reassortment or exchange of segments. Our findings demonstrate the instability and the potential of AIV (H9N2) circulating in poultry flocks in Pakistan to affect humans and can be used as a reference for further studies, which may involve *in vivo* studies and detailed 3D structure and function analysis of surface proteins thus facilitating the development of better treatment and prevention approaches.

Acknowledgements

We are thankful to the Director, National Centre of Excellence in Molecular Biology (CEMB) for encouragement and providing facilities to carry out this work.

References

1. Hiromoto Y, Saito T, Lindstrom SE, Li Y, Nerome R, Sugita S, Shinjoh M, Nerome K (2000) Phylogenetic analysis of the three polymerase genes (PB1, PB2 and PA) of influenza B virus. J Gen Virol 81: 929-937.
2. Fouchier RA, Munster V, Wallensten A, Bestebroer TM, Herfst S, Smith D, Rimmelzwaan GF, Olsen B, Osterhaus AD (2005) Characterization of a novel influenza a virus hemagglutinin subtype (h16) obtained from black-headed gulls. J Virol 79: 2814-2822.
3. Kamal RP, Tosh C, Pattnaik B, Behera P, Nagarajan S, Gounalan S, Shrivastava N, Shankar BP, Pradhan HK (2007) Analysis of the pb2 gene reveals that Indian h5n1 influenza virus belongs to a mixed-migratory bird sub-lineage possessing the amino acid lysine at position 627 of the pb2 protein. Arch Virol 152:1637-1644.
4. Suzuki Y and Nei M (2002) Origin and evolution of influenza virus hemagglutinin genes. Mol Biol Evol 19: 501-509.
5. Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y (1992) Evolution and ecology of influenza A viruses. Microbiol Rev 56: 152-179.
6. A revision of the system of nomenclature for influenza viruses: a WHO memorandum. (1980)Bull World Health Organ 58: 585-591.
7. Gubareva LV, Kaiser L, Hayden FG (2000) Influenza virus neuraminidase inhibitors. Lancet 355: 827-835.
8. Guo YJ, Krauss S, Senne DA, Mo IP, Lo KS, Xiong XP, Norwood M, Shortridge KF, Webster RG, Guan Y (2000) Characterization of the pathogenicity of members of the newly established H9N2 influenza virus lineages in Asia. Virology 267: 279-288.
9. Alexander DJ (2000) A review of avian influenza in different bird species. Vet Microbiol 74: 3-13.
10. Alexander DJ (2003) Avian Influenza in the Eastern Hemisphere (Excluding the Pacific Basin) during 1992-1997. Avian Diseases 47: 9-13.
11. Stallknecht DE, Shane SM (1988) Host range of avian influenza virus in free-living birds. Vet Res Commun 12:125-141.
12. Bano S, Naeem K, Malik SA (2003) Evaluation of pathogenic potential of avian influenza virus serotype H9N2 in chickens. Avian Dis 47: 817-822.
13. Butt KM, Smith GJ, Chen H, Zhang LJ, Leung YH, Xu KM, Lim W, Webster RG, Yuen KY, Peiris JS, Guan Y (2005) Human infection with an avian H9N2 influenza A virus in Hong Kong in 2003. J Clin Microbiol 43: 5760-5767.
14. Iqbal M, Yaqub T, Reddy K, McCauley JW (2009) Novel genotypes of H9N2 influenza A viruses isolated from poultry in Pakistan containing NS genes similar to highly pathogenic H7N3 and H5N1 viruses. PLoS One 4:e5788.
15. Kishida N, Sakoda Y, Eto M, Sunaga Y, Kida H (2004) Co-infection of Staphylococcus aureus or Haemophilus paragallinarum exacerbates H9N2 influenza A virus infection in chickens. Arch Virol 149: 2095-2104.
16. Lin YP, Shaw M, Gregory V, Cameron K, Lim W, Klimov A, Subbarao K, Guan Y, Krauss S, Shortridge K, Webster R, Cox N, Hay A (2000) Avian-to-human transmission of H9N2 subtype influenza A viruses: relationship between H9N2 and H5N1 human isolates. Proc Natl Acad Sci USA 97: 9654-9658.
17. Lu X, Renshaw M, Tumpey TM, Kelly GD, Hu-Primmer J, Katz JM (2001) Immunity to influenza A H9N2 viruses induced by infection and vaccination. J Virol 75: 4896-4901.
18. Aamir UB, Wernery U, Ilyushina N, Webster RG (2007) Characterization of avian H9N2 influenza viruses from United Arab Emirates 2000 to 2003. Virology 361: 45-55.
19. Cameron KR, Gregory V, Banks J, Brown IH, Alexander DJ, Hay AJ, Lin YP (2000) H9N2 subtype influenza A viruses in poultry in Pakistan are closely related to the H9N2 viruses responsible for human infection in Hong Kong. Virology 278: 36-41.
20. Naeem K, Ullah A, Manvell RJ, Alexander DJ (1999) Avian influenza A subtype H9N2 in poultry in Pakistan. Vet Rec 145: 560.
21. Tosh C, Nagarajan S, Behera P, Rajukumar K, Purohit K, Kamal RP, Murugkar HV, Gounalan S, Pattnaik B, Vanamayya PR, Pradhan HK, Dubey SC (2008) Genetic analysis of H9N2 avian influenza viruses isolated from India. Arch Virol 153: 1433-1439.
22. Bao Y, Bolotov P, Dernovoy D, Kiryutin B, Zaslavsky L, Tatusova T, Ostell J, Lipman D (2008) The Influenza Virus Resource at the National Center for Biotechnology Information. J Virol 82: 596-601.
23. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947-2948.
24. de Castro E, Sigrist CJ, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, Bairoch A, Hulo N (2006) ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res 34: W362-365.

25. Bhasin, M., Raghava, G. P. S. (2004) Prediction of ctl epitopes using qm, svm and ann techniques. Vaccine 22 (23-24): 3195-3204.
26. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24: 1596-1599.
27. Matrosovich MN, Matrosovich TY, Gray T, Roberts NA, Klenk HD (2004) Human and avian influenza viruses target different cell types in cultures of human airway epithelium. Proc Natl Acad Sci U S A 101:4620-4624
28. Matrosovich, M, Gambaryan, A, Klenk, H (2008) Receptor Specificity of Influenza Viruses and Its Alteration during Interspecies Transmission. In Avian Influenza Klenk H-D, Matrosovich MN, Stech J, eds. Monogr Virol. Basel, Karger, 2008, vol 27. pp 134-155.
29. Gamblin SJ, Haire LF, Russell RJ, Stevens DJ, Xiao B, Ha Y, Vasisht N, Steinhauer DA, Daniels RS, Elliot A, Wiley DC, Skehel JJ (2004) The structure and receptor binding properties of the 1918 influenza hemagglutinin. Science 303: 1838-1842.
30. Wan H, Perez DR (2007) Amino acid 226 in the hemagglutinin of H9N2 influenza viruses determines cell tropism and replication in human airway epithelial cells. J Virol 81: 5181-5191.
31. Igarashi M, Ito K, Kida H, Takada A (2008) Genetically destined potentials for N-linked glycosylation of influenza virus hemagglutinin. Virology 376: 323-329.

Corresponding Author

Yigang Tong, Ph.D
Professor
State Key Laboratory of Pathogen and Biosecurity,
Beijing Institute of Microbiology and Epidemiology,
Beijing, 100071, PR China
Email: tong.yigang@gmail.com

Conflict of interests: No conflict of interests is declared.